

GESELLSCHAFT
FÜR INFORMATIK



Volker Wohlgemuth, Stefan Naumann, Grit Behrens,
Malgorzata Burchard-Dziubinska (Editors)

EnviroInfo 2026
Short-/Work in Progress-Papers

September 09-11, 2026
Łódź- Poland

Gesellschaft für Informatik e.V. (GI)

Lecture Notes in Informatics (LNI) - Proceedings

Series of the Gesellschaft für Informatik (GI)

Volume P-385

ISSN 2944-7682

Volume Editors

Volker Wohlgemuth	HTW Berlin – University of Applied Sciences, Berlin, Germany
Stefan Naumann	Trier University of Applied Sciences, Trier, Germany
Grit Behrens	University of Applied Sciences Bielefeld, Minden, Germany
Małgorzata Burchard-Dziubińska	University of Łódź, Łódź, Poland

Series Editorial Board

Andreas Oberweis, KIT Karlsruhe, (Chairman, andreas.oberweis@kit.edu)
Torsten Brinda, Universität Duisburg-Essen, Germany
Dieter Fellner, Technische Universität Darmstadt, Germany
Ulrich Frank, Universität Duisburg-Essen, Germany
Barbara Hammer, Universität Bielefeld, Germany
Falk Schreiber, Universität Konstanz, Germany
Wolfgang Karl, KIT Karlsruhe, Germany
Michael Koch, Universität der Bundeswehr München, Germany
Heiko Roßnagel, Fraunhofer IAO Stuttgart, Germany
Kurt Schneider, Universität Hannover, Germany
Andreas Thor, HFT Leipzig, Germany
Ingo Timm, Universität Trier, Germany
Karin Vosseberg, Hochschule Bremerhaven, Germany
Maria Wimmer, Universität Koblenz-Landau, Germany

Dissertations

Rüdiger Reischuk, Universität Lübeck, Germany

Thematics

Agnes Koschmider, Universität Kiel, Germany

Seminars

Judith Michael, RWTH Aachen, Germany

© Gesellschaft für Informatik, Bonn 2025



This book is licensed under a Creative Commons BY-SA 4.0 license.

Preface

This book contains the short and work-in-progress contributions of the 40th International and Interdisciplinary Conference on Environmental Informatics – EnviroInfo 2026.

Held from 9 to 11 September 2026 at the University of Łódź, Poland, the conference was organized by the technical committee of Environmental Informatics of the Gesellschaft für Informatik e.V. (GI). It took place in conjunction with the international scientific conference "Sustainable Development and Other Paradigms and Concepts in the Context of Contemporary Civilization Trends – Global and Local Dimensions" (SD26) and the Annual Meeting of the Polish Association of Environmental Economists.

Over four decades, the EnviroInfo series has brought together researchers and practitioners from across the world who work at the intersection of applied computer science and environmental sustainability. The shared ambition has always been to translate advances in information and communication technologies into concrete steps toward a more sustainable society.

The theme of EnviroInfo 2026 – "Sustainable Development and Other Paradigms and Concepts in the Context of Contemporary Civilization Trends – Global and Local Dimensions" – opens up the conversation beyond any single framework. It situates environmental informatics within a wider debate about how societies, at both local and global scales, respond to the pressures of our time: through green computing and sustainable software, through circular economy models, through business intelligence, and through the critical examination of the paradigms that shape development itself.

The contributions gathered here are works in progress or early-stage results – deliberately so. This volume is a space for approaches that are still taking shape, for questions not yet fully answered, and for methodological experiments that connect disciplines and cross the boundary between research and practice. Taken together, they map a field in motion: from data infrastructures and environmental management systems to green coding, sustainability indicators, and the interfaces between science, policy, and civil society.

We thank all authors for sharing their work at this stage of development. We equally thank the members of the program and organization committee for their thorough and careful reviews. We thank the University of Łódź as host, the organizers of SD26 and the Polish Association of Environmental Economists as co-organizers, and the Gesellschaft für Informatik e.V. We also thank the European universities of the UNIC alliance (European University of Cities in Post-industrial Transition – UNIC 2.0), for their cooperation in making this conference possible.

Finally, we thank our sponsors, whose support was essential to making this conference happen.

Łódź, September 2026

On behalf of the Organizing Committee

Volker Wohlgemuth, Stefan Naumann, Grit Behrens, Malgorzata Burchard-Dziubińska

Sponsors

We want to thank the following companies for supporting the conference.

Hewlett Packard Enterprise



German Informatics Society



Disy Informationssysteme GmbH



General Chair

Volker Wohlgemuth	HTW Berlin – University of Applied Sciences, Berlin, Germany
Stefan Naumann	Trier University of Applied Sciences, Trier, Germany
Grit Behrens	University of Applied Sciences Bielefeld, Minden, Germany
Małgorzata Burchard-Dziubińska	University of Łódź, Łódź, Poland

Program Committee

Achim Guldner, Trier University of Applied Sciences, Germany
Amna Ramzy, German University in Cairo (GUC), Egypt
Andi H. Widok, HTW Berlin – University of Applied Sciences, Berlin, Germany
Andreas Abecker, Disy Informationssysteme GmbH, Germany
Andreas Winter, University of Oldenburg, Germany
Anna Zagorski, Umweltbundesamt - German Environment Agency, Germany
Barak Fishbain, Technion – Israel Institute of Technology, Israel
Benjamin Wagner vom Berg, University of Applied Sciences Bremerhaven, Germany
Benno Schmidt, Bochum University of Applied Sciences, Germany
Benoît Otjacques, Luxembourg Institute of Science and Technology (LIST), Luxembourg
Denis Havlik, Austrian Institute of Technology (AIT), Austria
Dennis Junger, HTW Berlin – University of Applied Sciences, Berlin, Germany
Dieter Kranzlmüller, Leibniz Supercomputing Centre (LRZ) of the Bavarian Academy of Sciences and Humanities, Germany
Frank Fuchs-Kittowski, HTW Berlin – University of Applied Sciences, Berlin, Germany
Friedhelm Hosenfeld, DigSyLand, Germany
Hans-Knud Arndt, Otto-von-Guericke-Universität, Magdeburg, Germany
Heiko Thimm, Hochschule Pforzheim, Germany
Horst Kremers, Horst Kremers UNDRR SEM - RIMMA CoE Community, Germany
Jochen Wittmann, HTW Berlin – University of Applied Sciences., Berlin, Germany
Jorge Marx Gomez, Carl von Ossietzky University of Oldenburg - Department of Computing Science, Germany
Karol Mrozik, Poznań University of Life Sciences, Poland
Klaus Greve, University of Bonn, Germany
Kostas Karatzas, Aristotle University of Thessaloniki, Greece
Margaret M. MacDonell, Argonne National Laboratory, USA
Martina Willenbacher, HTW Berlin – University of Applied Sciences., Berlin, Germany
Maximilian Höb, Leibniz Supercomputing Centre (LRZ) of the Bavarian Academy of Sciences and Humanities, Germany
Patricia Lago, Vrije Universiteit Amsterdam (VU), The Netherlands
Pawel Bartoszczuk, SGH Warsaw School of Economics, Poland
Ralf Klischewski, German University in Cairo (GUC), Egypt, Germany
Rüdiger Schaldach, University of Kassel - Center for Environmental Systems Research, Germany
Stanisław Czaja, Wrocław University of Economics and Business, Poland
Stefan Naumann, Trier University of Applied Sciences, Germany
Thomas Rose, RWTH Aachen University - Media Informatics / Processes, Germany

Organizing Committee

Dennis Junger	HTW Berlin – University of Applied Sciences, Berlin, Germany
Grit Behrens	University of Applied Sciences Bielefeld, Germany
Małgorzata Burchard-Dziubińska	University of Łódź, Łódź, Poland
Stefan Naumann	Trier University of Applied Sciences, Germany
Volker Wohlgemuth	HTW Berlin – University of Applied Sciences, Berlin, Germany

Contents

Sustainable development and other paradigms and concepts in the context of contemporary civilization trends – global and local dimensions

Friedhelm Hosenfeld, Kay Krüger

Web-based management and analysis of nature conservation measures in Schleswig-Holstein – current developments 15

Frank Fuchs-Kittowski, Maximilian Deharde, David Starzl, Marius Poppel, Simon Burkard

Common and Domain-Specific Design Patterns for Mobile AR Applications to Visualize Measures on Building Facades in an Environmental Context 25

Sandhya Koravi

Sustainability Context Inheritance: From Trusted Evidence to Product Sustainability Profiles in Circular Economy Decision-Making 35

Artificial Intelligence and Sustainability

Paweł Kobiela, Jarosław Stańczak, Paweł Bartoszczuk

AI model of a photovoltaic installation 47

Larissa Koch de Souza, Martin Engstler

Environmentally Informed GenAI Use in ITSM: A Problem Analysis of Organisational Use-Case Decision-Making 57

Nicole Werner, Heiko Thimm, Leonie Hammann

Application of Artificial Intelligence in German Municipal Wastewater Facilities 67

Doctoral Workshop

Engelbert Fellingner
Assessing the Maturity of Sustainability-aware Data Management Strategies that minimise Data Volumes and reduce Energy Consumption: A Study Approach 79

Sarah Flohr, Paul Bruck, Jan Kleinmann, Grit Behrens, Andreas Kahlfeld
Linking Water Quality to Corrosion Behaviour at Steel Sheet Pile Walls: A Machine Learning Study in the Port of Hamburg 89

Environmental Management Software

Łukasz Szalata, Grit Behrens, Cezary Orłowski, Rafał Grzeszczyk
Air Quality Knowledge as a Driver of Computing Competences: An International Project-Based Learning Project 101

Green Coding

Maximilian Jahns, Dennis Junger, Volker Wohlgemuth, Geerd-Dietger Hoffmann
Windows Integration into the Green Metrics Tool: Energy Measurement of Native Windows Applications via RAPL 115

Nick Kriesing Garcia, Dennis Junger, Volker Wohlgemuth
Joint Evaluation of the Green Coding Roadmap Training 125

Danila Valko, Jorge Marx Gómez
Toward a Framework for Performance and Carbon Footprint Assessment of Payment Channel Network Solutions 139

Max Westing, Stefan Naumann
Towards a Taxonomy of ICT and Software Sustainability Assessment Instruments 149

List of Authors

Sustainable development and other paradigms and
concepts in the context of contemporary civilization
trends – global and local dimensions

Web-based management and analysis of nature conservation measures in Schleswig-Holstein – current developments

Friedhelm Hosenfeld ¹, and Kay Krüger²

Abstract: The database for nature conservation measures (MDB) has been in operation since 2010 within the Environment Department of the State of Schleswig-Holstein as an internal web application for managing measures planned for nature conservation. This article describes recent key developments that have made the measures database application more user-friendly, more interoperable and better aligned with current requirements.

This includes a new module on the management of invasive alien species (IAS), which not only supports the implementation of EU requirements but also, beyond that, optimises the management of invasive species in Schleswig-Holstein.

Furthermore, the migration to the innovative Disy Cadenza Workbooks platform is discussed; whilst this platform offers comprehensive analysis and presentation options utilising linked dashboard displays, it also required the complete redevelopment of all existing analyses within the nature conservation measures database.

Keywords: nature conservation measures, invasive alien species, Disy Cadenza Workbooks, dashboard views, nature conservation management, OI DC

1 Introduction

The *nature conservation measures database* (MDB) described in this article was originally developed in 2010 as an internal web application for the nature conservation departments of the Ministry of the Environment (MEKUN) and State Agency for the Environment (LfU) in Schleswig-Holstein [Ho12].

The core application has been in use ever since, but has been continuously refined over the years, adapted to the latest technical requirements and optimised from a technical perspective.

A major change took place as early as 2019, when the application was migrated from the Oracle database system to PostgreSQL and operations were transferred from the LfU to the central twin data centre (TDC) of the state IT service provider Dataport. The switch to

¹ DigSyLand - Institute for Digital System Analysis & Landscape Diagnosis, Zum Dorfteich 6, D-24975 Husby, hosenfeld@digsyland.de, <https://orcid.org/0000-0002-0538-7684>

² MEKUN - Ministry for Energy Transition, Climate Protection, Environment and Nature Schleswig-Holstein, Mercatorstraße 3, D-24106 Kiel, kay.krueger@mekun.landsh.de

the open-source database PostgreSQL was carried out based on Schleswig-Holstein's open-source strategy [MP24].

This article discusses further extensive modernisations and enhancements that have recently been carried out or are currently being implemented, covering both technical aspects and subject-specific requirements:

- In terms of content, the measures database has been expanded to include a module for the management of invasive species, which not only covers the introduction of a new category for specific measures against invasive species but also takes invasive species into account within other categories of measures.
- As a major technical innovation – one which also opens up new technical possibilities in terms of analysis capabilities – a migration was carried out to the Disy Cadenza *Workbooks* analysis and GIS platform, having previously used the predecessor solution, Cadenza *Classic*.

Section 2 below outlines the current situation and the requirements for the latest developments, beginning with the introduction of the nature conservation measures database as a web application. The same section then sets out the rationale for the required enhancements and explains the requirements for both key areas of development.

Section 3 then examines the design and implementation of measures to manage invasive species. The final focus in Section 4 is on describing the migration from Cadenza Web Classic to Cadenza Workbooks, including the implementation of specialist analysis requirements in modern dashboards and the integration of the management application for data maintenance into Cadenza Workbooks. Finally, Section 5 provides a summary and outlook.

2 Background, motivation and requirements

2.1 Nature Conservation Measures Database (MDB) in Schleswig-Holstein

As the first implementation phase of a cross-cutting data management concept developed in 2009 for all nature conservation data within the environmental department, the nature conservation measures database (MDB) was developed in 2010 on behalf of what is now the Ministry for Energy Transition, Climate Protection, Environment and Nature (MEKUN), as until then no suitable tool for managing measures had been available within the state authorities.

From a technical standpoint, the MDB was implemented as a web application using the PHP scripting language (<https://www.php.net/>) and integrated into the Disy Cadenza analysis platform [DI26]. The PHP management application handles the recording, maintenance, and management of nature conservation measures, while analyses and map

visualizations were implemented in Cadenza Web. Furthermore, the GIS-based capture of measure geometries was one of the first applications realized in Cadenza Web using the Cadenza application interface [Ho12].

The application is used by over 80 staff members from the nature conservation departments of the LfU and MEKUN to manage currently more than 16,000 measures. Various categories of measures are distinguished – examples include species conservation measures, protection and development measures, and biotope creation measures. Access rights are determined based on two factors: the origin of the measure and the responsible authority (LfU or MEKUN).

The data is stored in a relational database management system (RDBMS) based on PostgreSQL and PostGIS at the central twin data centre of Dataport, the state's IT service provider (see Figure 1). This applies equally to two other key applications in the nature conservation sector: the Register of Protected Areas (PAR) and the Peatland Database, both of which are closely linked to the MDB. The Protected Area Register records all nature conservation-relevant protected areas, such as nature conservation areas, bird sanctuaries, and other Natura 2000 sites [Ho21]. The Peatland Database, in turn, not only records basic data on peatlands and specific peatland-related measures but also integrates all measures managed in the MDB that pertain to peatlands, thereby avoiding duplicate entries [Ho23] (see Figure 1). Both the PAR and the Peatland Database share the same technical architecture as the MDB, utilizing PHP data maintenance applications integrated into Cadenza.

2.2 Need for the management of invasive species

The importance of combating invasive species was recognized years ago and is increasingly gaining public attention. As early as 2008, the EU estimated invasive alien species were having a massive negative impact on the economy and the environment [EU09] which led in 2014 to EU Regulation 1143/2014 on the prevention and management of the introduction and spread of invasive alien species [EU14].

This regulation has two primary objectives:

- *prevention*: averting the introduction and spread of species that threaten biodiversity
- *control*: detecting such species that have already spread as early as possible, and subsequently eradicating them or, at the very least, monitoring and regulating them

In addition to collecting data on invasive species and implementing measures, the EU regulation requires Member States to report on species occurrences and the measures taken. The nature conservation departments of MEKUN and LfU worked on a technical concept deciding that data collection should take place within the existing nature conservation measures database in order to fulfill the needs of EU Regulation No. 1143/2014 and further local requirements concerning invasive alien species in Schleswig-Holstein.

2.3 Fundamentals of the Disy Cadenza Analysis and GIS Platform

Disy Cadenza is a comprehensive analysis platform offering a wide range of options for querying and visualizing non-spatial data, as well as map-based analyses. A particular strength lies in its functionality for the combined analysis and visualization of non-spatial and geospatial data; these features are tailored to the specific requirements of the environmental sector and are well-suited for the MDB (see also [DI26]). The universal analysis functions – which can be combined like building blocks and have proven highly effective in the field of environmental data – were a key reason for using Cadenza for the MDB.

Cadenza enables the seamless integration of dedicated management applications for data maintenance, while Cadenza itself focuses on providing read-only access for analysis purposes. Consequently, it serves as a modular system for designing analyses and visualizations (tables, charts, maps, reports, dashboards) based on standardized components, while a specialized management application tailored to the specific use case handles data maintenance – yet to the end-user, both components appear as a single, unified application (details on the current Cadenza Workbooks can be found in [JA26]).

2.4 Reasons and requirements for modernizing analyses and switching to Cadenza Workbooks

Cadenza Workbooks is designed to empower experts to customize dashboards and analysis views according to their needs, reducing their reliance on the centrally defined configurations typical of the earlier Cadenza Classic variant. Furthermore, the system's more flexible analysis and visualization options – particularly the ability to combine different views within dashboards – are driving the migration of existing applications to Cadenza Workbooks, like the MDB web application. While Cadenza Workbooks offers additional, useful analysis capabilities to optimize MDB data analysis, Cadenza Classic will receive support and further development for only a limited time; these were compelling reasons to switch to Cadenza Workbooks.

Analyses compiled in Cadenza for a specific business context are referred to as a "repository." The migration from Cadenza Classic to Cadenza Workbooks necessitated the complete reconstruction of all MDB related analyses and visualizations within the new environment due to the differing technical foundations. The new MDB workbooks repository was developed entirely by the responsible expert at MEKUN.

In contrast, the core MDB management application itself could continue to be used. Nevertheless, two significant adjustments were required to integrate the MDB management application with Cadenza Workbooks:

Since Cadenza Workbooks does not have its own user management system but instead relies on a central identity management solution like Keycloak via OpenID Connect (OIDC), a corresponding OIDC integration had to be established for the MDB management application (see Figure 1).

In addition, the management application utilized the geometry capture functionality previously provided by Cadenza Web Classic; this interface had to be migrated within the PHP application to the new Cadenza API provided by Cadenza Workbooks.

3 Design and implementation of the expansion to include invasive species

3.1 Technical aspects regarding Invasive Alien Species (IAS) data

As part of the requirements analysis, decisions were made regarding the necessary data fields, new reference lists, and the potential integration of the new invasive species module.

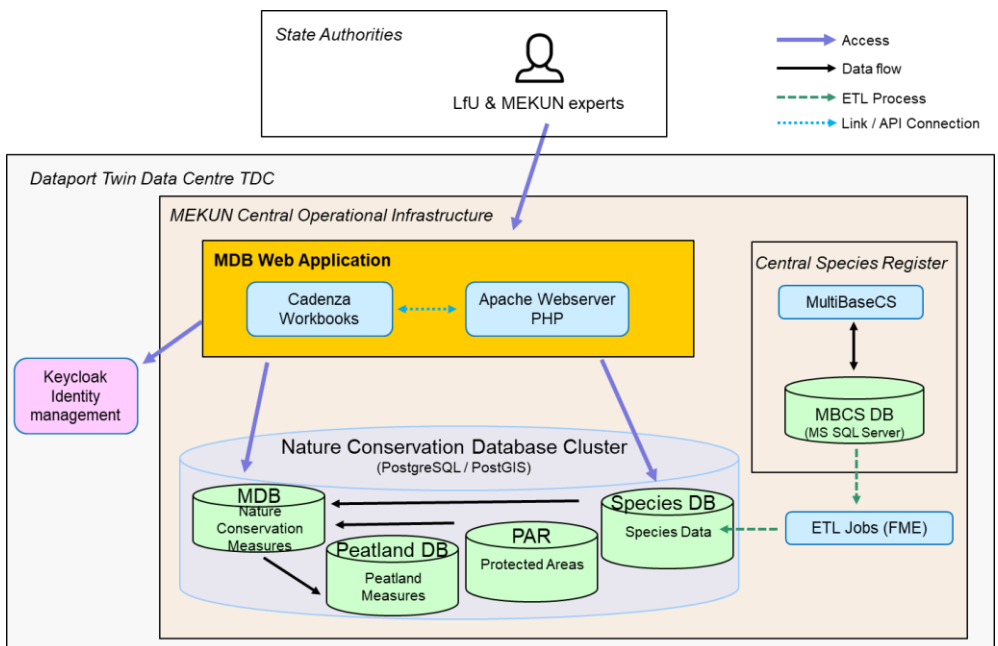


Figure 1: Main components used by MDB web application including key data flows, links, and access paths

A new measure category designated as "IAS measures" was introduced to capture all measures specifically planned to address invasive species that do not fall under other measure categories. Additionally, for other categories – such as conservation and development measures – a question was added to determine whether they also target invasive species. If so, the input fields for the respective measure category are supplemented with the additional fields required for invasive species management.

Monitoring impacts is essential for evaluating the success of invasive species control efforts and should be conducted at specific milestone dates. To accommodate this so-called *milestone monitoring*, another measure category designated as "monitoring measure" was introduced. This allows multiple monitoring measures to be linked to a single IAS measure, documenting the measure's success at the respective milestone dates.

3.2 Linking to specific occurrence locations of invasive species

A key requirement for measures addressing invasive species was establishing a data link between the actual occurrences of invasive species and the location data managed in the Central Species Register implemented using MultiBaseCS software (<https://www.multibasecs.de/>) [Ar24]. Consequently, each IAS measure in the MDB can be linked to one or more species occurrence records from MultiBaseCS. Since MultiBaseCS operates in a different technical environment, a direct link between the two applications could not be established. However, species data from the MultiBaseCS database is regularly transferred via ETL process (Extract – Transform – Load) to the nature conservation database, which is also used by the MDB (see Figure 1). This data also forms the basis for the species list within the MDB.

The MDB offers two methods for specifying the occurrence of an invasive species to which an IAS measure relates: either entering the data ID of the occurrence location from MultiBaseCS, or by selecting the occurrence location on the map.

For direct entry of the MultiBaseCS data ID, a filterable selection list of species occurrences is provided; this allows for direct ID entry while displaying the year and species name. This method is suitable for specialists who work in MultiBaseCS concurrently and can transfer the ID directly from that system.

When selecting via the map, the relevant species can be specified beforehand; this ensures that only occurrences of that specific species are displayed on the map and made available for selection and inclusion in the measure.

Additionally, every measure in the MDB includes one or more measure geometries, utilizing point, line, or area geometry types. A transfer function can be used to automatically adopt the point geometry of the species occurrence from MultiBaseCS as the measure geometry.

4 Migration from Cadenza Web Classic to Cadenza Workbooks

4.1 Setting up a new Workbooks repository with a dashboard and interactive analyses

The analyses compiled for the MDB – the MDB repository in Cadenza Classic – consisted of approximately 60 "filter forms" used to run queries; these typically presented search results as tables, but also as reports, charts, and maps.

The differing technical architecture of Cadenza Workbooks meant that all repository elements had to be rebuilt from scratch, as automatic migration was not possible. This reconstruction process provided an opportunity to align the analyses – which had evolved organically since their introduction in 2010 – with current requirements and to review them in light of the extensive new visualization options introduced with Cadenza Workbooks.

The new MDB repository developed by MEKUN in Cadenza Workbooks comprises 14 "workbooks" (see Figure 2), each containing multiple "worksheets."

The worksheets within a workbook are often linked by shared filter criteria, enabling different perspectives on a search (e.g., switching from an overview display to a detailed view). A worksheet can contain a dashboard comprising various individual views – such as tables, charts, indicators, and maps – as well as explanatory text with embedded current MDB data (see Figure 2). Key metrics – serving as so-called indicators – can be highlighted on a dashboard based on research findings, making the most important facts graspable at a glance. These newly available presentation options made it possible to tailor the analyses precisely to the functional requirements for using the MDB.

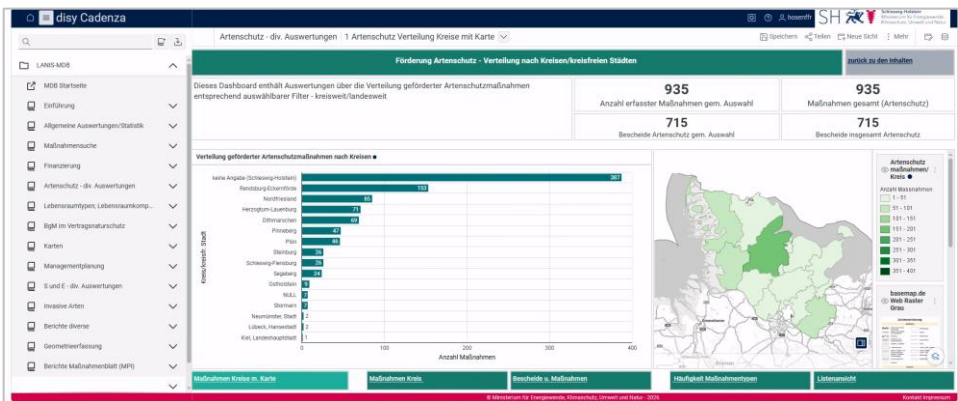


Figure 2: Cadenza „worksheet“ featuring a dashboard which combines various views displaying diagrams, tables, indicators, and maps (here: fictitious test data); the 14 „workbooks“ are listed on the left-hand side

4.2 Migration of the MDB Nature Conservation Management Application

The MDB management application, implemented in PHP, required adjustments for integration into Cadenza Workbooks regarding the following aspects:

- Connection to a central identity management system based on OpenID Connect, involving modifications to session management and access rights groups
- Utilization of the geometry capture functionality provided by Cadenza Workbooks via the new Cadenza API (`cadenza.js`)
- Addition of entry points to enable navigation from the Workbooks repository into the management application

In Schleswig-Holstein, the Keycloak software (<https://www.keycloak.org/>) is used for central identity management at the state data centre (see Figure 1). This Keycloak solution enables the connection of user credentials for all state agency employees via the statewide Active Directory (AD) while also allowing for the integration of external users through a unified interface.

The previous direct connection of the MDB to Cadenza's user management had to be replaced by an OpenID Connect (OIDC) interface to Keycloak. A recently developed suitable OIDC interface for another PHP-based management application in the geology sector [Ho25] – also created for the environmental department – could be reused and adapted to meet the specific requirements of the MDB application. It ensures that the user groups determining access rights within the MDB are appropriately mapped in both Keycloak and Cadenza and offers the convenience of Single Sign-On via Keycloak.

Another necessary adjustment involved reconfiguring the geometry capture function to use the Cadenza Workbooks API. The new JavaScript-based Cadenza API offers significantly more flexible integration and control options. However, it provides only client-side geometry capture. Consequently, the management application had to be extended with its own functionality for validating and saving geometries in the PostgreSQL (with PostGIS) database system. An advantage of this approach is that geometry capture can be embedded directly into the PHP-based management application.

Another benefit of switching to Cadenza Workbooks for the MDB is the availability of enhanced integration options between Cadenza dashboards and the management application; specifically, "drill-through" functions allow users to navigate directly from any Cadenza analysis view to the corresponding data entry forms in the PHP-based management application. This enables users to perform complex searches based on domain-specific criteria within Cadenza and then sequentially open and process all matching measures in the management application.

5 Summary and outlook

The nature conservation measures database (MDB) has been an important tool for the state authorities in Schleswig-Holstein since 2010, enabling them to plan, manage and monitor nature protection measures. The described recent enhancements and modernization efforts ensure continued efficient support for nature conservation tasks.

The addition of an invasive species module fulfills not only the recording and reporting obligations of EU Regulation No. 1143/2014 on invasive species [EU14] but also addresses additional requirements for invasive alien species (IAS) management in Schleswig-Holstein. Not only a dedicated category for IAS measures was added as well as a monitoring measure category in order to document results in combating alien species, but the IAS details could also be provided for measures of other categories.

As a further step in modernization, the MDB is being migrated from the previous analysis and GIS platform, *Cadenza Classic*, to the current web-based variant, *Cadenza Workbooks*. By rebuilding all analyses to be optimized for Cadenza Workbooks, the MDB's extensive data is being made accessible in a modern way and presented clearly. The new repository of MDB related analyses and visualizations unlocks newly available options and forms of presentation, such as dashboards and interconnected worksheets. The necessary OpenID Connect interface added to the PHP management application enables the user to utilize Single Sign-On. Greater flexibility regarding the integration of Cadenza and the management application is achieved by the Cadenza Workbooks API – for example to edit the geometries of measures.

Plans for the near future include optimising the integration of the PHP based management application into Cadenza Workbooks, as well as further expanding the analysis functions by utilising the new display formats available in Cadenza Workbooks.

Bibliography

- [Ar24] Artenerfassung im Zentralen Artenkataster Schleswig-Holstein (ZAK-SH) (01.07.2024): <https://www.schleswig-holstein.de/DE/fachinhalte/A/artenschutz/artenerfassungMultibase.html>, accessed: 18/08/2026.
- [DI26] Disy Informationssysteme GmbH: Software for Data Analytics, Business & Location Intelligence. <https://www.disy.net/en/products/disy-cadenza/overview/>, accessed: 18/08/2026.
- [EU09] European Union: Combating invasive species. Communication from the Commission to the Council, the European Parliament, the European Economic and Social Committee and the Committee of the Regions of 3 December 2008 – Towards an EU strategy on invasive species. Last updated: 30.01.2009. <https://eur-lex.europa.eu/EN/legal-content/summary/combating-invasive-species.html>, accessed: 18/08/2026.

- [EU14] European Union: Regulation (EU) No 1143/2014 of the European Parliament and of the Council of 22 October 2014 on the prevention and management of the introduction and spread of invasive alien species, 2014, <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32014R1143>, accessed: 18/08/2026.
- [Ho12] Hosenfeld, F.; Kleinbub, J.; Krüger, K.; Kumer, D.; Langner, K.; Rinker, A.: Development of a Database Supporting the Management of Nature Conservation Measures in Schleswig-Holstein. In (Arndt, H.-K.; Knetsch, G.; Pillmann, W. eds.): *Light up the Ideas of Environmental Informatics*, Dessau, pp. 623–629, 2012.
- [Ho21] Hosenfeld, F.; Rinker, A.; Langner, K.; Krüger, K.: Governmental Information Systems Supporting Nature Conservation in Schleswig-Holstein. In: (Wohlgemuth, V.; Naumann, S.; Arndt, H.-K.; Behrens, G., eds.): *Environmental Informatics. A bogeyman or saviour to achieve the UN Sustainable Development Goals? Adjunct Proceedings of the 35th edition of the EnviroInfo – the long standing and established international and interdisciplinary conference series on leading environmental information and communication technologies Berlin, Germany, September 27–29, pp. 133–140, 2021.*
- [Ho23] Hosenfeld, F.; Tiffert, J.; Jugelt, M.; Bretschneider, A.; Krummheuer, Y.: Management of climate protection measures in peatland areas in Schleswig-Holstein. *EnviroInfo 2023*. Bonn: Gesellschaft für Informatik e.V.. PISSN: 1617-5468. ISBN: 978-3-88579-736-4. pp. 143–152. *Innovative Approaches and Solutions*. Garching, Germany. 11.-13. Oktober 2023, 2023, <https://doi.org/10.18420/env2023-013>, last accessed: 18/08/2026
- [Ho25] Hosenfeld, F.; Willer, J.; Dzieran, L.; Tiffert, J.: Management of geological data and notifications in accordance with the Geological Data Act in Schleswig-Holstein. *EnviroInfo 2025*. Bonn: Gesellschaft für Informatik e.V.. PISSN: ISSN 2944-7682. pp. 155–162. *Open Data, Open Science and Open Society for a Sustainable Future*. Potsdam, Germany. 17.-19. September, 2025. https://doi.org/10.18420/env2025_13, last accessed: 18/08/2026
- [JA26] Janßen, J.,; Abecker, A.: Eine Datenplattform zum Austausch von Starkregenvorsorge-daten. In (Fuchs-Kittowski, F.; Abecker, A.; Hosenfeld, F.; Reineke, A.; Wachsmuth, M., eds.): *Umweltinformationssysteme – Digitale Innovationen für eine nachhaltige Zukunft*. Tagungsband des 32. Workshops „Umweltinformationssysteme (UIS2025)“ der Fachgruppe „Umweltinformationssysteme“ der Gesellschaft für Informatik (GI), 2026, pp. 35–56. Springer Vieweg, Wiesbaden, https://doi.org/10.1007/978-3-658-50065-8_3, last accessed: 18/08/2026
- [MP24] The Minister-President, State Chancellery: Open Innovation and Open Source Strategy of Land Schleswig-Holstein. Version 1.00. 20.11.2024. https://www.schleswig-holstein.de/DE/landesregierung/themen/digitalisierung/linux-plus1/Service/Downloads/_dateien/open-source-strategy_EN.pdf, last accessed: 18/08/2026

Common and Domain-Specific Design Patterns for Mobile AR Applications to Visualize Measures on Building Facades in an Environmental Context

Frank Fuchs-Kittowski^{1,2} , Maximilian Deharde¹ , David Starzl¹, Marius Poppel¹,
and Simon Burkard² 

Abstract: Mobile augmented reality (mAR) applications make it possible to visualize planned facade projects on-site with georeferencing and to scale. This paper presents two independently developed mAR prototypes and systematically compares them: Facade-Greening-AR for planning green facades (urban climate adaptation) and Facade-PV-AR for planning facade photovoltaics (energy transition). Both address the same physical domain, are based on an identical technical core (Unity, ARFoundation, Google Geospatial API, Google Streetscape Geometry API), and are aimed at citizens without specialized knowledge. The structured comparison across eight dimensions yields a catalog of seven design patterns: four cross-domain (georeferenced facade anchoring, visual in-situ assessment, local data persistence, domain-knowledge interface) and three domain-specific (discrete object placement, continuous area coverage, quantitative potential modeling). The catalog forms a generalizable foundation for future mAR applications in urban environmental contexts.

Keywords: Augmented Reality, green facades, facade photovoltaics, design patterns, requirements analysis, climate adaptation, renewable energy transition

1 Introduction

Cities must both reduce greenhouse gas emissions and adapt to the already unavoidable consequences of climate change [SM21]. Facade measures play a growing role in both areas of action: green facades, as nature-based solutions, mitigate urban heat islands, improve air quality, and enhance biodiversity [Pf16], while photovoltaics on facades contribute to decentralized electricity generation and thus to the decarbonization of the building sector [SS25].

A common challenge across both areas lies in communicating planned measures: stakeholders find it difficult to visualize what greening or PV systems would look like on an existing building and what ecological, economic, or aesthetic effects they would entail [St25, De26]. Mobile augmented reality (mAR) technology offers a promising solution: it enables a georeferenced, true-to-scale overlay of virtual objects onto the smartphone

¹ HTW Berlin, Berlin, Germany, frank.fuchs-kittowski@htw-berlin.de, <https://orcid.org/0000-0002-5445-3764> (Fuchs-Kittowski), <https://orcid.org/0009-0004-0335-3108> (Deharde)

² Fraunhofer FOKUS, Berlin, Germany, <https://orcid.org/0000-0001-6038-0891> (Burkard)

camera view directly on the facade on-site, making planned measures tangible [Az97].

mAR applications have been developed for various environmental planning tasks, such as visualizing urban greening measures [OSR23, Ci21], evaluating sites for renewable energy facilities [Ca16, Me23], or facilitating participation in planning processes [Bu22]. Although isolated design recommendations for urban AR planning exist, there has been no comparative analysis to date that distinguishes common patterns from domain-specific ones, thereby creating a transferable foundation for future developments.

This paper therefore addresses the following research question: What common and domain-specific design patterns can be identified for mAR applications used to visualize interventions on building facades in an environmental context? To answer it, two mAR prototypes developed in separate contexts at HTW Berlin are systematically compared: Facade-Greening-AR, an app for planning and visualizing facade greening [St25], and Facade-PV-AR, an app for planning and visualizing facade photovoltaics [De26]. The primary scientific contribution is a catalog of seven design patterns (M1 - M7) derived from two real-world implementations for mAR applications aimed at visualizing facade measures in an environmental context.

2 Background and Related Work

Green facades refer to the targeted greening of vertical building surfaces using soil- or wall-mounted planting systems [Pf23]. As a nature-based solution they block thermal radiation, cool the ambient temperature via evapotranspiration, filter pollutants and fine particulate matter from the air and contribute to biodiversity [Pf16], yet their adoption remains limited by planning barriers, a lack of imagination and insufficient information and advisory services [KKZ23]; assessing their potential quantitatively requires species-specific and seasonal parameters and is therefore difficult to standardize [THK19]. Facade PV, in turn, integrates solar modules into vertical building surfaces, either as curtain-wall facade elements or as building-integrated components (BIPV) [Hü18]. It converts incident solar radiation directly into electrical energy and can replace additional building materials [SM21], but is likewise held back by high investment costs, complex building code requirements and a lack of expertise among planners [SS25, Hü18]; its potential analysis covers technical, economic and environmental metrics, with orientation, tilt and shading as key influencing factors [SS25].

Augmented reality (AR) refers to the real-time overlay of physical reality with computer-generated, context-specific content [Az97]. For urban greening initiatives, mAR applications such as Cityscaper [Ci21] and GLARA [OSR23] visualize greening scenarios in street spaces, but do not address green facades. In building-integrated photovoltaics, AR has so far been used sporadically to highlight suitable facade surfaces [Ca16, Me23] and for shading simulation [VR15], while tools suitable for everyday use by laypeople are largely lacking [De26]. Comparative studies on design patterns in both domains are not yet available.

A georeferenced AR representation that accurately reflects the plan requires precise determination of the device's position [KH04], which GPS and compass typically cannot reliably achieve in urban environments [FFF16, St19]. Visual positioning services (VPS) offer significantly higher accuracy by aligning camera images with georeferenced reference images: the Google Geospatial API achieves position errors of approximately one meter and orientation errors of less than 5 degrees wherever Street View data is available [De26], and the complementary Streetscape Geometry API provides simple 3D meshes of surrounding buildings as a geometric basis for precise object placement on facades [St25, De26]. Both are part of the ARCore Extensions developed by Google.

3 Method: Comparative Design Study

To answer the research question, a comparative design study is conducted [SMM12]: two real-world system developments (Facade-Greening-AR and Facade-PV-AR) are systematically analyzed using a common comparative framework. The selection of systems follows the principle of maximum thematic coherence while ensuring maximum domain difference: both applications address the same physical object domain (vertical building facade), are based on a largely identical technical core, and target the same primary audience (citizens without specialized knowledge), yet operate in different environmental domains (climate adaptation vs. energy transition). This contrast enables the systematic separation of cross-domain design patterns from domain-specific ones.

The comparison is conducted across eight dimensions: requirements analysis (D1), georeferencing and facade detection (D2), object placement interaction (D3), object representation (D4), potential modeling (D5), data persistence (D6), UX and target audience suitability (D7), and domain-specific extensions (D8). The dimensions were inductively derived from the requirements documents of both prototypes, cover the entire development cycle from requirements gathering to UX, and were agreed upon in a workshop with all authors.

Both apps were developed within the same research group at HTW Berlin but in separate contexts: Facade-PV-AR by a research assistant as part of a research project [De26], Facade-Greening-AR as a bachelor's thesis [St25] supervised by that same research assistant. This ensures conceptual independence while sharing technical expertise, but does not completely rule out implicit assumptions arising from the supervisory relationship (see Section 7.2).

4 Facade-Greening-AR: Visualizing Green Facades

Domain, target audience and requirements. Facade-Greening-AR addresses urban climate adaptation through green facades. Its primary target audience consists of citizens without specialized knowledge who wish to explore the aesthetic potential of green

facades on buildings; secondary audiences include municipalities, environmental organizations and professional planners. The requirements analysis is based on structured workshops with environmental agencies, municipalities and potential end users, and on a systematic review of related literature and comparable apps, whose gaps were formulated as negative requirements; requirements are assigned unique identifiers and prioritized (MUST/SHOULD/CAN). The core functional requirements are realistic AR visualization of greening objects on facades, interactive placement and editing of 3D plant objects, georeferenced anchoring and persistent design management (FR-1 to FR-4), while plant catalog, collaborative functions and climate data integration (FR-5 to FR-7) are optional.

Architecture, placement and rendering. The application is based on Unity 6 with ARFoundation and ARCore Extensions, utilizing the Google Geospatial API and the Streetscape Geometry API, and follows a three-tier architecture (presentation, logic, data); localization, facade recognition and anchor placement reside in the logic layer, and the data layer persists designs as local JSON files. Placement is performed with a simple tap: raycasting against the facade geometry provided by the Streetscape Geometry API determines the intersection point, where a georeferenced AR anchor is instantiated. Each anchor corresponds to a discrete plant object, scaled via a slider and rendered from a library of pre-built, static plant models (including ivy and wild grapevine) with lighting and shadows computed by Unity's AR rendering system. The geometric complexity of organic plant forms poses a particular challenge and requires level-of-detail (LOD) and billboard rendering techniques, which should be addressed in future work.

Potential modeling and implementation status. Potential modeling was deliberately not designed for Facade-Greening-AR: for the primary target audience visual appearance is the decisive factor, and the ecological potential of a green facade encompasses heterogeneous parameters that are difficult to standardize due to seasonal and species-specific influences, in contrast to the uniform energy yield of facade PV. The prototype has been implemented for iOS and Android with localization, facade detection, plant placement and local data persistence; the plant catalog, gesture-based scaling/rotation, cloud/backend synchronization and a share function have not yet been implemented.

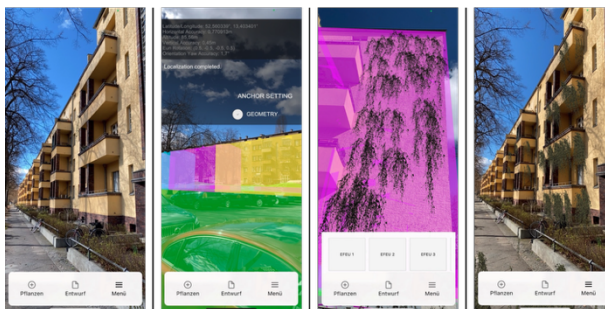


Fig. 1: mAR app Facade-Greening-AR (left: facade without AR greening; center-left: localization; center-right: plant selection; right: facade with AR greening)

5 Facade-PV-AR: Visualizing Facade Photovoltaics

Domain, target audience and requirements. Facade-PV-AR addresses the transition to renewable energy production using photovoltaics on building facades. Its primary target group consists of citizens who, without in-depth technical knowledge, wish to assess the visual appearance and energy potential of a facade PV system and thus generate a basis for deciding for or against its installation; secondary target groups include building owners, energy consultants and planning firms. The requirements were identified from the same two complementary sources and are likewise structured and prioritized [De26]. The MUST requirements are automatic capture of facade geometry and orientation-correct placement in the AR scene, interactive marking of sub-areas, automatic area-optimized module placement (conceptual; simplified in the prototype), manual adjustment of module type, color and spacing, and real-time calculation of technical, economic and ecological metrics (FR-01 to FR-05).

Architecture, placement and rendering. Facade-PV-AR uses the same technical core as Facade-Greening-AR: Unity 6, ARFoundation, Google Geospatial API and Streetscape Geometry API. Users mark the corner points of the area to be covered using tap gestures; the application then automatically determines an area-optimized layout from a catalog of real-world module types, with module spacing and type adjustable. The resulting layout is rendered in AR using static 3D models of the selected module type, placed automatically and rendered with standard shaders.

Potential modeling and implementation status. Every design change triggers an immediate recalculation of all key metrics: the energy potential of the facade is derived from the technical information of the panels, their orientation and the intensity of the sun's radiation at the given location, and from this, economic metrics (payback period, revenue) and environmental metrics (CO₂ savings compared to fossil fuel-based electricity generation) are derived and clearly presented. The prototype has been implemented for iOS and Android. A simplified version of the automatic layout generation lets the user mark a rectangle instead of a polygon, which is then textured with a PV module texture and displayed without real module type parameters; further implemented features are a rough facade potential calculation (approximation formula: $200 \text{ kWh/m}^2 \times \text{factor} \times \text{area of PV modules}$) and local storage of project files.



Fig. 2: mAR app Facade-PV-AR (left: mesh adjustment; center: facade marking; right: module and potential)

6 System Comparison and Design Patterns

6.1 Cross-Dimension Comparison

Dimension	Facade-Greening-AR	Facade-PV-AR
D1 Requirements analysis	Literature review and structured workshops; structured, prioritized requirements	Literature review and structured workshops; structured, prioritized requirements
D2 Georeferencing and facade detection	Geospatial API (VPS) + Streetscape Geometry API; raycasting via tap; manual calibration (transparent outlines)	Geospatial API (VPS) + Streetscape Geometry API; raycasting via crosshairs; manual calibration (transparent outlines)
D3 Object placement interaction	Location-specific: one tap creates one anchor per plant object	Area-continuous: corner points define an area that is filled with modules
D4 Object representation	Organic, complex meshes; many variants; high rendering effort; LOD/billboards; growth and seasonality	Rectangular, simple models; few variants; low rendering effort; standard shaders; static
D5 Potential modeling	None (deliberate design decision); aesthetic focus	Annual output, payback period, revenue, CO ₂ savings; recalculated on every change; investment focus
D6 Data persistence	Local JSON (plant type, geolocation, rotation)	Local JSON (module coordinates, parameters, potential metrics)
D7 UX and target audience suitability	Citizens; usability a prioritized non-functional requirement; single-tap placement; no formal evaluation	Citizens; usability a prioritized non-functional requirement; marking of several points; no formal evaluation
D8 Domain-specific extensions	Plant catalog (location, care, ecology); optional climate and weather data	Module specifications (type, color, efficiency, dimensions)

Tab. 1: Comparison of Facade-Greening-AR and Facade-PV-AR across the eight dimensions D1 - D8

The two applications are largely congruent in D1, D2, D6, D7 and D8: they share the requirements process, the georeferencing stack, local JSON persistence, the primary target audience and the absence of formal usability testing. The substantive differences lie in D3, D4 and D5, and all three are functionally determined by the physical nature of the subject

object rather than by design preference: a plant is location-specific, whereas PV modules cover a contiguous area (D3); organic plant geometry demands far more rendering effort and optimization than rectangular modules (D4); and only the investment-oriented PV case requires quantitative metrics alongside the visualization (D5). These three differences form the basis for M5, M6 and M7.

6.2 Derived Design Patterns M1 - M7

M1: Georeferenced facade anchoring (shared). *Context:* virtual objects must be anchored to real building facades precisely and persistently. *Solution:* a combination of VPS-based georeferencing (Google Geospatial API) and a facade geometry service (Google Streetscape Geometry API); raycasting against the facade geometry determines the exact placement point, and global coordinates from the Geospatial API persistently anchor it. *Evidence:* both systems use this pattern as their technical foundation.

M2: Visual in-situ assessment (shared). *Context:* non-experts should be able to assess the aesthetic impact of a facade measure without specialized software or static visualizations. *Solution:* AR overlay of virtual objects onto the real-time camera image on-site; the in-situ perspective under real lighting conditions creates an impact that static renderings cannot match. *Evidence:* both systems cite visual assessment of the building's appearance as their primary added value over existing planning tools.

M3: Local data persistence (shared, domain-specific expression). *Context:* planning drafts (anchor positions, object references, parameters) should be saved and restored between sessions. *Solution:* serialization of all planning-relevant data as human-readable JSON files in local storage. *Evidence:* both systems implement this pattern using an identical technical approach; only the JSON structure differs by domain.

M4: Domain knowledge interface (shared, domain-specific expression). *Context:* the placement decision depends on technical information that is not accessible to the primary target audience (plant suitability, module performance). *Solution:* structured linking of each placeable AR object to domain-specific technical information; shared as a pattern, domain-specific in its concrete form: plant catalog vs. module specifications. *Evidence:* both systems define such an interface.

M5: Discrete object placement (domain-specific: greening). *Context:* the object to be visualized is location-specific; each object exists at an individual, independent point on the facade. *Solution:* tap-based placement paradigm: a finger tap on the facade geometry generates a georeferenced anchor point for a single 3D object, intuitive and requiring no domain-specific prior knowledge. *Evidence:* Facade-Greening-AR; transferable to other location-specific facade objects (individual lights, information boards).

M6: Continuous area coverage (domain-specific: PV). *Context:* the individual objects to be visualized are discrete, but the facade area they cover is continuous and is to be occupied as a whole. *Solution:* area marking paradigm: users define the area to be covered

by specifying corner points, and the application automatically fills it with the objects. *Evidence:* Facade-PV-AR; applicable to all facade measures covering a contiguous area (e.g. composite thermal insulation systems, flat solar thermal collectors).

M7: Quantitative potential modeling (domain-specific: PV). *Context:* the target group's decision is investment-oriented and requires quantitative metrics (yield, payback period, CO₂ savings). *Solution:* tight integration of AR visualization and real-time potential calculation; any change in the plan (area, orientation, module type) triggers an immediate, clearly presented recalculation of all relevant metrics. *Evidence:* Facade-PV-AR; applicable to other facade measures with measurable performance (thermal insulation: heating cost savings; green roofs: rainwater retention).

7 Discussion

7.1 Interpretation of the Design Patterns

The four common patterns (M1 - M4) form a stable, reusable foundation for future mAR applications for facade interventions: the combination of VPS-based georeferencing and Streetscape-Geometry-based facade geometry (M1) solves the fundamental technical problem of precise, persistent object anchoring to real building facades and, together with local data persistence (M3) and Unity/ARFoundation, forms a proven technical basis for an entire class of such applications. That this identical core supports fundamentally different interaction paradigms (M5 vs. M6) and information requirements (none vs. quantitative KPIs) underscores that technical convergence does not preclude conceptual divergence: the effort for domain-specific design (M4 - M7) is at least comparable to that for the shared technical foundation. For a new application it must therefore first be clarified whether the object is location-specific (tap) or area-continuous (area marking), and whether the primary decision is aesthetic (no metrics required) or investment-oriented (real-time potential calculation required); a structured link between the AR object and domain-specific information (M4) should be planned from the outset in either case.

7.2 Limitations

Lack of formal user studies: both prototypes were evaluated without formal user studies, so the derived patterns are based exclusively on requirements analyses and implementation decisions, not on empirical usage data; SUS-based usability tests with the primary target group are strongly recommended for further work. **Prototype nature:** both implementations are prototypes, and patterns M1 - M7 refer to the conceptual level, not the implementation level; this should be taken into account during interpretation. **Limited generalizability:** the pattern catalog is based on two case studies; additional mAR applications for other facade measures (e.g. thermal insulation or facade renovation) would strengthen the evidence base and potentially reveal further patterns. **Construct**

validity: the comparison dimensions were derived from the source documents of both systems, and independent validation by external experts is pending; since both systems were developed by the same team, implicit assumptions may have influenced the derivation of the patterns (confirmatory bias).

8 Summary and Outlook

This paper is the first to systematically compare two mobile AR applications for facade interventions in an environmental context (facade greening and facade photovoltaics). Seven design patterns (M1 - M7) were identified: four cross-domain (M1 - M4) and three domain-specific (M5 - M7). The common patterns demonstrate a stable, reusable technical and conceptual core for this class of applications: georeferenced facade anchoring via VPS and Streetscape Geometry, visual in-situ assessment as the primary added value, local data persistence, and a domain knowledge interface. The domain-specific patterns show that the physical nature of the subject object determines the placement paradigm (discrete vs. continuous) and that the target group's primary decision-making type (aesthetic vs. investment-oriented) determines the need for quantitative potential modeling.

The derived pattern catalog forms a generalizable basis for future mAR applications for facade interventions in an environmental context. Further work should empirically validate the patterns through formal user studies, deepen the requirements analysis for Facade-Greening-AR with more stakeholder surveys, and expand the catalog with case studies from related facade domains (e.g. thermal insulation, renovation).

Acknowledgments

The content presented is part of the collaborative project PAREK. The PAREK project is funded by the Federal Ministry of Education and Research (BMBF) as part of the "KMUinnovativ:IKT" funding program, funding code 16IS24055 B.

References

- [Az97] Azuma, R. T.: A Survey of Augmented Reality. *Presence* 6(4), pp. 355 - 385, 1997.
- [Bu22] Burkard, S. et al.: A Mobile Augmented Reality Application for Displaying Planned Wind Turbines. In: *Environmental Information Systems*. Springer, Wiesbaden, pp. 21 - 41, 2022.
- [Ca16] Carmo, M. B. et al.: Augmented Reality for Decision Support on Solar Radiation Harnessing. In: *23rd Portuguese Meeting on Computer Graphics and Interaction (EPCGI)*. IEEE, pp. 1 - 8, 2016.
- [Ci21] Cityscaper GmbH: CITYSCAPER - Digital Citizen Participation with AR in Urban

Planning. <https://page.cityscaper.de>, 2021.

- [De26] Deharde, M. et al.: Mobile AR Application for the Integrated Visualization and Potential Assessment of Building-Integrated Photovoltaics. In: Environmental Information Systems. Springer, Wiesbaden, 2026 (forthcoming).
- [FFF16] Fedorov, R.; Frajberg, D.; Fraternali, P.: A Framework for Outdoor Mobile Augmented Reality and Its Application to Mountain Peak Detection. In: AVR 2016, LNCS 9768. Springer, Cham, 2016.
- [Hü18] Hülsmeier, F. et al.: SOLAR.shell. Fraunhofer IRB Verlag, 2018.
- [KH04] Karimi, H. A.; Hammad, A. (eds.): Telegeoinformatics. CRC Press, Boca Raton, 2004.
- [KKZ23] Knifka, W.; Karutz, R.; Zozmann, H.: Barriers and Solutions to Green Facade Implementation. Buildings 13(7), 1621, 2023.
- [Me23] Meireles, C. et al.: Mobile Augmented Reality for Analysis of Solar Radiation on Facades. In: Proc. VISIGRAPP 2023, Vol. 1: GRAPP. SCITEPRESS, pp. 131 - 142, 2023.
- [OSR23] Othengrafen, F.; Sievers, L.; Reinecke, E.: Using Augmented Reality in Urban Planning Processes. GAIA 32(1), pp. 54 - 63, 2023.
- [Pf16] Pfoser, N.: Fassade und Pflanze. Dissertation, TU Darmstadt, 2016.
- [Pf23] Pfoser, N.: Grüne Fassaden. DETAIL Praxis, Munich, 2023.
- [SM21] Schröder, J.; Moebus, S.: Climate-Sensitive Urban Planning and Urban Development. In: Versorgungs-Report Climate and Health. MWV, 2021.
- [SMM12] Sedlmair, M.; Meyer, M.; Munzner, T.: Design Study Methodology: Reflections from the Trenches and the Stacks. IEEE TVCG 18(12), pp. 2431 - 2440, 2012.
- [SS25] Schram, W. L.; Shirazi, E.: PV on Facades: A Financial, Technical, and Environmental Assessment. Energy and Buildings 328, 2025.
- [St19] Stranner, M. et al.: A High-Precision Localization Device for Outdoor AR. In: ISMAR-Adjunct. IEEE, pp. 37 - 41, 2019.
- [St25] Starzl, D.: Digital Design of Greening Measures: Concept and Prototype Implementation of a Mobile AR App. Bachelor's thesis, HTW Berlin, 2025.
- [THK19] Tudiwer, D.; Hollands, J.; Korjenic, A.: Berechnung der Kühlgestehungskosten von fassadengebundenen Begrünungssystemen im städtischen Raum. Bauphysik 41, pp. 120 - 124, 2019.
- [VR15] Veldhuis, A. J.; Reinders, A. H. M. E.: Shadow Analysis for BIPV and PIPV Systems in a Virtual Environment. In: 42nd IEEE PVSC, pp. 1 - 5, 2015.

Sustainability Context Inheritance: From Trusted Evidence to Product Sustainability Profiles in Circular Economy Decision-Making

Sandhya Koravi ¹

Abstract: Circular economy decisions increasingly depend on sustainability evidence distributed across product passports, certifications, supplier disclosures, lifecycle assessments and regulatory systems. This paper introduces the Sustainability Context Inheritance Architecture (SCIA), a conceptual architecture for determining whether heterogeneous evidence is sufficient for a product-level sustainability claim in a specific Decision Context. SCIA links evidence to product structures, exposes Coverage Context and Sustainability Gaps, applies decision-specific attribution rules, and produces a Product Sustainability Profile (PSP). An illustrative electric-scooter case shows how weighted evidence coverage, explicit gaps and a mandatory assurance condition distinguish evidence availability from decision admissibility. The contribution is a decision-scoped, coverage- and gap-aware approach that reports claims as Supported, Partially Supported or Unsupported without treating evidence sufficiency as sustainability performance.

Keywords: Circular economy, environmental informatics, environmental decision support, sustainability attribution, decision context, Digital Product Passport, sustainability evidence, Sustainability Lineage Graph, Product Sustainability Profile

1 Introduction

European circular-economy policy and ISO guidance emphasize the transition toward sustainable and circular production [EC20; IS24]. ESPR and DPP initiatives improve structured product-information availability and exchange [CI24; EU24]; battery-passport guidance defines relevant passport content [BP23]; and corporate sustainability reporting rules expand sustainability disclosure [EU22]. However, information availability alone does not make a product-level claim decision-ready.

Evidence remains distributed across supplier declarations, certifications, product carbon footprints, lifecycle assessments, audits and proprietary systems. The European Commission reports that 53% of assessed environmental claims were vague, misleading or unfounded and 40% lacked supporting evidence [EC23]. We refer to this decision problem as Environmental Information Asymmetry, extending information asymmetry to sustainability contexts [Ak70]. SCIA asks a specific question: for a given sustainability claim and decision, is the available evidence sufficient, incomplete or missing?

¹ Lloyds Technology Centre, Hyderabad, India,
koravi.sandhya@gmail.com,  <https://orcid.org/0009-0005-5087-793X>

SCIA is intended for decision roles such as OEMs, regulators, financial institutions, downstream actors and recyclers. In this paper, the Decision Context captures who is making or consuming the assessment (the actor) and the purpose for which the sustainability claim is being evaluated. Decision Context establishes the assessment scope; the Applicable Policy determines admissibility, materiality weights, mandatory conditions and thresholds.

The paper asks: RQ1: How can product-level sustainability claims be attributed from distributed evidence? RQ2: How can sufficiency, coverage limitations and evidence gaps be made explicit? RQ3: How can multi-organization and cross-jurisdiction evidence be consolidated into decision-ready PSPs?

The contribution comprises: (1) SCIA as an evidence-sufficiency architecture; (2) a Sustainability Lineage Graph linking product structures and evidence; (3) Coverage Context, Sustainability Gaps and Traceability Context as explicit sufficiency constructs; and (4) the PSP as a measurable, decision-scoped output reporting Supported, Partially Supported or Unsupported with supporting coverage, gaps and traceability. The contribution aligns primarily with SDG 9, SDG 12 and SDG 13.

2 Related Work and Research Gap

Existing sustainability mechanisms address complementary needs. DPPs provide structured mechanisms for associating product and lifecycle information with products and components [CI24; EU24]; GS1 EPCIS provides cross-organizational visibility-event data that can support lineage and traceability [GS1E]; GS1 Digital Link / Resolver connects identifiers to distributed resources [GS1R]; and digital credential mechanisms such as verifiable credentials can carry machine-verifiable assertions together with issuer, validity and status information [W3C25]. SCIA neither replaces nor depends on these mechanisms. Evidence may originate from them or from certifications, supplier disclosures and lifecycle assessments, with provenance and assurance metadata retained for later policy evaluation.

The research problem addressed by SCIA is specific: given heterogeneous and potentially incomplete evidence, is the evidence material to a particular product-level claim and decision sufficient to support that claim? SCIA represents the evidence through a Sustainability Lineage Graph, makes Coverage Context and Sustainability Gaps explicit, and applies Inheritance Assessment to produce a PSP. It is therefore a decision-oriented interpretation layer rather than another repository or traceability infrastructure; fitness-for-use research similarly emphasizes context rather than mere data existence [WS96]. The remaining gap is how missing or incomplete decision-relevant evidence is measured and carried into product-level attribution, not merely how available evidence is represented or exchanged.

Tab. 1: Positioning of SCIA with complementary sustainability information mechanisms.

Mechanism	Primary role	Relationship to SCIA
Digital Product Passports	Structured product/lifecycle information	Potential sustainability-evidence source
GS1 EPCIS	Cross-organizational events/traceability	Potential event/lineage source
GS1 Digital Link / Resolver	Identifier resolution to distributed resources	Optional discovery mechanism
Trust/conformity mechanisms	Certification, validation, assurance	Can provide qualification/assurance information
SCIA	Decision-scoped evidence sufficiency/attribution	Produces PSP from relevant evidence, coverage, gaps and rules

3 Research Methodology

This research follows Design Science Research (DSR) to design and evaluate a conceptual artifact addressing a socio-technical problem [He04; Pe07]. The problem is Environmental Information Asymmetry; the artifact is SCIA and its Lineage Graph, Coverage Context, Sustainability Gaps, Inheritance Assessment and PSP; evaluation uses an illustrative electric-scooter scenario.

The evaluation is conceptual, examining internal consistency, gap visibility and decision discrimination through an author-defined policy instantiation. Decision Context scopes the claim and relevant evidence paths; the illustrative policy supplies materiality weights, admissibility requirements, mandatory conditions and thresholds. These parameters are neither universal nor empirically calibrated. Prototype performance, calibration, stakeholder validation and empirical decision-quality evaluation remain future work; SCIA remains metric-agnostic across product categories, domains and regulations.

4 SCIA: Evidence Sufficiency Architecture

SCIA asks whether available evidence is sufficient to support a sustainability claim for a particular decision. The Sustainability Lineage Graph maps evidence to product structures and lifecycle relationships; Coverage Context describes evidence completeness; Sustainability Gaps expose missing or unresolved evidence; and Traceability Context links evidence to its source and product path. These generic evidence-state constructs are interpreted during Context-Specific Inheritance Assessment using Decision Context and Applicable Policy, producing a PSP.

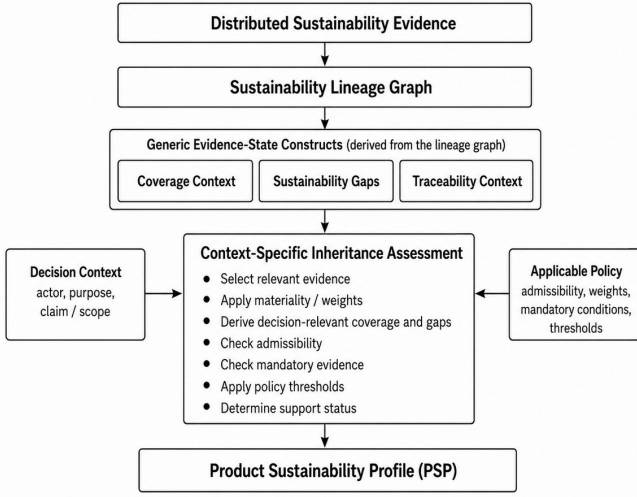


Fig. 1: Sustainability Context Inheritance Architecture (SCIA).

SCIA assesses evidence sufficiency, not sustainability performance, and does not prescribe how environmental metrics are calculated. Product carbon footprints or other metrics may be derived using applicable methods such as the GHG Protocol Product Standard [WR11] or Product Environmental Footprint (PEF) [EC21]. SCIA treats the resulting metric, methodological information and supporting evidence as inputs and assesses whether they are sufficient and admissible for the Decision Context and Applicable Policy. Its design principles are to trace evidence to source, make attribution explainable and expose rather than hide gaps.

5 Mapping Evidence, Coverage and Gaps

SCIA models the evidence landscape as a Sustainability Lineage Graph $G = (V, E)$, where V includes Product, Component, Material, LifecycleStage, Evidence and Credential nodes, and E includes relations such as contains, sourcedFrom, processedAt, evidencedBy and appliesTo. RDF or property-graph structures can represent the graph, with provenance, coverage and qualification information as attributes.

The graph supports three checks: what product/constituent structure is represented; what evidence and qualification information support each path; and where Sustainability Gaps remain. Evidence metadata may include provider or issuer, signer where applicable, product/component scope, provenance, validity or expiry status, and verification or assurance basis. Source-specific validation or verification results, where available, are retained as

qualification metadata; SCIA does not replace those mechanisms. Such metadata does not prove the sustainability claim; it enables later admissibility assessment.

Coverage Context describes completeness of the available evidence landscape and may combine composition, evidence, lifecycle, assurance and provenance coverage; SCIA does not prescribe a universal generic formula. Sustainability Gaps identify missing, incomplete, inaccessible, unverifiable or unknown evidence and do not imply poor sustainability performance. Traceability Context links evidence and gaps to their source and product path.

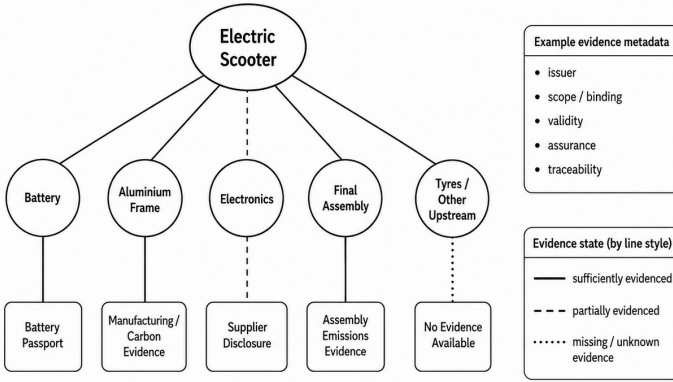


Fig. 2: Sustainability Lineage Graph showing evidence coverage and Sustainability Gaps.

A practical SCIA implementation can remain technology-agnostic: evidence could be resolved from available sources and matched to product, component, process or lifecycle nodes using identifiers and semantic mappings, while provenance and qualification metadata are retained. GS1 Digital Link / Resolver can support resource discovery [GS1R], and EPCIS can provide visibility-event information for lineage [GS1E]. The lineage itself may be represented using RDF or property graphs; SCIA does not require a specific implementation stack.

6 Inheritance Assessment

Inheritance Assessment interprets the generic evidence state for a Decision Context. For the illustrative low-carbon manufacturing assessment, let $c_i \in [0, 1]$ be evidence coverage of relevant path i and w_i its policy-defined materiality weight, with $\sum_i w_i = 1$. Decision-relevant coverage and gaps are:

$$C_D = \sum_i w_i c_i \quad g_i = w_i (1 - c_i) \quad G_D = \sum_i g_i.$$

With normalized weights, $C_D + G_D = 1$. These values quantify evidence sufficiency, not sustainability performance; weights represent decision relevance, not evidence reliability. Coverage measures the presence and completeness of decision-relevant evidence, but not whether that evidence satisfies the conditions under which a decision maker may rely on it. Evidence with identical coverage may differ in admissibility: policy may require independent assurance rather than self-attestation, an accepted accounting/assessment method, a recognized issuer or scheme, or a credential valid on the assessment date. SCIA therefore separates three questions: is relevant evidence available; is it admissible under the Applicable Policy; and is the admissible evidence sufficient for product-level inheritance? The gap terms g_i and G_D quantify evidence deficits; formal probabilistic uncertainty propagation remains future work. All values below are author-defined illustrations.

Tab. 2: Illustrative decision-relevant coverage and gap assessment.

Evidence path	w_i	c_i	g_i	Evidence state
Battery manufacturing	30%	100%	0%	Available
Aluminium/frame production	20%	100%	0%	Available
Electronics manufacturing	25%	50%	12.5%	Partial
Final assembly	20%	100%	0%	Available
Tyres/other upstream	5%	0%	5%	Unavailable
Total	100%	$C_D = 82.5\%$	$G_D = 17.5\%$	–

The illustrative policy requires $C_D \geq 90\%$ for Supported, $60\% \leq C_D < 90\%$ for Partially Supported, and $C_D < 60\%$ for Unsupported. Coverage alone therefore yields Partially Supported at $C_D = 82.5\%$, with $G_D = 17.5\%$ from electronics (12.5%) and tyres/other upstream (5%). This answers only how much decision-relevant evidence is available; it does not establish whether that evidence satisfies policy conditions for reliance on the claim. Section 7 applies those admissibility conditions and determines the final inheritance outcome.

7 Illustrative Decision Assessment and Product Sustainability Profile

The scenario asks whether an OEM has sufficient manufacturing evidence to substantiate, and therefore publish or rely on, a whole-product low-carbon manufacturing claim for an electric scooter. Section 6 established 82.5% decision-relevant coverage. The second assessment step tests whether available evidence is admissible under the Applicable Policy and whether the admissible evidence permits whole-product inheritance. The illustrative OEM policy requires $C_D \geq 90\%$ for Supported status and valid independent assurance for battery-manufacturing evidence from a provider accepted under the policy. This mandatory condition is author-defined for the illustration, not prescribed universally by SCIA.

Battery evidence is present, traceable and therefore 100% covered in the lineage, but its qualification metadata shows that the required verification expired before the assessment

date. Thus, evidence presence and coverage are distinct from decision-specific admissibility. The mandatory assurance failure overrides the coverage-based result and makes the final attribution Unsupported.

Mandatory admissibility condition failed $\Rightarrow$ Unsupported.

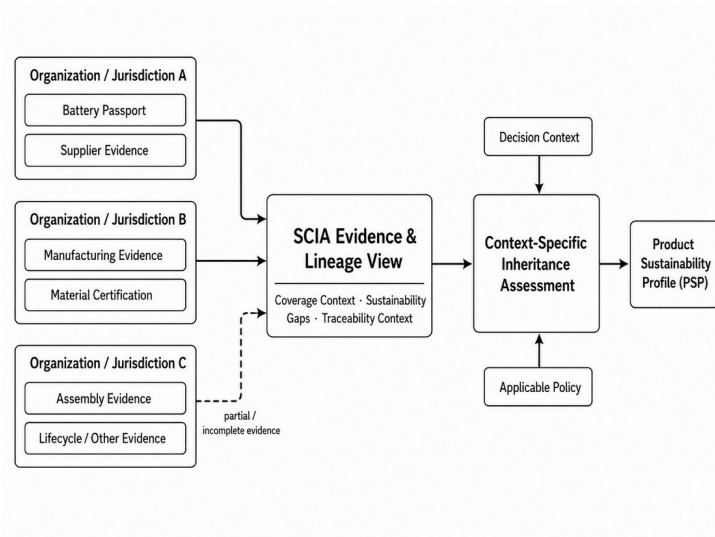


Fig. 3: Federated sustainability evidence contributing to a decision-scoped PSP.

Tab. 3: Illustrative decision-ready Product Sustainability Profile.

Decision information	PSP result
Claim being assessed	Manufacturing evidence is sufficient to substantiate a whole-product low-carbon manufacturing claim for the electric scooter
Decision	Whether the OEM should publish or rely on the claim
Applicable OEM policy	Supported: $C_D \geq 90\%$; valid independent battery-manufacturing assurance is mandatory
Observed evidence state	$C_D = 82.5\%$, $G_D = 17.5\%$; battery evidence present and traceable, but required verification expired
Final PSP outcome	Unsupported - mandatory battery-assurance condition failed; coverage alone: Partially Supported
Key coverage gaps	Electronics manufacturing: 12.5%; tyres/other upstream: 5%
Required action	Obtain valid battery assurance and close remaining upstream gaps before reassessment
Traceability	Evidence, gaps and qualification status remain linked to product/-component paths

The PSP prioritizes the decision-blocking condition before diagnostic coverage detail. Because it is decision-scoped, the same evidence may yield a different outcome for another actor; for example, a recycler assessing end-of-life recoverability may apply different evidence paths, weights, mandatory conditions and thresholds. The scenario addresses RQ1 through lineage-based attribution, RQ2 through explicit coverage and gaps, and RQ3 through a decision-ready PSP. Numerical values and policy conditions are illustrative and demonstrate internal decision transparency, not empirical improvement in human decision quality.

8 Limitations and Future Work

SCIA is conceptual and does not claim prototype, empirical, stakeholder or expert validation. The scooter weights, thresholds and mandatory assurance rule are illustrative rather than universal or calibrated. SCIA also does not establish the substantive truth of evidence; it preserves provenance and assurance information while Applicable Policy determines admissibility. Practical deployments would need to integrate, as appropriate, with relevant mechanisms and account for DPP access and interoperability requirements [EU24] and evidence validity or status changes such as credential revocation [W3C25]. Attribution accountability remains a deployment concern.

As the next phase of this research, we plan to empirically evaluate SCIA using externally sourced product and sustainability evidence, including real-product prototyping and controlled evaluation of decision-specific coverage and admissibility conditions. Subsequent work will address calibration, stakeholder evaluation, formal uncertainty propagation, dynamic policy management and context-dependent evidence-path selection.

9 Conclusion

SCIA provides an evidence-sufficiency architecture for product-level sustainability attribution. The Sustainability Lineage Graph connects distributed evidence while preserving traceability and qualification information; Coverage Context and Sustainability Gaps expose evidence completeness and limitations; and Inheritance Assessment applies Decision Context and policy to determine decision-relevant coverage, admissibility and final status. The PSP reports the outcome, blocking conditions, gaps and traceability needed to understand the decision.

The OEM illustration shows that evidence may be present and traceable yet inadmissible when a mandatory assurance requirement fails. SCIA is therefore not another sustainability metric, evidence repository or trust mechanism, but a structured approach for determining whether evidence is sufficient and admissible for a product-level sustainability claim in a particular decision.

References

- [Ak70] Akerlof, G. A.: The Market for “Lemons”: Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics* 84 (3), pp. 488–500, 1970.
- [BP23] Battery Pass Consortium: Battery Passport Content Guidance: Achieving Compliance with the EU Battery Regulation and Increasing Sustainability and Circularity, Battery Pass, 2023.
- [CI24] Wautelet, T.; Ayed, A.-C.: Exploring possible Digital Product Passport (DPP) use cases in battery, electronics and textile value chains, version 2.0, CIRPASS Consortium, 2024.
- [EC20] European Commission: A New Circular Economy Action Plan: For a Cleaner and More Competitive Europe, COM(2020) 98 final, Brussels, 2020.
- [EC21] European Commission: Commission Recommendation (EU) 2021/2279 of 15 December 2021 on the use of the Environmental Footprint methods to measure and communicate the life cycle environmental performance of products and organisations, *Official Journal of the European Union*, L 471, 30 December 2021, pp. 1–396, 2021.
- [EC23] European Commission: Green claims: New criteria to stop companies from making misleading claims about environmental merits of their products and services, Directorate-General for Environment, 2023, https://environment.ec.europa.eu/topics/circular-economy-topics/green-claims_en.
- [EU22] European Parliament and Council: Directive (EU) 2022/2464 as regards corporate sustainability reporting, *Official Journal of the European Union*, L 322, 16 December 2022, pp. 15–80, 2022.
- [EU24] European Parliament and Council: Regulation (EU) 2024/1781 establishing a framework for the setting of ecodesign requirements for sustainable products, *Official Journal of the European Union*, L 2024/1781, 28 June 2024, 2024.
- [GS1E] GS1 AISBL: EPCIS Standard, version 2.0.1, Published 1 July 2025, GS1, 2025.
- [GS1R] GS1 AISBL: GS1-Conformant Resolver Standard, version 1.2.1, Published 19 August 2026, GS1, 2026.

- [He04] Hevner, A. R. et al.: Design Science in Information Systems Research. *MIS Quarterly* 28 (1), pp. 75–105, 2004.
- [IS24] International Organization for Standardization: ISO 59004:2024 Circular economy – Vocabulary, principles and guidance for implementation, Geneva: ISO, 2024.
- [Pe07] Peffers, K. et al.: A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems* 24 (3), pp. 45–77, 2007.
- [W3C25] W3C: Verifiable Credentials Data Model v2.0, W3C Recommendation, 15 May 2025, 2025, <https://www.w3.org/TR/vc-data-model-2.0/>.
- [WR11] World Resources Institute; World Business Council for Sustainable Development: Product Life Cycle Accounting and Reporting Standard, Greenhouse Gas Protocol, 2011.
- [WS96] Wang, R. Y.; Strong, D. M.: Beyond Accuracy: What Data Quality Means to Data Consumers. *Journal of Management Information Systems* 12 (4), pp. 5–33, 1996.

Artificial Intelligence and Sustainability

AI model of a photovoltaic installation

Paweł Kobiela ¹, Jarosław Stańczak ^{2,3}, and Paweł Bartoszczuk ⁴

Abstract: This work aims to develop two models to simulate the operation of a photovoltaic system and predict electricity production using weather data, including solar radiation and system operating temperature. A prototype measurement system was developed to collect data from an operating photovoltaic installation. The obtained data set was prepared for developing linear regression and LSTM models simulating the operation of the installation. The models can be applied to estimate the value of electricity production for a specific installation in the future and the resulting ability to plan the production and the redistribution of this energy.

Keywords: : photovoltaic system modeling, LSTM, neural network, linear regression.

1 Introduction

This work describes the process of developing artificial intelligence models that simulate the operation of a photovoltaic panel system depending on the seasonal data (such as solar radiation and system operating temperature) and can be used for predicting electricity production. The problem is quite new, but there are some works that show similar models using neural or agent approaches, but rather taking into account economic aspects [Ag22], [DD26], and no one has tried to describe collecting data using a device designed for this purpose. Our approach presented in this paper focuses on presenting the task from scratch, from constructing the data collection equipment, through data processing, to creating models using selected methods. To build such a model, a special device has been constructed, which can store data directly from PV and can observe weather conditions. While there are computational models and formulas (shown in the next section) that describe PV panel parameters depending on changes in weather and other parameters, they are difficult to use directly because the specific parameters of the panels used are often unknown, as is the impact of their installation method and real weather conditions on their values. Therefore, their experimental identification is necessary. The next step was to collect the data, then process it, and use it to build models. Two models were used in this work: one simple, based on the least squares method, and the other, a much more complex one, employing an LSTM neural network. By creating two models, it was possible to compare their performance as predictors of electricity production.

¹ WIT Academy, Newelska 6, Warsaw, 01-447, Poland, P.Kobiela@wit.edu.pl, <https://orcid.org/0009-0003-7786-5934>

² Systems Research Institute, Polish Academy of Sciences, Newelska 6, Warsaw, 01-447, Poland, stanczak@ibspan.waw.pl, <https://orcid.org/0000-0003-3722-0085>

³ WIT Academy, Newelska 6, Warsaw, 01-447, Poland, J.Stanczak@wit.edu.pl

⁴ SGH Warsaw School of Economics, al. Niepodległości 16, Warsaw, 202-554, Poland, pbarto1@sgh.waw.pl, <https://orcid.org/0000-0002-2208-5960>

2 Basic concepts

The operation of a photovoltaic cell is based on the photovoltaic effect, a key phenomenon in solid-state physics that occurs in semiconductors (see [DLM24] or [CO17]). The basis of photovoltaic phenomena is a semiconductor material, most commonly crystalline silicon (Si). In pure silicon, the valence band is filled, and the conduction band is empty. The energy required for an electron to jump between the bands is determined by the band gap width (E_g). To facilitate conductivity, such a semiconductor is doped. To achieve n-type conductivity (electron conductivity, negative charges), it is doped with elements from group V of the periodic table, such as phosphorus (P). P-type conductivity (hole conductivity, positive charges) is obtained by doping with elements from group III, such as boron (B). At the boundary of the n and p layers, a layer of space charge is formed - the p-n junction - which is crucial for the operation of the photovoltaic cell. The photovoltaic effect is a process in which light is directly converted into electrical energy. To be absorbed and generate a carrier pair, a photon must have an energy greater than or equal to the band gap width (E_g). For crystalline silicon, $E_g \approx 1.12$ eV. Absorption leads to the transition of an electron from the valence band to the conduction band. The electric field in the p-n junction acts as a barrier. Immediately after generation, electron-hole pairs are separated. It is this process of creating and separating charges of different signs at opposite ends of the cell that generates the voltage. Total solar power at the average distance from Earth to the Sun, often described as AM0 (Air Mass Zero) power, is approximately 1361 W/m² (the so-called solar constant). However, after passing through the Earth's atmosphere, this radiation weakens, and at the Earth's surface it ranges from about 1200 W/m² at the equator, about 900 W/m² in Central Europe, and about 500-700 W/m² at the poles (at the highest position of the Sun). These parameters are called AM1.5 (Air Mass 1.5) values with the spectral composition of light appropriately modified by the atmosphere. Irradiance (G) is a fundamental indicator of the solar power reaching the module. Short-circuit current (I_{sc}) is almost perfectly linearly proportional to irradiance. This relationship is described by the formula (1):

$$I_{sc}(G) \approx \frac{I_{sc_{STC}} * G}{G_{STC}} \quad (1)$$

An increase in irradiance causes a logarithmic increase in the open-circuit voltage (V_{oc}), which is a weaker relationship than that of current.

$$V_{oc} = \frac{n * k * T}{q} * \ln \left(\frac{I_{sc}}{I_o} + 1 \right) \quad (2)$$

G , G_{STC} – irradiance, irradiance at Standard Test Conditions (STC, $T=25^\circ\text{C}$, $P_{light}=1000\text{W/m}^2$, AM1.5 light spectrum), I_{sc} , $I_{sc_{STC}}$ – short-circuit current, short-circuit current at STC, I_o - saturation current of the p-n junction, q - the charge of an electron ($1.602 \cdot 10^{-19}$ C), V_{oc} – open circuit voltage, n - diode quality factor (for Si cells ranges from 1 to 2), k - Boltzmann constant ($1.38 \cdot 10^{-23}$ J/K), T - cell temperature in Kelvin.

The power (P) of the electric current delivered at a certain moment is the product of the current and voltage (3), and the electric energy (E) is the integral of the current power during the operation of the panel (4).

$$P(t) = I(t) * V(t) \quad (3)$$

$$E = \int_{t=t_1}^{t_2} P(t) dt \quad (4)$$

$I(t)$ - instantaneous value of electric current, $V(t)$ - instantaneous value of electric voltage, t , t_1 , t_2 – time, start and end time of the system operation.

Instantaneous values of I and V are not exactly equal to those calculated from formulas (1) and (2); those given above are rather their upper estimates. The panels produce direct current (DC) with variable parameters. Connecting them to the power grid requires an inverter that converts the DC to alternating current (AC) (see Fig. 1).

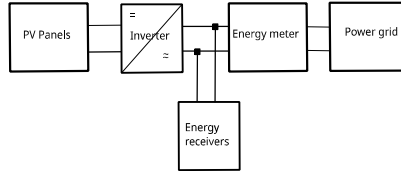


Fig. 1: Diagram of connecting PV panels and an inverter to the power grid.

Generally speaking, the values of t_1 and t_2 are determined by sunrise and sunset, but in reality, by the switching on and off of the inverter. All the above formulas show the most important factors that influence the power of the panel, such as irradiance, length of day, height of the sun above the horizon, junction temperature, and the material from which it is made, and more unpredictable factors such as cloud cover, dust, snow, hoarfrost, rain, and similar. It is not possible to examine their overall impact analytically, and experimental research is needed.

3 The system used to collect data during PV operation

Collecting data on the performance of PV panels requires a significant amount of data. Some of this data can be collected from the PV-inverter system. But data such as irradiance and panel temperature must be collected separately using a specially designed device, resembling a weather station, placed next to the panels.

3.1 The hardware

The device consists of a measurement system controller and connected light (irradiation) and temperature sensors. Access to the inverter is via a Wi-Fi router, which provides access to the local LAN and internet services (see Fig. 2). The main device of the system is a well-supported Raspberry PI minicomputer, which serves as the measurement system controller. This device is quick and easy to configure thanks to prebuilt drivers. It supports both C and Python compilers. Both languages have the necessary libraries for communication via serial interfaces, such as I2C and 1-Wire. The device is connected to sensors via GPIO ports and their associated interfaces, as well as a built-in Wi-Fi module for communication with a router and LAN. The minicomputer monitors the operation of temperature and light sensors, collecting data at predefined intervals, and reads production results from the inverter. The inverter transforms the output voltage and also monitors the operation of the photovoltaic system. Communication with the inverter is via TCP/IP using a local router, to which both the controller and the inverter are connected. The local router also provides internet connectivity for remote services such as Gmail, GitHub, and SSH.

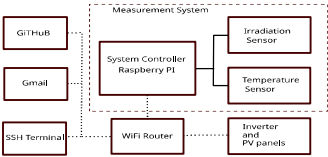


Fig. 2: System structure diagram

The measurement system reads data from the sensors, collects a 24-hour data set, which the controller saves to a CSV file kept locally on the SD card. The file is closed at midnight, and a daily report is generated and sent to Gmail. Then, the measurement system restarts the controller, logs into the LAN via Wi-Fi, and updates the software from GitHub. Next, it automatically launches the measurement script, and another daily cycle is executed.

Device ⁵	Type	Bus	Parameters
Termometer	DS18B20	1-WIRE	-55°C – 125°C
Light sensor ⁶	B-LUX-V30	I2C	0 – 200 klm
Microcomputer	Raspberry Pi Zero 2W 512MB RAM - WiFi	WHWiFi, Bluetooth, 1- WIRE, I2C	
Inverter	FRONIUS SYMO 6.0-3-M	WiFi	P _{max_out} ⁷ = 6kW
Photovoltaic modules (22pcs.)	Q.ANTUM Q.PEAK DUO-G5 305-330		P _{MPP} ⁸ = 330W

Tab. 1: Devices from which the system is built

5 Detailed information about the devices used can be found, in [Al26], [Df26], [Ra26], [Fr26], and [Mp26].

6 The light sensor used measures illuminance in [lm], but the energy of the light is connected with irradiance [W/m²]. However, these quantities are linearly related for a given light source (solar light spectrum) and can be recalculated using a formula: 685 Lux [lm/m²] = 1 W/m² or 1 Lux = 0.0079 W/m².

7 P_{max_out} - maximum inverter power fed into the power grid.

8 P_{MPP} – maximum electrical power generated under STC conditions (Maximum Power Point).

3.2 The software

The measurement system software consists of a set of scripts run in a planned order, placed together with libraries on the Raspberry PI device. The project has been placed in a GITHUB repository. You can download its latest version from the PVLogger.git repository. This solution provides great flexibility in the development and creation of code corrections. The measurement system has been programmed mainly in Python, C++, and shell commands to carry out individual subroutines in a certain logical order, which enables the implementation of the daily measurement cycle. Starting at midnight, the system executes scripts from MYSCRIPT.SERVICE containing scripts controlling the measurement system. The script AUTOSTART starts the internet connection, updates the software, and runs the main program. PVLOGGER is the main program that collects data in a loop at a specified time interval. Finally, each working day, the program sends the results and status notifications by email. Then, at midnight, it restarts and clears the data again. The block diagram of the system operation is shown in Fig. 3.

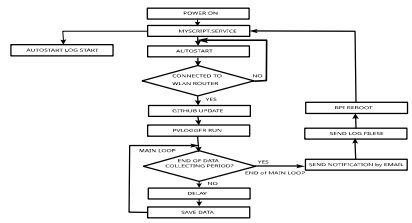


Fig. 3: The block diagram of the automatic measurement system operation

4 Results obtained

4.1 The data preprocessing

The first stage of processing the data obtained from the constructed device was to appropriately filter the data to eliminate sensor errors, interference, and data obtained without lighting. Raw data collected from the sensor during one day (2024/07/15) are presented in Fig. 4a). Apart from the data series, one can notice two series that are certainly erroneous because they contain data with impossible illuminance values. Therefore, the following data preprocessing was performed (results are shown in Fig. 4b)). The erroneous data were deleted, and the remaining data were filtered to remove noise and outliers using a Gaussian filter (5) (see [OS094]).

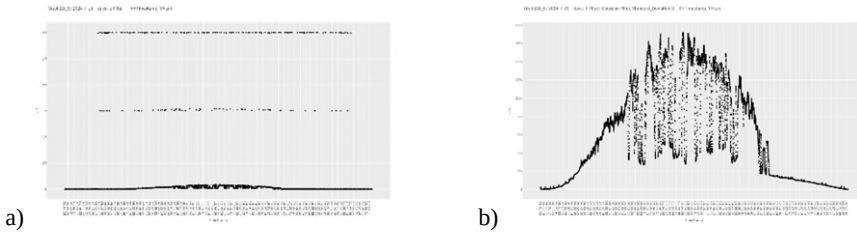


Fig. 4: a) The noisy raw data obtained from the light sensor and b) the data after preprocessing using a Gaussian filter ($\sigma=2$)

$$G(x)=\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(\frac{-x^2}{2\sigma^2}\right) \quad (5)$$

$G(x)$ – the value of the Gaussian function at point x , σ – standard deviation (controls the "blurring" of the filter), x – distance from the center (can be, for example, -1, 0, 1 for a 3-element filter).

4.2 The models

Linear regression model. The method of least squares (OLS – Ordinary Least Squares) [Co03] consists of fitting a simple regression line/surface to minimize the sum of squares of differences between the observed and predicted values of the dependent variable. The coefficients of the regression surface (6) are determined from the formula (7).

$$y=a_1 * x_1+a_2 * x_2+b \quad (6)$$

$$A=(X' * X)^{-1} * X' * Y \quad (7)$$

Y is a dependent data vector, and X is a matrix of independent data supplemented with a unit vector, $A = [a_1, a_2, b]$ – a vector of searched regression parameters.

Time	a_1	a_2	b	R^2	Number of days
	(temperature)	(irradiation)			
07-2024	10.282	0.061775	-2942.5	0.9637	31
08-2024	14.243	0.062507	-4109.9	0.9571	12
09-2024	15.437	0.077128	-4444.7	0.9509	12
10-2024	4.7887	0.071981	-1332.1	0.9117	12
11-2024	57.765	0.050808	-1594.9	0.7200	6
12-2024	5.7580	0.077986	-1605.7	0.9424	2

2024	3.5798	0.065990	-947.85	0.94643	78
01-2025	1.1715	0.035135	-341.31	0.3700	13
02-2025	4.9480	0.076804	-1401.0	0.9325	17
03-2025	5.7610	0.083579	-1565.0	0.9731	24
04-2025	4.1314	0.082172	-1084.8	0.9502	30
2025	5.6454	0.083158	-1566.3	0.9622	85
Total	-4.3980	0.075880	1348.1	0.9381	163

Tab. 2: Summary of linear regression results for individual months, parts of 2024 and 2025, and for all data (a_1 , a_2 , and b – regression parameters, R^2 - the coefficient of determination)

The model parameters presented in Tab. 2 are quite surprising, because while for all shorter periods they are quite similar, especially the dependence on irradiance, for the entire data set, the obtained model is different and, interestingly, more similar to the physical dependence of panel efficiency on temperature. Fig. 5 shows graphs presenting the features of selected models.

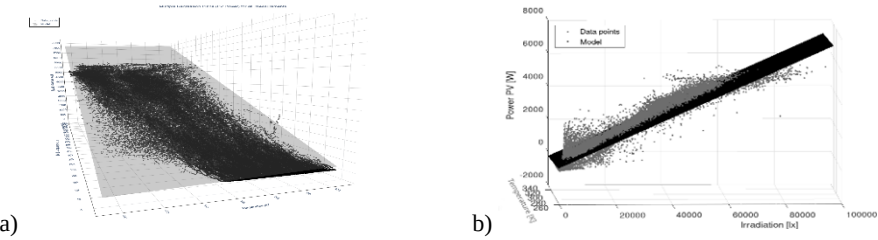


Fig. 5: 3D linear regression model obtained for all data, the dependence of electricity production a) on panel temperature; b) on solar radiation and temperature of panels

LSTM neural network. The LSTM (Long Short-Term Memory) neural network (see Fig. 6) belongs to the category of recurrent networks, which are characterized by the ability to process data in a time-dependent manner based on current and past values ([GSC00], [CO17]). This enables deeper analysis of dependencies and introduces an element of memory or inertia, which is clearly present in the operation of PV panels. The operation of the LSTM network is described by formulas (8)–(11).

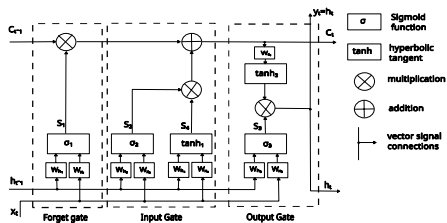


Fig. 6: Schematic diagram of the structure of one segment of the LSTM network

$$s_i = \sigma_i(W_{x_i} \cdot x_t + W_{h_i} \cdot h_{t-1} + b_i) \quad i=1,2,3. \quad (8)$$

$$s_4 = \tanh_1(W_{x_4} \cdot x_t + W_{h_4} \cdot h_{t-1} + b_4) \quad (9)$$

$$c_t = s_1 \cdot c_{t-1} + s_2 \cdot s_3 \quad (10)$$

$$h_t = y_t = s_4 \cdot \tanh_2(W_{c_t} \cdot c_t + b_{c_t}) \quad (11)$$

W_{x1} , W_{x4} – weights for the input signal; W_{h1} , W_{h4} , W_{c1} – weights for the short-term and long-term signal; b_1 , b_4 , b_{c1} – threshold values (biases); s_1 , s_4 – auxiliary symbols for the output signals from the corresponding activation function modules; c_t – long-term signal; h_t – short-term input signal; t , $t-1$ – subsequent time moments.

The input values were: irradiance, panel temperature, time, day of the week, and day of the year. The vector of standard output signals was the PV power production. The data was indexed at a time resolution of 1 minute. The data was appropriately preprocessed and normalized, and divided into training and validation parts. The data selected for learning and testing the considered models were selected randomly, but using data from the whole days. The key challenge is selecting appropriate network parameters and training procedures. The well-chosen parameters yield good results with an R^2 value better than 0.99. Results obtained using the LSTM network are shown in Fig. 7. The figures show that the model obtained using the LSTM network is very accurate.

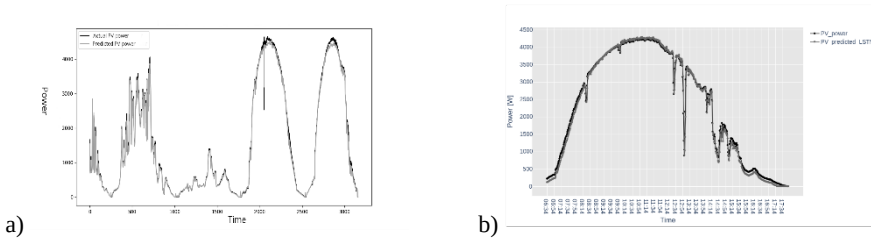


Fig. 7: The comparison of real and obtained results, a) real data and test data; b) LSTM predictions for March 21, 2025.

4.3 PV output power prediction

Tab. 3 presents the results of predicting the electricity production for selected days of the training data. The significant advantage of the LSTM-based model is evident, although the OLS method also gives acceptable results, considering the effort required to create the LSTM model. Tab. 4 shows the results obtained for testing data. The obtained accuracies (12) are lower than for training data, but for the LSTM, they are still very high, demonstrating the enormous potential of using it for modeling.

$$S_p = \frac{(E_r - |E_p - E_r|) * 100\%}{E_r} \quad (12)$$

S_p – daily prediction accuracy, E_r – produced energy, E_p – prediction of the evaluated model.

Date	Energy [kWh]	Prediction: OLS [kWh]	Prediction: LSTM [kWh]	Accuracy OLS [%]	(S)Accuracy LSTM [%]	(S)
15-08-2024	32.195	34.792	31.302	91.934	97.226	
15-10-2024	5.094	5.599	4.68	90.086	91.873	
15-02-2025	4.028	5.375	3.689	66.559	91.584	
21-03-2025	26.621	23.231	26.026	87.266	97.765	

Tab. 3: Prediction results of daily electricity production for the training data.

Date	Energy [kWh]	Prediction: OLS [kWh]	Prediction: LSTM [kWh]	Accuracy OLS [%]	(S)Accuracy LSTM [%]	(S)
14-06-2025	40.911	32.546	39.769	79.355	97.209	
15-07-2025	30.513	24.701	28.717	80.952	94.113	
11-08-2025	29.960	24.142	28.406	80.581	94.813	

Tab. 4: Prediction results of daily electricity production for the testing data.

5 Summary

The developed device allows for estimating production data from any time range and can be used as a predictor to estimate energy loss in the event of a forced shutdown. It is also possible to use the model as a predictor of energy production for a specific time period in the future. In this case, the forecast of irradiance and temperature for that time period is required. These data can be supplied either from external sources or estimated using another prediction model. It should be emphasized, that the system is quite simple, containing only two basic sensors.

6 Bibliography

[Ag22] Agga, A. et al.: CNN-LSTM: An efficient hybrid deep learning architecture for predicting short-term photovoltaic power production. Electric Power Systems Research 204, article no 107908, Elsevier, 2022.

[CO17] Carroll, B., Ostile, D.: An Introduction to Modern Astrophysics, Cambridge University Press, Cambridge, 2017.

[Ch23] Charu, A.: *Neural Networks and Deep Learning*, Springer, Hamburg, Germany, 2023.

- [Co03] Cohen, J., et al.: Applied multiple regression/correlation analysis for the behavioral sciences, Hillsdale, NJ: Lawrence Erlbaum Associates, 2003.
- [DD26] Dragomir, F., Dragomir, O.: *Photovoltaic Power Forecasting with AI: A Cost-Benefit Framework Across Multiple Time Horizons*, *Future Internet* 18(6):291, MDPI, 2026.
- [DLM24] Dyndał, K., Lewińska, G., Marszałek K.: Fotowoltaika, Kraków Open AGH e-Podręczniki, 2024, <https://epodreczniki.open.agh.edu.pl/handbook/35>, accessed 2026/04/13
- [GSC00] Gers, F., Schmidhuber, J., Cummins, F.: Learning to Forget: Continual Prediction with LSTM, *Neural Computation*, 12 (10): 2451–2471, (2000)
- [Al26] Alldatasheet, www.alldatasheet.com/datasheet-pdf/pdf/58557/DALLAS/DS18B20.html, accessed 2026/04/14
- [Df26] Dfimg, dfimg.dfrobot.com/wiki/18648/SEN0390_b-lux-v30b-ambient-light-sensor_datasheet_V1.0.pdf, accessed: 2026/04/14
- [Ra26] Raspberry Pi, www.raspberrypi.com/documentation/computers/os.html#gpio-and-the-40-pin-header, accessed 2026/04/14
- [Fr26] Fronius, www.fronius.com/en-gb/uk/search-page?searchword=fronius-symo-3-0-3-m&activeDepartment=solarenergy&categoryFacet=category-downloads, accessed 2026/04/14
- [Mp26] Mpe-online, mpe-online.ie/wp-content/uploads/2018/11/Hanwha_Q_CELLS_Data_sheet_QPEAK_DUO-G5_305-330_2017-10_Rev01_EN.pdf, accessed 2026/04/14
- [OS09] Oppenheim, A., Schaffer, R.: Discrete-Time Signal Processing, Pearson, 2009.

Environmentally Informed GenAI Use in ITSM

A Problem Analysis of Organisational Use-Case Decision-Making

Larissa Koch de Souza ¹ and Martin Engstler ²

Abstract: Generative AI (GenAI) is increasingly considered for IT Service Management (ITSM), promising automation, predictive support, service optimisation and improved knowledge processing. At the same time, AI systems create environmental burdens through energy use, emissions, water consumption and hardware-related lifecycle effects. This paper examines environmentally informed GenAI use-case decisions in ITSM. As a problem-analysis echelon within a broader Design Science Research project, it reports a Systematic Literature Review and synthesises literature into six problem characteristics. The study clarifies the focal construct and formulates a literature-grounded problem statement for subsequent objectives and requirements definition.

Keywords: Generative AI (GenAI), IT Service Management (ITSM), digital sustainability, AI governance, problem analysis

1 Introduction

Generative Artificial Intelligence (GenAI) is increasingly adopted across organisations and used to reshape organisational processes, structures and use cases [WEF24]. While GenAI is mainly associated with productivity gains [WEF24], it intensifies a multifaceted governance debate. In particular, existing research highlights the need to consider its consequences, alongside requirements for fairness, transparency and explainability [BI23]. These requirements become more pressing when ecological sustainability is considered. For example, training the underlying deep neural networks is energy-intensive and requires a significant amount of resources [RG25]. Moreover, ecological impacts extend beyond electricity-related emissions to water consumption and impacts across hardware supply chains [RG25], potentially jeopardising sustainability goals [Ga21].

These tensions become particularly actionable in IT Service Management (ITSM), a service-oriented governance domain that aligns business needs with IT capabilities [IE13] and structures decision-making along the IT service lifecycle to improve coordination and service quality [GSF16; Or14]. More recent studies explore deep learning for predictive

¹ Hochschule der Medien Stuttgart, 70569 Stuttgart, kochdesouza@hdm-stuttgart.de,  <https://orcid.org/0009-0003-7862-1851>

² Hochschule der Medien Stuttgart, 70569 Stuttgart, engstler@hdm-stuttgart.de,  <https://orcid.org/0000-0002-0687-1814>

analyses in ITSM, while underscoring prerequisites such as ensuring data quality, process standardisation and system integration [He25]. In contrast, ecological sustainability remains weakly embedded in ITSM research and practice, with only a few contributions addressing this topic [e.g. CT11; SR13]. Overall, related research offers important but fragmented foundations. Green information systems (IS) literature conceptualises IS as socio-technical means for supporting sustainable business practices [Ij10], while digital sustainability broadens this view to the development and deployment of digital resources for environmental, societal and economic welfare [KOS23]. Within this domain, sustainable AI research further highlights energy efficiency, lifecycle management, data management and reduced carbon emissions [Ko25; Le25]. However, these streams do not yet clarify how environmentally relevant evidence should impact concrete organisational decisions on GenAI use cases in ITSM. This creates a problem of translation between environmental evidence and ITSM-specific decision-making.

This paper therefore examines the problem space of environmentally informed GenAI use-case decision-making in ITSM, understood as the organisational process through which responsible stakeholders consider environmentally relevant consequences when deciding on a GenAI-enabled ITSM use case. Accordingly, this study focuses on the following research question: *What characterises the problem of environmentally informed decision-making regarding the use of GenAI in ITSM?* This study is part of a broader DSR project that aims to develop and evaluate decision support for environmentally informed GenAI use-case decisions in ITSM. Following the echeloned DSR logic of [TWV24], the present paper reports the problem-analysis echelon and provides a literature-grounded snapshot of the initial problem understanding and the derived problem characteristics. The aim of this paper is thus to clarify the focal construct, synthesise fragmented literature into problem characteristics and formulate a literature-grounded problem statement. By addressing environmentally responsible decisions on resource-intensive digital technologies, this paper relates to the Sustainable Development Goals (SDGs) [UN15], primarily SDG 12 (Responsible Consumption and Production) and secondarily to SDG 13 (Climate Action).

2 Conceptual background and construct clarification

ITSM is commonly understood as a service-oriented governance domain [IE13] that aligns business requirements with IT capabilities and structures the delivery of IT services in a scalable and quality-assured manner [AMB17]. In this sense, ITSM is associated with improvements in strategy, transparency and service quality across complex organisational settings [GSF16; MT23]. Within this ITSM context, GenAI is increasingly discussed as a transformative technology, e.g. for predictive automation and improved service availability [Si25] or for the usage of AIServiceOps with an emphasis on resilience, proactive insights, customer satisfaction and business wellbeing [VSM23]. Overall, AI in ITSM is mainly discussed in terms of automation, efficiency and service quality – oftentimes in the context of incident management [e.g. VSM23]. The ecological sustainability perspective of including AI into ITSM is only weakly addressed.

In IS research, environmental sustainability is commonly connected to digital sustainability and Green IS. Digital sustainability refers to the development and deployment of digital resources and artefacts toward improving environmental, societal and economic welfare [KOS23], while Green IS focuses on information systems that contribute to sustainable business processes and stresses their socio-technical nature [Ij10]. Although AI for sustainability is already discussed in IS research [Sc23], explicit work on sustainable ITSM or sustainable GenAI use in ITSM remains scarce. Sustainable IT services have been argued to extend beyond economic performance [HA09], ITSM has been discussed as a lever for Green IT [CT11], and practitioners have called for sustainability-oriented extensions of ITSM standards [SR13]. More recent guidance, such as the ITIL sustainability module, acknowledges sustainability in digital and IT contexts but remains largely descriptive [Ax21]. Regarding AI, sustainable AI research highlights energy efficiency, lifecycle management, sustainable data management and reduced carbon emissions [Ko25], while lifecycle-oriented approaches emphasise that sustainable AI cannot be reduced to a single metric or isolated lifecycle phase [Le25]. However, these streams do not yet clarify how environmental evidence should inform concrete organisational decisions on GenAI use cases in ITSM.

Against this background, the focal construct of this study is *environmentally informed GenAI use-case decision-making in ITSM*. It consists of the organisational process through which responsible stakeholders consider environmentally relevant consequences when deciding on the introduction, configuration, continued operation, modification or retirement of a GenAI-enabled ITSM use case. Following the principle of construct clarity [Su10], this construct is deliberately narrower than Sustainable AI, Green AI, AI for Sustainability or Responsible AI. It focuses on the organisational decision problem of relating environmental evidence to ITSM-specific governance responsibilities and decision alternatives rather than on the general sustainability of GenAI systems or AI-enabled sustainability goals. This study addresses this gap by structuring the problem space before deriving solution objectives or a later decision-support artefact.

3 Research Design

Methodologically, this study is positioned within a broader DSR project on Sustainable ITSM. [He04] provide the artefact- and problem-oriented DSR paradigm, while [Pe07] provide the process logic from problem identification to design and evaluation. This paper focuses on reporting a self-contained problem analysis echelon following the **echeloned DSR logic** [TWV24]. The central output is a literature-grounded problem statement that clarifies the characteristics, prior coverage and apparent solvability of environmentally informed GenAI use-case decision-making in ITSM. Accordingly, literature coverage captures which parts of the focal problem are already addressed, whereas remaining problem and solvability capture what remains unresolved and whether the literature indicates plausible structures or mechanisms through which it could subsequently be addressed. The focal construct (section 2) served as the scoping device for the literature

analysis. Publications were assessed according to whether they provided direct or transferable evidence on environmentally relevant organisational decisions concerning GenAI-enabled ITSM use cases. To identify this evidence, a focused Systematic Literature Review (SLR) was conducted following a nine-step protocol based on [KC07]:

First, the **guiding research question** was formulated: “What characterises the problem of environmentally informed decision-making regarding the use of generative AI in IT Service Management?” Second, the **key concepts** were harmonised across the three core components of ITSM, GenAI and ecological sustainability. Broader AI terminology was included to account for heterogeneous wording in the literature, however, the subsequent content-based screening focused on relevance for GenAI-enabled ITSM decision-making. Third, **corresponding keywords** were constructed. Fourth, the **databases** Scopus, IEEE Xplore, ScienceDirect and AIS eLibrary (AISEL) were selected. Fifth, the **search strategy** was defined as title, abstract and keyword searches, restricted to peer-reviewed journal and conference publications. The search string was adapted to database-specific syntax and combined the three components as follows: (*ITSM OR “IT Service Management” OR “incident management”*) AND (*“generative AI” OR “artificial intelligence”*) AND (*sustainab* OR ecologic**). Sixth, **inclusion and exclusion criteria** were specified. The review was limited to English- and German-language publications from 2020 to 2026. Publications were included if at least two of the three core components were central topics and if they contained extractable evidence relevant to the focal construct. Publications were excluded when relevant terms appeared only in venue or conference metadata, when “incident management” referred to non-IT contexts, when sustainability was used without an environmental dimension, or when no extractable evidence for the focal construct was available. These criteria were necessary because the search string intentionally captured directly relevant but also transferable literature. Seventh, the studies were **screened in stages** using a PRISMA-inspired screening flow adapted from [Pa21]. The first stage removed duplicates, records not matching the title or keyword criteria, as well as 12 publications without full-text access. The large reduction of 1621 papers primarily resulted from duplicates and title/keyword screening, particularly because the broad search term “incident management” retrieved numerous non-IT contexts. Subsequent abstract and full-text screening resulted in a total of 22 publications included in the analysis (see Figure 1).

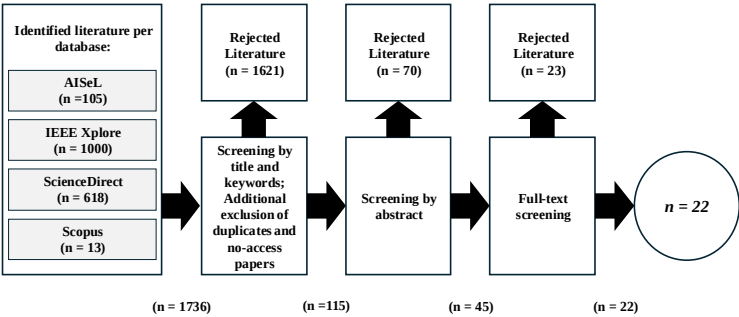


Fig. 1: PRISMA-screening stages based on [Pa21]

Eighth, the 22 publications were **extracted and coded** in a problem-oriented evidence repository covering literature stream, unit of analysis, ITSM relevance, AI scope, environmental scope, problem claims, causes or mechanisms, consequences and remaining gaps. Ninth, the coded evidence was **conceptually synthesised** by comparing recurring problem claims, causes, consequences and remaining gaps across publications and grouping related evidence. These preliminary problem characteristics were iteratively consolidated where they addressed the same decision problem. The resulting problem characteristics were assessed for literature coverage and the apparent solvability of the remaining problem following the problem-analysis echelon logic [TWV24]. The output is therefore a set of literature-grounded problem characteristics and a synthesised problem statement that informs subsequent solution objectives and artefact design requirements.

4 Results of the problem analysis

Overall, the SLR confirms that the intersecting literature of GenAI, ITSM and ecological sustainability is still fragmented rather than absent. Given the narrow focal construct and applied screening criteria, the 22 retained publications provide a focused view of the relevant literature rather than an indication of the field’s overall maturity. They provide direct, transferable or contextual evidence on parts of the problem space, but do not form a coherent research stream on environmentally informed GenAI use-case decisions in ITSM. The inductive synthesis leads to six problem characteristics, assesses their literature coverage and evaluates the apparent solvability of the remaining problem (see Table 1).

Problem characteristic	Literature coverage	Remaining problem and solvability
<i>PC1: Temporal integration deficit</i>	[Be25; Le25; Si25; VSM23]	Environmental evidence is not positioned within ITSM decision points. Moderately solvable through lifecycle structures.
<i>PC2: Lifecycle and boundary fragmentation</i>	[DJ24; Ko25; Le25; LMD21; Ra26; Ro24]	AI impacts are assessed at model or infrastructure level, while ITSM decisions concern use cases. Partly solvable due to boundary and transparency issues.
<i>PC3: Environmental–governance disconnect</i>	[Be25; Ko25; Ma23; MJ24; Yu21; ZT25]	Environmental evidence is not assigned to ITSM roles or authorities. Moderately solvable through role-based governance.
<i>PC4: Ambiguity of direct and indirect effects</i>	[DJ24; Ho24; Os24; Sc23; Vi23]	Direct AI burdens and indirect process effects risk being conflated. Conceptually solvable, empirically difficult.
<i>PC5: Decision operationalisation deficit</i>	[Bo23; GMC24; KB20; Mo25; Si25; VSM23]	Evidence, review points and documentation logic remain unspecified. Moderately solvable through later design work.
<i>PC6: Evidence and capability deficit</i>	[Bo23; Fa24; Ko25; Ma23; Mo25; Ra26; Ro24; ZT25]	Robust indicators, comparable data and interpretive capabilities are insufficiently specified. Partly solvable.

Tab. 1: Problem characteristics (PCs) derived from the SLR

The six problem characteristics reconstruct the focal problem as a set of interrelated translation deficits: environmental evidence needs to be positioned earlier in ITSM decisions (*PC1*), related to explicit lifecycle and system boundaries (*PC2*), connected to ITSM governance responsibilities (*PC3*), differentiated into direct and indirect environmental effects (*PC4*), translated into review points and documentation logic (*PC5*) and supported by comparable evidence and organisational capabilities (*PC6*).

Furthermore, across the reviewed literature, three patterns become visible. AI-enabled ITSM literature explains why organisations adopt AI for automation, service availability, predictive support, resilience, and operational performance, but only weakly connects these decisions to ecological sustainability. Sustainable AI, Green AI, and AI-for-Sustainability literature provide relevant environmental concepts, but often at the level of models, infrastructures, lifecycle phases, or industrial sustainability effects rather than ITSM use-case decisions. Governance- and implementation-oriented literature clarifies adjacent issues such as responsibility, accountability, readiness, and feasibility, but does not yet provide an integrated decision logic for environmentally informed GenAI use in ITSM. The resulting problem statement for this problem analysis echelon can therefore be formulated as follows: *Organisations increasingly consider GenAI-enabled applications for ITSM activities, yet they lack a conceptually and operationally coherent basis for integrating use-case-specific environmental evidence into decisions on introducing, configuring, continuing, modifying or retiring such applications.*

Following the solvability criterion of the problem-analysis echelon [TWV24], this problem is only partially addressed in existing literature but appears moderately solvable because ITSM already provides lifecycle-oriented processes, role-based governance and decision structures into which environmental evidence could potentially be integrated. This indicates plausible starting points for subsequent solution development rather than an already validated solution. However, the reviewed literature does not yet justify the derivation of design requirements or the selection of a specific decision-support artefact.

5 Discussion

The problem analysis suggests that environmentally informed GenAI decision-making in ITSM is not only an unresolved practical challenge, but also reflects a broader conceptual misalignment in the current literature. Environmental AI research often treats impacts as properties of models, infrastructures or lifecycle phases, whereas ITSM research treats decisions as organisational choices embedded in services, roles and governance structures. The central difficulty is therefore not merely to add environmental criteria to ITSM, but to translate evidence produced at one level of analysis into decisions made at another. The relative novelty of organisational GenAI adoption may further explain why this translation problem has only partially been addressed so far, as empirical experience with GenAI-specific ITSM applications is still emerging. However, the review cannot establish this explanation causally and also points to the fragmentation of relevant knowledge. This also

marks the main limitation of the present study. The literature-grounded problem analysis reveals this misalignment and structures its recurring patterns, but it cannot determine how organisations actually prioritise environmental evidence under operational, economic and governance constraints. The apparent solvability of the problem is therefore conditional. ITSM provides promising structures for embedding environmental considerations, but whether these structures can support meaningful decision-making depends on stakeholder responsibilities, data availability and quality, organisational maturity and the willingness to treat ecological sustainability as decision-relevant. The next research step is thus not simply to design a tool, but to validate whether the identified problem characteristics correspond to organisational experience and to determine which form of an artefact can realistically make environmental evidence actionable in GenAI-related ITSM decisions.

6 Conclusion

This paper examined the problem of environmentally informed decision-making regarding GenAI use in ITSM. Based on a SLR, the results show that the problem is not the absence of a relevant knowledge base, but its fragmentation across AI-enabled ITSM, Green IS, Sustainable AI and AI-for-Sustainability research. The analysis identified six interrelated problem characteristics: *temporal integration deficits, lifecycle & boundary fragmentation, an environmental-governance disconnect, ambiguity between direct & indirect environmental effects, decision operationalisation deficits and evidence & capability deficits*. By clarifying the focal construct and formulating a literature-grounded problem statement, the paper contributes a problem-structuring foundation for subsequent DSR, in which the problem statement should be validated with organisational stakeholders before deriving objectives, requirements and an appropriate decision-support artefact.

Bibliography

- [AMB17] Abdelkebir, S.; Maleh, Y.; Belaissaoui, M.: An agile framework for ITS management in organizations. A case study based on DevOps. In Zbitou, J.; Larbi, M.; En-Naimi, E.: ICCWCS '17 Proceedings. Association for Computing Machinery, pp. 1-8, 2017. DOI: 10.1145/3167486.3167556
- [Ax21] AXELOS: ITIL 4. Sustainability in Digital and IT. The Stationery Office, 2021.
- [Be25] Berney, M. et al.: Ethics-by-Design Canvas. A visual inquiry tool to reduce ethical blindness in digital innovation projects. In: ICIS 2025 Proceedings, 1, 2025.
- [Bl23] Blockeel, H. et al.: Decision trees. From efficient prediction to responsible AI. *Frontiers in Artificial Intelligence* 6, 2023. DOI: 10.3389/frai.2023.1124553
- [Bo23] Bosić, Z. et al.: Smart Work Order. In: 2023 46th MIPRO ICT and Electronics Convention (MIPRO), pp. 441-447, 2023. DOI: 10.23919/MIPRO57284.2023.10159836

- [CT11] Cater-Steel, A.; Tan, W.-G.: The role of IT service management in green IT. *Australasian Journal of Information Systems* 17(1), pp. 107-125, 2011. DOI: 10.3127/ajis.v17i1.609
- [DJ24] Di Berardino, D.; Jamshed, K.: Digital transition and sustainability in SMEs. A structured literature review. In: *ITAIS 2024 Proceedings*, 24, 2024.
- [Fa24] Fazlurrahman, H. et al.: E-recruitment with AI in GHRM for corporate university sustainability to improve organization performance. Systematic literature review. In: *CITSM 2024 Proceedings*, pp. 1-5, 2024. DOI: 10.1109/CITSM64103.2024.10775838
- [Ga21] Gailhofer, P. et al.: The role of artificial intelligence in the European Green Deal. *European Parliament*, 2021. DOI: 10.2861/882830
- [GMC24] Grigorev, A.; Mihaita, A.-S.; Chen, F.: Traffic incident duration prediction. A systematic review of techniques. *Journal of Advanced Transportation*, Article 3748345, 2024. DOI: 10.1155/atr/3748345
- [GSF16] Goldberg, M.; Satzger, G.; Fromm, H.: Adapting IT service management for successful multi-sourcing service integration. In: *ECIS 2016 Proceedings*, 2016.
- [HA09] Harmon, R.; Auseklis, N.: Sustainable IT services. Assessing the impact of green computing practices. In: *PICMET 2009 Proceedings*, pp. 1707-1717, 2009. DOI: 10.1109/PICMET.2009.5261969
- [He04] Hevner, A. R. et al.: Design science in information systems research. *MIS Quarterly* 28(1), pp. 75-105, 2004. DOI: 10.2307/25148625
- [He25] Hennig, M. C.: Towards accurate predictions in ITSM. A study on transformer-based predictive process monitoring. In: *ICPM 2024 Proceedings. Lecture Notes in Business Information Processing* 533, 2025.
- [Ho24] Hong, Z. et al.: A data-driven approach for improving energy efficiency in a semiconductor manufacturing plant. *IEEE Transactions on Semiconductor Manufacturing* 37(4), pp. 475-480, 2024. DOI: 10.1109/TSM.2024.3483781
- [IE13] Iden, J.; Eikebrokk, T. R.: Implementing IT service management. A systematic literature review. *International Journal of Information Management* 33(3), pp. 512-523, 2013. DOI: 10.1016/j.ijinfomgt.2013.01.004
- [Ij10] Ijab, M. T. et al.: Seeking the 'Green' in 'Green IS'. A spirit, practice and impact perspective. In: *PACIS 2010 Proceedings*, 2010.
- [KB20] Kaparathi, S.; Bumblauskas, D.: Designing predictive maintenance systems using decision tree-based machine learning techniques. *International Journal of Quality & Reliability Management*, 2020. DOI: 10.1108/IJQRM-04-2019-0131
- [KC07] Kitchenham, B.; Charters, S.: Guidelines for performing systematic literature

reviews in software engineering. EBSE Technical Report EBSE-2007-01, 2007.

[Ko25] Kosasi, S. et al.: Measurement of responsible AI and sustainable AI influence on IT service management. In: 2025 13th CITSM Proceedings, pp. 1-6, 2025. DOI: 10.1109/CITSM67730.2025.11291344

[KOS23] Kotlarsky, J.; Oshri, I.; Sekulic, N.: Digital sustainability in information systems research. Conceptual foundations and future directions. *Journal of the Association for Information Systems* 24(4), pp. 936-952, 2023. DOI: 10.17705/1jais.00825

[Le25] Lee, D. et al.: Towards sustainable AI development. Challenges, opportunities, and the sustainable AI development card. In: UK Academy for Information Systems Conference Proceedings 2025, 22, 2025.

[LMD21] Li, M.; Molla, A.; Duan, S.: AI affordances perception for the transformation of mobility ecosystems. In: ACIS 2021 Proceedings, 76, 2021.

[Ma23] Maccioni, S. et al.: HRM and sustainability at crossroad. A performance perspective from the Italian context. In: ITAIS 2023 Proceedings, 18, 2023.

[MJ24] Marchack, N.; Joseph, E. A.: Centering sustainability in process development through improved characterization of HFC-PFC byproducts. *IEEE Transactions on Semiconductor Manufacturing* 37(4), pp. 453-457, 2024. DOI: 10.1109/TSM.2024.3439271

[Mo25] Mokashi, D. et al.: Lifecycle vulnerability in urban water infrastructure for predictive maintenance planning. *Waterlines* 43(1), pp. 24-33, 2025. DOI: 10.3362/waterlines.v43i1.35

[MT23] MacLean, D.; Titah, R.: Implementation and impacts of IT service management in the IT function. *International Journal of Information Management* 70, Article 102628, 2023. DOI: 10.1016/j.ijinfomgt.2023.102628

[Or14] Orta, E. et al.: Decision-making in IT service management. A simulation based approach. *Decision Support Systems* 66, pp. 36-51, 2014. DOI: 10.1016/j.dss.2014.06.002

[Os24] Osowiecki, W. T. et al.: Achieving sustainability in the semiconductor industry. The impact of simulation and AI. *IEEE Transactions on Semiconductor Manufacturing* 37(4), pp. 464-474, 2024. DOI: 10.1109/TSM.2024.3438622

[Pa21] Page, M. J. et al.: The PRISMA 2020 statement. An updated guideline for reporting systematic reviews. *BMJ* 372, Article n71, 2021. DOI: 10.1136/bmj.n71

[Pe07] Peffers, K. et al.: A design science research methodology for information systems research. *Journal of Management Information Systems* 24(3), pp. 45-77, 2007.

[Ra26] Raina, K. et al.: Artificial intelligence-driven management. Bridging innovation, knowledge creation, and sustainable business practices. *Journal of Innovation &*

Knowledge 11, Article 100860, 2026. DOI: 10.1016/j.jik.2025.100860

[RG25] Rojahn, M.; Grum, M.: Green AI. A systematic review and meta-analysis of its definitions, lifecycle models, hardware and measurement attempts. 2025.

[Ro24] Rowan, W. et al.: Teaching tip. Embedding sustainability in information systems design education. *Journal of Information Systems Education* 35(2), pp. 122-137, 2024. DOI: 10.62273/HBHX1382

[Sc23] Schoormann, T. et al.: Artificial intelligence for sustainability. A systematic review of information systems literature. *CAIS* 52, 2023. DOI: 10.17705/1CAIS.05209

[Si25] Sivakumer, D. et al.: Intelligent automation and predictive ITSM for service optimisation. In: 2025 ICSIT Proceedings, pp. 1-7, 2025. DOI: 10.1109/ICSIT65336.2025.11294337

[SR13] Schneider, J.; Reiter, M.: Green IT-Service-Management. Eine empirische Voruntersuchung der konzeptionellen Grundlagen. In: *Wirtschaftsinformatik 2013 Proceedings*, 110, 2013.

[Su10] Suddaby, R.: Editor's comments. Construct clarity in theories of management and organization. *Academy of Management Review* 35(3), pp. 346-357, 2010.

[TWV24] Tuunanen, T.; Winter, R.; vom Brocke, J.: Dealing with complexity in design science research. A methodology using design echelons. *MIS Quarterly* 48(2), pp. 427-458, 2024.

[UN15] United Nations: Transforming our world. The 2030 agenda for sustainable development. United Nations, New York, 2015.

[Vi23] Vijayakumar, H.: Unlocking business value with AI-driven end user experience management (EUEM). In: *MSIE '23 Proceedings*. Association for Computing Machinery, New York, pp. 129-135, 2023. DOI: 10.1145/3603955.3604004

[VSM23] Vijayakumar, H.; Seetharaman, A.; Maddulety, K.: Impact of AIServiceOps on organizational resilience. In: 2023 15th ICCAE Proceedings, pp. 314-319, 2023. DOI: 10.1109/ICCAE56788.2023.10111409

[WEF24] World Economic Forum: Unlocking value from generative AI. Guidance for responsible transformation. AI Governance Alliance Briefing Paper Series, 2024.

[Yu21] Yuniarto, D. et al.: An alternative model to measure smart space implementing success. In: 2021 9th CITSM Proceedings, pp. 1-6, 2021. DOI: 10.1109/CITSM52892.2021.9588920

[ZT25] Zulkifli, D. U.; Tungkagi, D. Q.: Artificial intelligence and Islamic ethical guidelines. A systematic review. In: 2025 13th CITSM Proceedings, pp. 1-9, 2025. DOI: 10.1109/CITSM67730.2025.11291209

Application of Artificial Intelligence in German Municipal Wastewater Facilities

Nicole Werner ¹, Heiko Thimm ² and Leonie Hammann ³

Abstract: This paper investigates the current state, perceived potential, and barriers to the adoption of AI-based solutions in German municipal wastewater treatment plants, based on an online survey of 105 respondents and ten expert interviews. AI adoption remains limited despite high awareness of its potential: respondents expect the greatest benefits in operational safety, process stability, regulatory compliance, and improved documentation, and existing monitoring systems provide favorable conditions for AI integration, whereas concerns regarding technical reliability, system integration, costs, and IT security hinder wider deployment. Bivariate analyses reveal that awareness increases with plant size, whereas implementation remains consistently low across all facility categories; the perceived benefits are largely consistent regardless of plant size or prior AI experience, with AI-supported detection considered most valuable for process-critical events. To narrow this gap, the paper derives recommendations pertaining to pilot projects, knowledge transfer, and economically viable, easily integrated solutions, and outlines future research directions.

Keywords: Artificial Intelligence in Wastewater Management, AI Adoption Barriers, German Municipal Wastewater Treatment Plants

1 Introduction

Digitalization is increasingly transforming wastewater management. Rising requirements pertaining to efficiency, resource conservation, and the protection of water bodies, an increasing volatility within the sewer network, more frequent heavy rainfall events, and a growing shortage of skilled personnel render data-based approaches increasingly important. Methods of artificial intelligence (AI), in particular, offer considerable potential to optimize operational processes, for instance through improved forecasting, adaptive control, or automated anomaly detection. Particularly since the effectiveness of the European wastewater sector has been called into question [BJBB22], [Euro23].

Within environmental information systems, AI moreover opens up novel possibilities to systematically evaluate large volumes of data from measurement, control, and monitoring systems for operational decision-making. Camera-based AI solutions, for instance, may serve as an additional online monitoring method for the detection of anomalies such as discolorations, foaming, and oil films [APBS24], [ICMB23], [WCLG23].

¹ Pforzheim University, School of Engineering, 75175 Pforzheim, Germany, nicole.werner@hs-pforzheim.de

² Pforzheim University, School of Engineering, 75175 Pforzheim, Germany

³ Pforzheim University, School of Engineering, 75175 Pforzheim, Germany

Notwithstanding this potential, the extent to which AI applications are actually deployed in German wastewater facilities has received scant empirical attention to date. Against this background, a nationwide online survey with more than one hundred participating representatives of municipal wastewater treatment plants was conducted, complemented by expert interviews; the present paper summarizes the principal findings and depicts the current status of AI deployment together with central potentials and barriers.

2 Related Work

A substantial body of international research on AI applications in wastewater management exists. Hocaoglu et al. [HRGM25] differentiate three principal classes of AI-based models: predictive models, detective models, and AI-based optimization models.

Predictive models are concerned with the forecasting of relevant parameters, such as wastewater quality parameters and emissions from treatment plants (e.g. [ShAH20]). Detective models address the identification of faults and the detection of anomalies, whose principal causes are illicit discharges, defects within the treatment system, and leakages in the sewer network. Camera-based AI solutions for anomaly detection have been examined in several research works [APBS24], [MCHJ18] and integrated into initial commercial applications, one exemplary solution being DeepFlow by CETAQUA [Ceta26], a Spanish center for water technology.

AI-based optimization models encompass performance, cost, and multi-objective optimization. One exemplary solution is the digital operating assistant for wastewater treatment plants developed by Nerou [Nero26], a company based in Cologne, which forecasts plant behavior by AI-based means and provides recommendations for the optimization of biological processes. Frequently investigated fields of application further include the prediction of influent volumes and loads [AnDB22], the optimization of energy-intensive processes such as aeration [LWHJ26], and predictive maintenance [XCYJ25]. These studies predominantly demonstrate that significant efficiency gains and environmental relief can be attained.

At the same time, a multitude of publications point out that the transfer into operational practice is associated with challenges [ICMB23], in particular, the availability and quality of suitable data [DIBH26], the integration into existing control systems, and the acceptance among operating personnel [WCLG23]. Questions regarding the transparency and traceability of AI decisions ("explainable AI") are likewise gaining importance within this safety-critical environment [WCLG23]. At the national level, the deployment of AI for wastewater treatment has hitherto been confined to a limited number of research projects and pilot installations, such as LiveSewer [RBFK25] and KikKa [Plaz23]. The empirical evidence pertaining to the actual deployment of AI in German municipal wastewater treatment plants therefore remains deficient, in particular regarding the extent of routine use, the objectives pursued, and the barriers impeding wider dissemination.

3 Survey Methodology

The data collection was conducted by means of an anonymized online survey employing a non-representative sample. The questionnaire was developed in cooperation with the specialists of a wastewater treatment plant (current load size of approximately 171,000 design population equivalents (PE)) and subjected to repeated testing.

Wastewater treatment plants in Germany are classified into five size classes according to their population equivalents (PE), as defined in Annex 1 of the German Wastewater Ordinance [Fede04], based on a standard load of 60 g BOD5 per PE and day. This classification is commonly referenced by the German Association for Water, Wastewater and Waste (DWA). Class 1 plants serve up to 1,000 PE, Class 5 plants more than 100,000 PE, and Classes 2–4 cover intermediate capacities.

Of the approximately 8,900 public wastewater treatment facilities existing in Germany according to the Federal Statistical Office [BdDe24] the email addresses of employees from among the operational management were researched by means of web scraping and manual investigation. This resulted in 507 facilities, which were invited by email to the survey created with the software solution LamaPoll. 301 invitations were dispatched directly to facility employees.

The questionnaire combined closed, semi-open, and scaled questions. The general part encompassed the activity profiles of the respondents, structural company characteristics such as the respective DWA class, and wastewater-related key figures, the degree of digitalization, existing monitoring methods and control mechanisms, and past critical situations. The AI section ascertained the knowledge and deployment of AI-based solutions for technical processes of wastewater treatment, the potential benefits and the concerns towards such solutions, and the relevance of camera-based AI solutions.

4 Survey Results

The following, primarily descriptive analysis provides, by means of frequency analyses and cross-tabulations, an overview of the current status of AI deployment among the respondents. Although the results are not representative of all German public wastewater facilities, they provide indications of initial tendencies. Of the 507 facilities invited, 105 online questionnaires were completed in full. With 56 respondents from facilities serving 10,001 to 100,000 PE (DWA size class 4) and 39 from facilities serving more than 100,000 PE (DWA size class 5), the questions were predominantly answered by large wastewater facilities.

The respondents perform differing functions within their facilities, ranging from executive management and engineering to laboratory, operations, and IT. It may nonetheless be assumed that, they possess the knowledge relevant to the questionnaire. 83 individuals concur at least in part that the technical and organizational tasks pertaining to wastewater

treatment exhibit a high degree of digitalization within their company.

4.1 Company-Wide Use of AI Applications

The survey results reveal a heterogeneous picture with respect to the company-wide use of AI-based applications. Most frequently, other fields of application were named (33 responses), which points to a broad yet imprecisely delineated use. 17 respondents employ AI in the control and/or monitoring of technical installations, and 16 for commercial or organizational tasks, whereas the measurement of wastewater quality has hitherto played a subordinate role (6 responses). It is furthermore noteworthy that 38 respondents possess no knowledge regarding the deployment of AI within their company, which points to a limited transparency and/or internal communication deficits.

4.2 AI Applications for Technical Processes of Wastewater Treatment

The respondents were asked to what extent AI-based solutions for the support and/or optimization of the technical process of wastewater treatment are known to them, or are even already in deployment within their company.

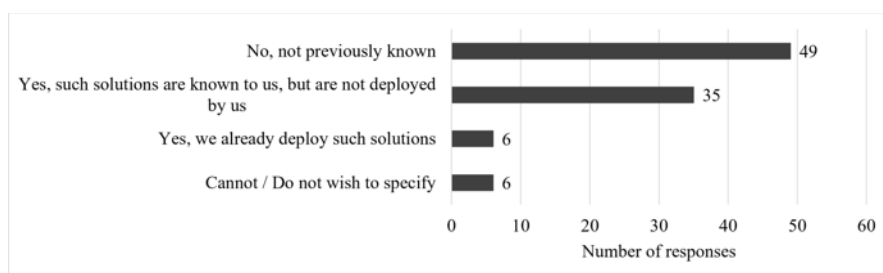


Fig. 1: Awareness and use of AI-based solutions for the support and/or optimization of the technical process of wastewater treatment.

Figure 1 illustrates the distribution of the responses. The largest group (49 respondents) reports that such applications have hitherto been unknown to them, which permits the inference of a distinct information deficit within the sector. 35 respondents are familiar with corresponding solutions yet have not implemented them, and merely 6 of the 96 facilities surveyed already actively employ such technologies. A further 6 respondents provided no information.

4.3 Potential Benefits of AI-Based Applications

The respondents perceive a multitude of potential benefits, whereby multiple responses were possible. Most frequently, the enhancement of operational safety is named (64 responses), followed by the improvement of the process stability (59 responses), the

prevention or reduction of exceedances of limit values (46 responses), and improved documentation and reporting (42 responses). A reduction of manual work steps is anticipated by 34 respondents, whereas further aspects, such as the detection and attribution of problematic indirect discharges, the reduction of false alarms, and environmental aspects, are named less frequently. 10 respondents perceive no advantages.

4.4 AI-Based Monitoring of Critical Situations

Among the challenges faced by wastewater facilities is the timely identification of, and response to, critical situations such as blockages in the influent or sewer network, flooding, exceptional water temperatures, "invisible loads" caused by chemicals or viruses, and discernible loads such as foaming or discolorations. The survey therefore ascertained the process stages at which monitoring measures are carried out (multiple responses possible).

The monitoring predominantly takes place along several process stages: most frequently within the wastewater treatment process itself (76 responses), followed by influent points (52 responses) and direct discharges (48 responses), whereas indirect discharge points play a lesser role (29 responses). Overall, the findings indicate a predominantly established and multi-stage monitoring system.

Critical discharge situations at the influent points, such as fats, dyes, foam, suspended solids, or impermissible discharges are addressed with particular frequency in the literature [WCLG23]. According to the survey, such situations have hitherto occurred rather rarely: 77 of 104 respondents report extremely rare events (one to five cases per year), and a further 10 have never observed such situations, and more frequent occurrences remain the exception. Critical discharges thus do occur, yet do not constitute a regularly arising problem in most facilities. Important methods for the monitoring of the process flow are online measurements of wastewater parameters and automatic samplers. Camera-based AI systems such as DeepFlow [Ceta26] and LiveSewer [RBFK25], by virtue of their capacity for the real-time anomaly detection, may constitute an interesting complement. The respondents were therefore asked how important they deem such systems as complementary monitoring methods for certain critical situations.

The relevance is deemed particularly high for oil films, for which approximately 73% of respondents classify the detection as (very) important, followed by blockages (approximately 54%) and flooding (approximately 53%), whereas for foaming, suspended solids, turbidity, and discoloration a more balanced picture emerges. Immediately process-critical and potentially damage-relevant events thus stand at the focus of perception.

4.5 Barriers

Various research groups arrive, through empirical investigations, at the conclusion that the following barriers, above all, stand in the way of a comprehensive deployment of AI applications in wastewater facilities [ICMB23]: an insufficient availability of data, a lack

of specialist competencies, integration effort, and economic uncertainties. In the study, potential barriers were addressed with a focus on camera-based AI detection solutions. It may nonetheless be assumed that the reported body of opinion may, at least in part, also be transferred to the general deployment of AI. The results demonstrate manifold concerns (multiple responses possible): most frequently, technical reliability and robustness are named (55 responses), followed by the integration into existing IT and infrastructure systems (49 responses), cost and economic-efficiency aspects (48 responses), requirements pertaining to critical infrastructures, in particular IT security (42 responses), and the implementation effort (38 responses). Further aspects are named less frequently, and merely 6 respondents harbor no reservations.

5 Data Analyses and Interview Results

In order to further address the research questions, the survey results are analyzed by means of bivariate analyses of possible relationships between company size and the assessment of potential benefits, fields of application, and barriers. Subsequently, selected insights from 10 expert interviews, each of half an hour in duration, contextualize and deepen the quantitative results.

5.1 Data Analysis

The following evaluations consider the size of the facility (DWA size class) and the prevailing level of awareness and use of AI-based solutions. The size-class analyses are confined to size classes 3 to 5, since the remaining classes were too sparsely represented for a meaningful group-wise interpretation.

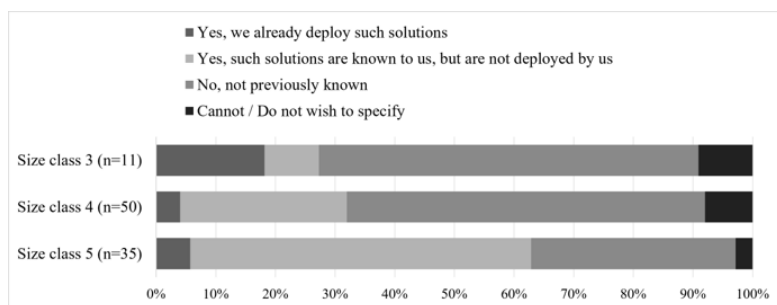


Fig. 2: Awareness and use of AI-based solutions for the technical processes of wastewater treatment, by DWA size class (percentage refers to the share of respondents within each size class; multiple responses possible).

As displayed in Figure 2, the response "not previously known" diminishes as facility size increases: from about 64% in size class 3 (n=11) and about 60% in size class 4 (n=50) to about 34% in size class 5 (n=35), whereas "known but not yet employed" rises markedly, from about 9% to about 57%, and represents the predominant category in size class 5.

Actual deployment ("already in use") remains limited throughout (approximately 4% to 18%). Awareness thus increases with facility size, whereas actual implementation remains low irrespective of size. With respect to the company-wide fields of AI application, a comparable picture emerges: in each class, a considerable proportion of respondents (about 33% to 42%) report having no knowledge of where AI is employed within their company, corroborating the limited internal transparency already observed in Section 4.1, while concrete operational applications remain modest irrespective of facility size. In all three classes, the enhancement of operational safety and the improvement of process stability constitute the most frequently named benefits (each about 50% to 60%), congruent with Section 4.3, and the prevention or reduction of exceedances of limit values likewise features prominently (about 41% to 45%). The core benefit expectations are thus shared across facility sizes, while more specific benefits display a size-dependent emphasis. In order to examine whether the assessment of potential benefits is contingent upon a facility's prior engagement with AI, the respondents were grouped according to their reported awareness and use of AI-based solutions (already in use, $n=6$; known but not employed, $n=35$; not previously known, $n=49$; no specification, $n=6$). The enhancement of operational safety constitutes the most prominently named benefit across the relevant groups (67% among those already deploying such solutions, 63% among those familiar but not yet deploying, and 76% among those not previously aware), closely followed by the improvement of process stability (67%, 69%, and 59%, respectively) and the prevention or reduction of exceedances of limit values (67%, 54%, and 47%, respectively). These core expectations are largely independent of prior exposure to AI. The group already employing such solutions ($n=6$) rates the majority of benefits particularly highly, although the very small size of this group necessitates a cautious interpretation. The perceived potential benefits are thus broadly shared, irrespective of the prevailing level of awareness or implementation.

5.2 Results of the Expert Interviews

The expert interviews reveal a differentiated picture: overall, AI is perceived as a relevant topic for the future. Yet practical deployment has hitherto been predominantly selective and concentrates on supporting applications. A central finding pertains to the underlying data: several experts regard insufficient data quality and a lack of standardization as substantial barriers, whilst one of the greatest potentials of AI is perceived in the automated processing and analysis of large volumes of operational data. Reference is likewise made to the need for uniform data models and for KRITIS- or NIS2-compliant interfaces (i.e. compliant with the German critical-infrastructure regulation and the corresponding EU directive). As relevant fields of application, process and operational optimization are named above all, including the optimization of energy consumption, the dosing of operating supplies, predictive maintenance, and anomaly detection. AI-based monitoring and control systems, for instance on the basis of digital twins or camera-based systems, are likewise assessed as fundamentally promising, and a municipal joint wastewater treatment plant already reports positive experiences with a digital twin for the

intelligent control of stormwater overflow basins. At the same time, several experts emphasize that economic aspects and a comprehensible business case constitute decisive preconditions for adoption, and that municipal wastewater treatment plants, owing to their individual technical characteristics, are amenable to standardization only to a limited extent. In addition to technical challenges, organizational and human factors play an important role, including a lack of understanding of digital concepts, resistance to change, and concerns regarding possible job losses. Owing to the high responsibility for compliance with limit values under water law, the deficient traceability of many AI models (the "black-box problem") is regarded as a central barrier, and individual experts report erroneous decisions of learning-based systems in hitherto unknown operating states.

6 Discussion and Conclusion

The results indicate that the practical implementation of AI-based applications in German municipal wastewater treatment plants is still at an early stage, despite a widespread awareness of their potential. The findings identify structural, economic, and organizational barriers as the main reasons for this discrepancy, in particular technical reliability, integration into existing process control and IT systems, implementation and operating costs, as well as IT security and critical infrastructure requirements. At the same time, respondents expect AI solutions to provide significant benefits, particularly with regard to operational safety, process stability, efficiency, and improved documentation and reporting. The multi-stage monitoring systems already well established in wastewater treatment plants offer favorable conditions for the integration of AI-supported technologies. The bivariate analyses further reveal that awareness of AI increases with plant size, whereas actual implementation remains consistently low across all size classes. The expected benefits, in contrast, are perceived similarly regardless of plant size, prior AI experience, or the occurrence of critical operational events. Based on these findings, future efforts should focus on practical demonstration projects, transparent cost-benefit analyses, and the development of robust, user-friendly AI solutions that can be integrated into existing infrastructures with minimal effort. In addition, targeted training and knowledge transfer are essential to improve staff acceptance and foster confidence in AI technologies. The study is subject to several limitations, including the relatively small and non-representative sample, potential biases resulting from self-reported responses, and varying levels of respondents' knowledge about AI. Furthermore, some subgroup analyses are based on only a limited number of observations and should be interpreted as descriptive and exploratory rather than statistically generalizable. Future research should validate specific use cases under real operating conditions and provide robust assessments of their economic viability. In the long term, AI is expected to become a key enabler of the digital transformation of the sector, with the large volumes of archived operational and process data already available at most treatment plants representing a valuable foundation for training domain-specific AI models and accelerating the development of practical AI applications.

Bibliography

- [AnDB22] ANDREIDES, MARKÉTA ; DOLEJŠ, PETR ; BARTÁČEK, JAN: The prediction of WWTP influent characteristics: Good practices and challenges. In: *Journal of Water Process Engineering* Vol. 49 (2022), p. 103009
- [APBS24] ANTONINI, GIORGIO ; PEARCE, JOSHUA M. ; BERRUTI, FRANCO ; SANTORO, DOMENICO: A novel camera-based sensor for real-time wastewater quality monitoring. In: *Water Practice & Technology* Vol. 19 (2024), No. 9, pp. 3778–3793
- [BdDe24] BDEW ; DESTATIS: *Abwasserdaten Deutschland, Strukturdaten der Abwasserentsorgung*. Berlin : BDEW Bundesverband der Energie- u. Wasserwirtschaft e.V., 2024. URL https://www.bdew.de/media/documents/BDEW_Broschuere_Abwasserdaten_2025_6_Aufl_17072025_1.pdf. - accessed on 2026-04-20
- [BJBB22] BÜTTNER, OLAF ; JAWITZ, JAMES W. ; BIRK, SEBASTIAN ; BORCHARDT, DIETRICH: Why wastewater treatment fails to protect stream ecosystems in Europe. In: *Water Research* Vol. 217 (2022), p. 118382
- [Ceta26] CETAQUA: Detection of hydrocarbons in wastewater and rivers using cameras installed at critical points. Barcelona : CETAQUA Water Technology Centre, 2026. URL <https://www.cetaqua.com/en/proyectos/deepflow-eng/> - accessed on 2026-04-25
- [DIBH26] DE LA HOZ, JEIMY MARTINEZ ; ISLAM, M. D. SHAFIKUL ; BAPPY, MAHATHIR MOHAMMAD ; HAYES, MICHAEL P.: AI-enabled modeling for smart rural wastewater treatment systems: current practices and remaining gaps. In: *Applied Water Science* Bd. 16 (2026), No. 2, p. 40
- [Euro23] EUROPEAN COMMISSION, DIRECTORATE-GENERAL FOR ENVIRONMENT: *Are waste-water treatment plants failing to protect the ecological health of European streams?, 2023-02-15*. URL https://environment.ec.europa.eu/news/are-waste-water-treatment-plants-failing-protect-ecological-health-european-streams-2023-02-15_en. - accessed on 26-06-26 — Science for Environment Policy
- [Fede04] FEDERAL MINISTRY OF JUSTICE (GERMANY): *Anhang 1 AbwV – Häusliches und kommunales Abwasser [Annex 1 AbwV – Domestic and Municipal Wastewater]*. URL <https://www.gesetze-im-internet.de/abwv/BJNR056610997.html>. - accessed on 2026-06-26. — Gesetze im Internet
- [HRGM25] HOCAOGLU, SELDA MURAT ; ROGHANI, BARDIA ; GULCAN, HANDE ; MAGNA, DÉBORA JARETA ; AYDÖNER, CIHANGIR ; BARROS, VIRGINIA GRACE ; KOYUNLUOGLU AYNUR, SEBNEM ; RATNAWEERA, HARSHA ; ET AL.: Artificial intelligence in wastewater treatment plants: a review of current trends and adaptation strategies in the face of climate-driven rainfall fluctuations. In: *Journal of Water and Climate Change* Vol. 16 (2025), No. 9, pp. 2742–2759
- [ICMB23] IMEN, SANAZ ; CROLL, HENRY C. ; MCELLELLAN, NICOLE L. ; BARTLETT, MARK ; LEHMAN, GENO ; JACANGELO, JOSEPH G.: Application of machine learning at wastewater treatment facilities: a review of the science, challenges and barriers by level of implementation. In: *Environmental Technology Reviews* Vol. 12 (2023), No. 1, pp. 493–516
- [LWHJ26] LI, QUNLI ; WANG, KAI ; HORACE, GENEVIEVE FAITH ; JIANG, PEITING ; XIONG, YANYAN

- ; HUANG, FUJIAN ; WANG, YONG ; XU, WENLAI ; ET AL.: Research on an intelligent precise aeration control system for wastewater treatment based on LSTM models. In: *Journal of Water Process Engineering* Vol. 85 (2026), p. 109746
- [MCHJ18] MULLINS, DARRAGH ; COBURN, DEREK ; HANNON, LOUISE ; JONES, EDWARD ; CLIFFORD, EOGHAN ; GLAVIN, MARTIN: A novel image processing-based system for turbidity measurement in domestic and industrial wastewater. In: *Water Science and Technology* Vol. 77 (2018), No. 5, pp. 1469–1482
- [Nero26] NEROU: nerou elements. Die moderne Steuerung von Kläranlagen. Köln : nerou GmbH, 2026. URL <https://www.nerou.de/landing/produkt.html>. - accessed on 2026-04-25
- [Plaz23] PLAZZO, MICHAELA: BMUV fördert KI-Leuchtturmprojekt zu klimaneutralen Kläranlagen. In: *EUWID Wasser und Abwasser* Vol. 2023. Gernsbach (2023), No. 28. URL <https://www.euwid-wasser.de/news/wirtschaft/bmuv-foerdert-ki-leuchtturmprojekt-zu-klimaneutralen-kläranlagen-290623/>. - accessed on 26-06-26
- [RBFK25] ROHRBACH, MICHAELA ; BÖDECKER, JÖRG ; FUHRMANN, TIM ; KLÖCKER, PAULA ; LAU, PHILIP ; ZIMMERMANN, MARTIN: Kanalmonitoring mit KI - vom Potenzial zur Umsetzung. In: *gwf-Wasser/Abwasser International* Vol. 11. Essen, Vilkan Verlag (2025), No. 166, pp. 65–73
- [ShAH20] SHARAFATI, AHMAD ; ASADOLLAH, SEYED BABAK HAJI SEYED ; HOSSEINZADEH, MAJID: The potential of new ensemble machine learning models for effluent quality parameters prediction and related uncertainty. In: *Process Safety and Environmental Protection* Vol. 140 (2020), pp. 68–78
- [WCLG23] WANG, YI ; CHENG, YUHAN ; LIU, HE ; GUO, QING ; DAI, CHUANJUN ; ZHAO, MIN ; LIU, DEZHAO: A Review on Applications of Artificial Intelligence in Wastewater Treatment. In: *Sustainability* Vol. 15 (2023), No. 18, p. 13557
- [XCYJ25] XIONG, ZHAOYANG ; CHENG, LEIMING ; YUNUS, AHMED IBRAHIM ; JIAN, DEWU ; ZHANG, XINPING ; IGOU, THOMAS ; TONG, SHA ; CHEN, YONGSHENG: Machine Learning-Assisted Expert Control in Wastewater Aeration Treatment Process Application. In: *ACS ES&T Engineering* Vol. 5 (2025), No. 11, pp. 3140–3152

Doctoral Workshop

Assessing the Maturity of Sustainability-aware Data Management Strategies that minimise Data Volumes and reduce Energy Consumption: A Study Approach

Engelbert Fellingner ¹ and Patricia Lago ²

Abstract: Digital sustainability can be considered along four core dimensions: technical, social, economic, and environmental. We found a significant gap in research on assessing data management (DM) strategies aimed at minimising data volumes and reducing energy consumption. In this paper, we present a mixed-method study designed to create the SusDMMA-DVE model. Aimed to assess the maturity and effectiveness of sustainability-aware DM strategies, this model can help identify DM challenges, such as excessive data collection, thereby reducing energy consumption and operational costs while meeting business needs.

Keywords: Data management strategies, sustainability, energy efficiency, carbon dioxide emissions, data growth, data management maturity assessment

1 Introduction

Data is a valuable resource that provides insight and informs decision-making. However, excessive data collection can be counterproductive as it hinders business processes, consumes large amounts of energy, and increases costs for data processing and storage. As more data is generated at an exponential rate, concerns about growing energy demands and carbon dioxide emissions are increasing [Ca24]. According to research commissioned by Veritas, only 15% of organisational data is business-critical [Ve16]. Research by Netapp states that a significant two-fifths (41%) of the data stored by UK organisations is unused or unwanted, leading to unnecessary costs and emissions from storage and energy consumption [Wa23]. Data volumes can be minimised by applying sustainability-aware data management (DM) strategies such as compression, deduplication, and data classification that have different characteristics in terms of use and potential energy-savings [Fe26; FL25].

We define *sustainability-aware data management maturity* as the degree to which an organisation manages and improves its data processes to address sustainability goals, achieving both effectiveness and efficiency; this involves ensuring that data practices are not only efficient, cost-effective, legally compliant, secure, and measurable, but also minimise environmental impact, conserve resources, and support overall organisational sustainability goals.

¹ Vrije Universiteit Amsterdam, De Boelelaan 1111, 1081 HV Amsterdam, The Netherlands, e.h.fellinger@vu.nl,  <https://orcid.org/0009-0001-4331-5825>

² Vrije Universiteit Amsterdam, De Boelelaan 1111, 1081 HV Amsterdam, The Netherlands, p.lago@vu.nl,  <https://orcid.org/0000-0002-2234-0845>

Digital sustainability can be considered along four core dimensions: technical, social, economic, and environmental [La15], see Table 1. At its core, a Data Management Maturity Assessment (DMMA) is a systematic process used to evaluate an organisation's DM capabilities against established best practices. Although existing DMMA models, such as DAMA DMMA [DA24] and DCAM [ED26] address various aspects of DM and data governance, we observe a notable gap in research on assessing DM strategies that minimise data volumes and reduce energy consumption while addressing digital sustainability. Our study aims to fill this gap by introducing a model to assess sustainability-aware data DM strategies that address these challenges.

Maturity models describe a progression through levels of process characteristics. By understanding these characteristics, an organisation can evaluate its maturity level and create a plan to improve its capabilities[DA24]. A typical *staged* maturity model, such as the DAMA DMMA, assesses how organisations manage information assets across progressive levels [DA24]. They can be used to evaluate overall DM efforts, or can be used to focus on a single knowledge area or a single process [DA24]. A commonality of DMMA models is the measurement of current and future state [Be21]. Our novel contribution to sustainability-aware DM is using the four core digital sustainability dimensions to define goals, along with CSFs, KPIs, metrics, and measurements to reveal different levels of maturity and effectiveness in reducing data volumes and energy consumption. Our proposed Sustainability-aware Data Management Maturity Assessment model for minimising Data Volumes and Energy consumption (SusDMMA-DVE) can function as a stand-alone or as a supplement to widely used DM frameworks. This is achieved by linking our model to common areas of knowledge in DM, such as Data Governance, Data Architecture, Data Storage and Operations, Data Security, and Data Quality.

Our work in progress builds on previous research by Fellingner and Lago [FL25] that identified nine types of sustainability-aware DM strategies that are applicable in the cloud, data centres, and edge/fog. It also includes our previous research on factors that support or hinder the deployment of sustainability-aware DM strategies [Fe26]. Both studies inform the interviews and case studies that underpin the development of the proposed SusDMMA-DVE model.

A goal is defined as a long-term outcome that teams use to define strategies [FFL25]. Their CSFs are elements that determine long-term success, acting as a focal point for resource allocation, strategy execution, and performance measurement [OC25]. KPIs are the metrics used to measure performance according to the CSFs [OC25]. Based on Lago et al. [La15], DAMA DM-BOK [DA24], ISO25010:2023 [In23], Fellingner [Fe26], and the Cloud Sovereignty Framework of the European Commission [Co25], Table 1 defines the four core dimensions, goals, and CSFs of sustainability-aware DM that we will supplement with the findings of the literature study to build our Preliminary model. The KPIs are specific for the chosen DM strategy; they will be measured in the company context.

Our study design aims to produce the following main contributions:

Core dimension	Goal	CSF	KPI
Technical	Long-term sustainability of data systems and their architectures as they evolve while minimising data volumes and reducing energy consumption	Maintaining product performance within set time and throughput parameters while using resources such as CPU, memory, storage, and bandwidth efficiently along the data pipeline	In company context
Economic	Reducing costs and generating business value with data, ensuring data quality throughout its lifecycle while minimising data volumes and reducing energy consumption.	Lowering operational costs through energy efficiency and reducing data volumes, improving data quality, and maintaining agreed service levels.	In company context
Social	Adhering to laws and data governance principles while minimising data volumes and reducing energy consumption.	Attention to behaviour, risk control, legal compliance (GDPR), security while minimising data volumes and reducing energy use along with sovereignty of data assets and AI services.	In company context
Environmental	Addressing the ecological footprint of data systems while minimising data volumes and reducing energy consumption.	Reducing energy consumption, lowering carbon dioxide emissions, minimising data collection, data storage, and data transmissions	In company context

Tab. 1: Core digital sustainability dimensions, goals, critical success factors (CSFs) and key performance indicators (KPIs).

- A model to assess the maturity and effectiveness of DM strategies that minimise data volumes and reduce energy consumption while balancing the four core dimensions of digital sustainability.
- Our proposed model can function as a stand-alone or as a supplement to other main DM frameworks.

Paper structure Section 2 covers the background and related work. Section 3 details the design of our mixed-method study. Section 4 outlines potential validity threats. Section 5 states the conclusion.

2 Related work

Belghith et al. [Be21] reviewed 20 maturity models related to DM and data governance. Through a comparative analysis of the main concepts and features of these models, a meta model was developed that can serve as a tool for researchers to position future maturity models. However, information on older DMMA models and outdated references highlight the need for ongoing research in this area. Mettler [Me11] outlines the typical phases involved in the development and application of a maturity model from a design science research (DSR) perspective. Using artificial and naturalistic evaluations, we incorporate DSR in a mixed-method study design to define, iteratively refine, and validate our proposed model during development. Lautenschutz et al. [La18] reviewed green ICT maturity models to improve the social or environmental impact of ICT. They applied the knowledge gathered to extend an existing maturity model with the intention of applying it during an ICT sustainability audit in higher education. Their proposal covers a broad range of topics related

to energy and material efficiency. In contrast, our research focuses on reducing data volumes and energy consumption. Pörtner et al. [Pö25] researched the maturity levels of data-related challenges during the transition to a data-driven organisation. They state that existing maturity models for digital transformation, data management, and data-driven organisations lack a “comprehensive, industry-agnostic, and practically validated approach to addressing industry challenges”. Corallo et al. [Co23] researched measuring the level of maturity in big data management in manufacturing companies to assess their current status and define future actions in five areas: organisation, DM, data analysis, infrastructure, governance & security. Their DM area has three sub-dimensions: data source, data processing, and data storage, reflecting the data pipeline architecture. However, environmental digital sustainability concerns are not addressed. Yan et al. [YJ26] provides a critical evaluation of 28 data management capability maturity models published over the past decade. The authors highlight several gaps, including an often insufficient coverage of the entire data lifecycle, particularly in relation to data deletion. The Cloud Sovereignty Framework of the European Commission [Co25] aims to protect and ensure control over data and AI services within the European Union (EU). While the EU Cloud Sovereignty Framework is designed for procurement, it may influence DM decision-making within the EU in the years to come. Building on the findings of previous studies, our research aims to address key DM issues related to the four core dimensions of digital sustainability.

3 Study design

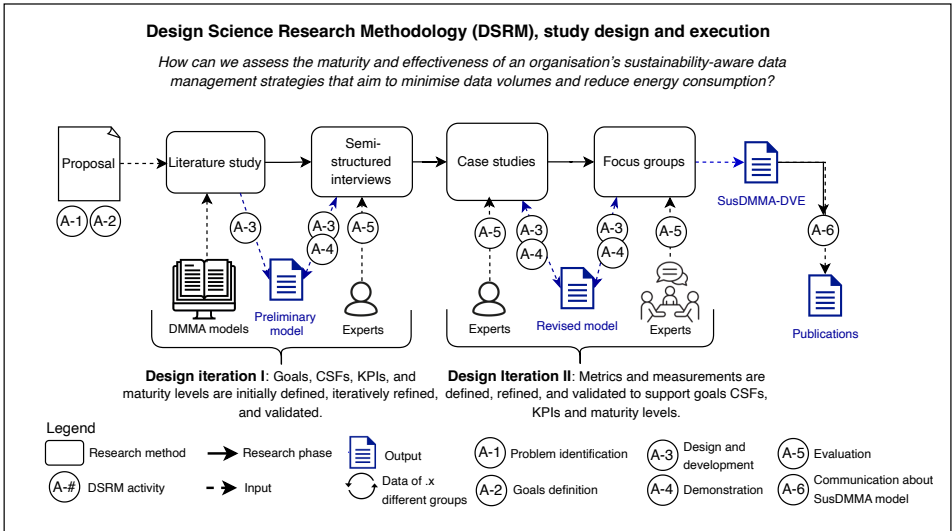


Fig. 1: Study design and execution

In this section, we describe the design of our mixed-method study to develop our proposed SusDMMA-DVE model (see Figure 1). This includes the goal and research question, the

collection of data through the literature study, interviews, focus groups, and data analysis. To guide the iterative development of our SusDMMA-DVE model, we use the Design Science Research Methodology (DSRM) in information systems as proposed by Peffers et al. [Pe07]. As shown in Figure 1, the stage of problem identification is covered in this proposal along with the definition of goals. Our proposed SusDMMA-DVE model evolves between two design iterations. During Design Iteration I, the Preliminary model and its goals, CSFs, KPIs, and initial maturity levels are defined through a literature study, then refined and validated with experts. A self-assessment questionnaire is delivered to evaluate sustainability-aware DM strategies aimed at reducing data volumes and energy consumption. Design Iteration II involves the iterative development of a toolkit of metrics and measurement methods through case studies to support goals, CSFs, KPIs, and maturity levels.

In conducting our literature study, we follow the systematic review search methods for grey literature outlined by Godin et al. [Go15]. The guidelines by Wohlin et al. [Wo24] are used to inform the design of the semi-structured interviews and case studies. For the design of the focus groups, we use the guidelines provided by Kontio et al. [KBL08].

3.1 Research goals and research question

The Goal Question Metric (GQM) strategy [BCR94] is used to formulate our research goal: **to analyse** goals, success factors, KPIs, metrics and measurements of DM strategies that focus on minimising data volumes and reducing energy consumption **for the purpose of** understanding their maturity **with respect to** technical, economic, social, and environmental goals **from the point of view of** researchers and practitioners **in the context of** an organisation.

To accomplish the research goal, the study addresses the following RQ:

How can we assess the maturity and effectiveness of an organisation's sustainability-aware DM strategies that aim to minimise data volumes and reduce energy consumption?

3.2 Literature study

To identify current literature on the rapidly evolving field of DMMA models from their developers, accredited partners, users, or critics, we intend to conduct a grey literature search. For a manageable number of search results and to minimise bias from a short time frame while reducing the risk of analysing outdated information, we propose a five-year period for our grey literature search. Grey literature searches can introduce bias in the results due to automated relevance ratings. To mitigate this threat, we will supplement the grey literature with related work (see Section 2), and use the widely recognised DAMA DMMA [DA24] as a reference point.

Identification: we perform our searches using regular Google, employing the search terms “data management maturity” to identify relevant websites that publish documentation on DMMA models. A filter will be applied to the search to capture the results of the last five years. The first ten pages of the search results are captured using the Google Search Results Scraper extension for Google Chrome [Go99]. In this step, the websites are identified and their URL will be entered into an Excel spreadsheet. Next, each of these websites will be manually searched to remove off-topic data, and to identify potentially relevant information on DMMA. The remaining potentially relevant records continue to the screening and selection process. During this step, the name of the website, DMMA model, URL, and year are recorded in an Excel spreadsheet.

Screening and selection process: our review begins with reading the available abstracts, executive summary, table of contents, or introductions to assess relevance to the research objectives using inclusion and exclusion criteria.

Inclusion criteria: I1: Topic: DMMA models I2: Language: written in English I3: Current: the content on DMMA models must have been developed, maintained, or revised within the past five years. I4: Accessible: DMMA documentation must be accessible to researchers. I5: Detail: DMMA documentation should provide details on goals, the knowledge or capability areas assessed, specify progressive levels, metrics, and the assessment process.

Exclusion criteria: E1: Multimedia content: we limit our research to text-based sources. E2: Duplicates: these are removed to avoid invalid conclusions or results. E3: Specific focus: models that focus on assessing a specific capability area, such as data analytics, governance, or security, are excluded.

The full texts of the items that will go to the second stage are then examined. If there is any uncertainty about an item’s eligibility, the reviewer will decide to proceed with further screening. The second author will review the selection using a random sample, ensuring that the set criteria were applied objectively and correctly by the first author.

Data collection process and synthesis of results: the remaining records are analysed using thematic analysis [BC06]. Our goal is to extract a superset of features that can be used to assess capabilities related to minimising data volumes and reducing energy consumption toward building the Preliminary version of our proposed SusDMMA-DVE model.

Results: the results of the literature study and our previous research on sustainability-aware DM strategies [FL25], and their supporting and hindering factors [Fe26], inform the first design iteration of the Preliminary model.

3.3 Interviews

The Preliminary model will be used to evaluate the qualitative data from the expert interviews and will be iteratively updated based on the feedback from the interviewees (i. e. expert

evaluation). For these interviews, we select experts from companies that are members of the Dutch National Coalition of Sustainable Digitalisation. We interview those who play key roles related to DM and DMMA, such as IT managers, domain experts, data architects, data engineers, and enterprise architects, based on availability (convenience sampling). We plan to use *semi-structured interviews* to balance structured data collection with the flexibility to explore participant insights in depth using the 5W (Who, What, When, Where, Why) and 1H (How) method. The list of DMMA interview topics and questions will be finalised after completing the literature study and creating the Preliminary model.

Informed by our previous work [Fe26; FL25], the literature discussed in the Introduction Section (1) and the Related Work Section (2), we propose to address the following topics for our semi-structured interviews: *Vision and business goals*: To gain an understanding of what our DMMA must support, we begin our inquiry with eliciting the *vision and business goals* in relation to the four core digital sustainability dimensions' goals and CSFs. *Scoping*: After an overview of the situation is made, we determine with the participants the scope of the DMMA: general initiatives, specific branches or departments, knowledge areas, processes. *Detail DM strategies*: To get to the core of the DMMA, we ask whether any DM strategies related to data volume minimisation or energy-savings are applied [FL25]. In addition, we ask for any other sustainability-aware DM efforts that may be relevant. *Data pipeline*: To identify possible overhead in data lineage, we will place the identified DM strategies along the data pipeline; data source, a data processing or data transformation stage, and a data destination or data storage location [HP]. *Factors that support and hinder*: To enrich the Preliminary model with contextual information, we ask about factors that support or hinder the achievement of the goals and CSFs of the four core dimensions of digital sustainability 1.

Data analysis: To identify, analyse and report patterns (themes) within the transcripts from the interviews, we use thematic analysis [BC06] to further refine the output of Design Iteration I. The results are then plotted as capability levels on a radar chart to visually compare them with goals and ambitions across various dimensions.

3.4 Case studies

Design iteration II contains case studies; empirical methods aimed at investigating contemporary phenomena in their context [RH08]. To build the Revised model with participants, we will use case studies to collect quantitative data on effectiveness by measuring KPIs before and after deploying sustainability-aware DM strategies. We propose the following steps, which will be further detailed during the research: **Define metrics**: First, we define metrics to measure the current state of DM strategies relating to the four dimensions' goals, CSFs, and KPIs. We also agree with the participants on the desired levels that will be compared to the actual levels after the measurements. **Measurements**: Participants perform measurements of DM strategies and their KPIs according to agreed and set methods to ensure internal validity. **Analysing data**: Excel is used to record and analyse the quantitative data on KPIs

to substantiate our findings. Our goal is to expose current strengths and opportunities for improvement, rather than solving problems. **Data visualisation:** The metrics are plotted as proficiency levels on a radar chart to visually compare the actual maturity with the desired levels and plan improvements to the effectiveness of current DM strategies.

3.5 Focus groups

For the focus groups, we invite experts who play key roles in DM or DMMA (expert sampling). Through the focus group sessions, we aim to further improve our Revised model and its practical application. We will address the suitability of the model for operational use, possible implementation challenges, and suggestions to improve its user-friendly deployment [KBL08]. Data will be captured through observer notes and affinity maps from the sessions. To mitigate validity risks from group dynamics, moderator bias, and contextual factors, we will carefully select participants, train the moderator, and incorporate focus groups into a multi-method research design. After processing the input from the focus groups, design iteration II is completed.

3.6 SusDMMA-DVE

Our proposed SusDMMA-DVE model aims to integrate qualitative and quantitative evidence to determine levels of maturity and effectiveness to reduce data volumes and energy consumption. It consists of a self-assessment questionnaire, descriptions of maturity levels, and a toolkit of metrics and measurement methods to systematically support the findings. We intend to provide two levels of detail to accommodate both managerial overview and specialist needs. To ensure that the proposed model is applicable across multiple types of organisation, we focus on the characteristics of the data involved, rather than the specific industries of the case studies, such as education, banking, retail, or healthcare.

4 Threats to validity

In this section, we discuss the most prominent threats to the validity of our study and how they are mitigated in the study design. We follow the classification of validity threats as discussed by Wohlin et al. [Wo24], based on the work of Cook and Campbell [CC79]. **Construct Validity:** to mitigate the risk of “Inadequate Preoperational Explication of Constructs” [CC79], we use clear definitions of goals, CSFs, and measurement subjects throughout our research. **Internal Validity:** to ensure that any changes are solely due to DM strategies that reduce data volumes and energy consumption, we will clearly define metrics and measurement methods with the participants in advance and calibrate these during our research. **External Validity:** in our interviews and case studies, we use convenience

sampling, which limits the ability to draw statistically representative samples. To enhance the generalisability of our findings, we will employ common DM knowledge areas throughout the data pipeline architecture. **Reliability:** to mitigate researcher bias, we will follow a systematic protocol for our literature study, interviews, case studies, and focus groups. We will also provide a public replication package to enable verification of our findings.

5 Conclusion

Our proposed mixed-method study design aims to create the SusDMMA-DVE model, which allows professionals to assess sustainability-aware DM efforts aimed at minimising data volumes and reducing energy consumption. Through gathering qualitative and quantitative evidence, the model can help identify issues such as excessive data collection, processing, and storage hotspots, ultimately reducing resource use and operational costs while meeting business needs.

References

- [BC06] Braun, V.; Clarke, V.: Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), pp. 77–101, 2006.
- [BCR94] Basili, V. R.; Caldiera, G.; Rombach, D. H.: *The Goal Question Metric Approach*. John Wiley & Sons, 1994.
- [Be21] Belghith, O. et al.: A Survey of Maturity Models in Data Management. In: *2021 IEEE 12th International Conference on Mechanical and Intelligent Manufacturing Technologies (ICMIMT)*. Pp. 298–309, 2021.
- [Ca24] Castro, V. et al.: Digital data demand and renewable energy limits: Forecasting the impacts on global electricity supply and sustainability. *Energy Policy*, vol. 195, p. 114404, 2024.
- [CC79] Cook, T. D.; Campbell, D. T.: *Quasi-experimentation*. Houghton Mifflin, Boston, MA, 1979.
- [Co23] Corallo, A. et al.: Evaluating maturity level of big data management and analytics in industrial companies. *Technological Forecasting and Social Change* 196, p. 122826, 2023.
- [Co25] Commission, E.: *Cloud Sovereignty Framework Version 1.2.1 – Oct. 2025*, tech. rep., European Commission Directorate-General for Digital Services, 2025.
- [DA24] DAMA: *DAMA-DMBOK2 Revision: Data Management Body of Knowledge*. DAMA International, 2024.
- [ED26] EDM Association: *DCAM Framework*, Available Online: <https://edmcouncil.org/frameworks/dcam/> (Accessed 11th March 2026), 2026.
- [Fe26] Fellingner, E.: Assessing the factors that support and hinder sustainability-aware data management strategies. In: *Joint Proceedings of ICT4S 2026 Doctoral Symposium, Demonstrations & Posters Track and Workshops*. CEUR Workshop Proceedings, in press, CEUR-WS.org, 2026.
- [FFL25] Fatima, I.; Funke, M.; Lago, P.: Providing Guidance to Software Practitioners: A Framework for Creating Key Performance Indicators. *IEEE Software* 42 (4), pp. 68–78, 2025.

- [FL25] Fellingner, E.; Lago, P.: Sustainability-driven Data Management Strategies. English, ACM Computing Surveys, (under review), 2025, <https://hdl.handle.net/1871.1/639baf52-e8c7-4946-89ea-55a6021913b3>.
- [Go15] Godin, K. et al.: Applying systematic review search methods to the grey literature: a case study examining guidelines for school-based breakfast programs in Canada. *en, Syst. Rev.* 4 (1), p. 138, 2015.
- [Go99] Google: Google Search Results Scraper, (Accessed May, 1 2026), 9999, <https://chromewebstore.google.com/detail/google-search-results-scr/cmdkhofenchpdlkbpkhdhdfgfdmmobcc?hl=en&pli=1>.
- [HP] HP: Data pipelines, (Accessed May, 1 2026), <https://www.hpe.com/nl/en/what-is/data-pipelines.html>.
- [In23] International Organization for Standardization: ISO/IEC 25010:2023 Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Product quality model, Available Online: <https://www.iso.org/standard/78176.html>, 2023.
- [KBL08] Kontio, J.; Bragge, J.; Lehtola, L.: The Focus Group Method as an Empirical Tool in Software Engineering. In (Shull, F.; Singer, J.; Sjøberg, D. I. K., eds.): *Guide to Advanced Empirical Software Engineering*. Springer London, London, pp. 93–116, 2008.
- [La15] Lago, P. et al.: Framing sustainability as a property of software quality. *Commun. ACM* 58 (10), pp. 70–78, 2015.
- [La18] Lautenschutz, D. et al.: A Comparative Analysis of Green ICT Maturity Models. In (Penzstadler, B. et al., eds.): *ICT4S2018. 5th International Conference on Information and Communication Technology for Sustainability*. Vol. 52. EPiC Series in Computing, EasyChair, pp. 153–167, 2018.
- [Me11] Mettler, T.: Maturity assessment models: a design science research approach. *Int. J. Soc. Syst. Sci.* 3 (1/2), p. 81, 2011.
- [OC25] O'Connor, S.: Critical success factors explained: a complete 2026 strategy guide, Available Online: <https://monday.com/blog/project-management/critical-success-factors/> (Accessed 9th March 2026), 2025.
- [Pe07] Peffers, K. et al.: A Design Science Research Methodology for Information Systems Research. *Journal of Management Information Systems* 24 (3), pp. 45–77, 2007.
- [Pö25] Pörtner, L. et al.: Data Management Maturity Model—Process Dimensions and Capabilities to Leverage Data-Driven Organizations Towards Industry 5.0. *Applied System Innovation* 8 (2), 2025.
- [RH08] Runeson, P.; Höst, M.: Guidelines for conducting and reporting case study research in software engineering. *Empirical Software Engineering* 14 (2), pp. 131–164, 2008.
- [Ve16] Veritas: The Databerg Report: see what others don't identify the value, risk and cost of your data, tech. rep., Veritas, 2016, https://www.veritas.com/content/dam/Veritas/docs/reports/veritas-strike-global-report_a4-sdc2.pdf.
- [Wa23] Watts, M.: Data Waste Index, Available Online: "https://www.netapp.com/pdf.html?item=/media/83312-data_waste_research_report_final_for_submission.pdf", 2023.
- [Wo24] Wohlin, C. et al.: *Experimentation in software engineering*. Springer, Berlin, Germany, 2024.
- [YJ26] Yan, S.; Jones, J.: Bridging the Gaps: A Comprehensive Review of Data Management Capability Maturity Models in the Digital Era. *IEEE Transactions on Big Data* 12 (2), pp. 307–320, 2026.

Linking Water Quality to Corrosion Behaviour at Steel Sheet Pile Walls: A Machine Learning Study in the Port of Hamburg

Sarah Flohr ¹, Paul Bruck ², Jan Kleinmann ¹, Grit Behrens ¹, and Andreas Kahlfeld²

Abstract: Corrosion of steel sheet pile walls poses a significant challenge for the maintenance of port and waterway infrastructure. This work investigates the relationship between water quality parameters measured at monitoring stations in the Port of Hamburg and site-specific corrosion behaviour, with the aim of identifying environmental drivers of material loss at hydraulic steel structures. A Random Forest model trained on water quality means, structure age, and spatial assignment groups achieves a cross-validated R^2 of 0.719 on an age-adjusted corrosion residual, confirmed by a hold-out R^2 of 0.680. SHAP-based analysis indicates that predictive performance is largely attributable to structure age and spatial group membership. This is consistent with the hydrological data structure: water quality varies between monitoring regions but remains constant within them, which limits the contribution of individual water quality parameters to site-specific corrosion prediction and points to the need for spatially finer-grained input data.

Keywords: Corrosion Prediction, Steel Sheet Pile Walls, Water Quality, Random Forest, Explainable AI

1 Introduction

Steel sheet pile walls are structural elements of port, canal, and coastal infrastructure. They serve to retain river banks, delimit harbour basins and locks, stabilise terrain, and provide flood protection, making them an indispensable component of federal waterway infrastructure. Due to their permanent exposure to water, sediment, and varying chemical conditions, steel sheet pile walls are subject to continuous corrosion [Bu22; HAH07]. The resulting material loss reduces the remaining wall thickness and, over several years, compromises the structural integrity and serviceability of these structures. If the residual wall thickness falls below critical thresholds, the consequences may extend to the stability of adjacent structures; in the event of wall failure, retained soil can escape and contaminants such as oil may leach directly into the waterway, entailing considerable risk of damage and rehabilitation costs.

The study area of this work is the Port of Hamburg. Within the harbour basin, sheet pile walls

1 Hochschule Bielefeld (HSBI), Angewandte Informatik, Artilleriestraße 9, 32427 Minden, Germany, sarah.flohr@hsbi.de,  <https://orcid.org/0009-0009-9901-7841>; jan.kleinmann1@hsbi.de,  <https://orcid.org/0009-0009-2185-5765>; grit.behrens@hsbi.de,  <https://orcid.org/0009-0009-0247-8204>

2 Hochschule Bielefeld (HSBI), Wasser- und Verkehrsbau, Artilleriestraße 9, 32427 Minden, Germany, paul.bruck@hsbi.de,  <https://orcid.org/0009-0005-8963-3182>; andreas.kahlfeld@hsbi.de

are exposed to the varying physical and chemical conditions of a tidally influenced system, where elevated corrosion rates pose a particular challenge for operation and maintenance. Direct corrosion assessment relies on periodic wall thickness measurements, which are both cost- and labour-intensive, resulting in spatially and temporally sparse data that preclude comprehensive condition monitoring. At the same time, continuous water quality time series are available from monitoring stations within the harbour basin, covering parameters such as dissolved oxygen, pH, electrical conductivity, and chloride concentration. This raises the central question of the present work: to what extent can publicly available water quality data help explain site-specific corrosion behaviour of sheet pile walls, and can they support data-driven corrosion assessment? The contribution of this work is a spatial-resolution diagnosis, showing that water quality data capture corrosion behaviour at the group level but not site-specific differences within groups, with direct consequences for monitoring-network design and maintenance planning.

2 Related Work

The corrosion of steel sheet pile walls represents a central challenge for the safety, durability, and economic viability of port and waterway infrastructure. Corrosion in aquatic environments is driven by a combination of physical, chemical, and biological factors (including, for example, dissolved oxygen availability, ionic conductivity, chloride concentration, and pH value [Me06]), interacting in coupled, partly non-linear ways that are difficult to capture through simplified empirical models alone.

The temporal evolution of corrosion loss is commonly described by a power-law model relating cumulative material loss to structure age, widely applied to steel structures under long-term immersion conditions [Ca18; Me03; Me14] and adopted in engineering standards for hydraulic steel structures [Bu22; DI10]. These approaches generally assume stationary boundary conditions and spatially aggregated variables, limiting their capacity to capture local variation and dynamic environmental changes [Ch17]. Climate change further compounds this limitation, as shifting temperature and salinity regimes alter the boundary conditions of corrosion and question the transferability of established empirical values [Fr21; PI22].

A promising alternative lies in exploiting continuously available water quality data. Monitoring stations along federal waterways routinely provide high-resolution time series for relevant parameters, and unlike direct corrosion measurements, which are costly and spatially sparse [HAH07; Me25], these data are available at scale without additional measurement effort.

Machine learning has emerged as a powerful tool for corrosion rate estimation. Yan et al. [Ya20] applied a Random Forest model to marine atmospheric corrosion data, identifying chloride deposition rate, precipitation, and relative humidity as dominant drivers ($R^2 = 0.73$). Yang et al. [Ya24] extended this to depth-varying marine exposure using Gradient Boosting on 856 field samples, finding dissolved oxygen and relative humidity most critical. A broader review by Coelho et al. [Co22] confirms that ensemble methods

consistently outperform classical regression approaches across pipelines, offshore structures, and marine steels [BHZ22; IJM24].

A consistent limitation is the lack of model interpretability. Yang et al. [Ya24] note that certain inputs improve accuracy yet lack physical plausibility, illustrating that predictive performance alone is insufficient for engineering practice. SHapley Additive exPlanations (SHAP), introduced by Lundberg and Lee [LL17], address this by assigning each feature a quantified contribution to individual predictions, enabling both global rankings and instance-level explanations. SHAP has been applied in corrosion contexts [Ya20] and in adjacent environmental domains including water quality assessment and flood susceptibility modelling [Ab26; Sa25].

The present work applies ensemble ML and SHAP-based explainability to corrosion data from steel sheet pile walls in the Port of Hamburg, with the aim of identifying environmental drivers of corrosion and assessing the conditions under which water quality monitoring data can support spatially resolved corrosion prediction.

3 Data and Methodology

The corrosion dataset comprises wall thickness measurements from steel sheet pile walls in the Port of Hamburg. Corrosion loss (in mm) is derived as the difference between the originally installed and the measured residual wall thickness, and related to structure age at the time of measurement (Years Since Construction, t). Figure 1 shows the temporal coverage of available measurements. For the large majority of sites, only a single observation is available; only one site was measured repeatedly. This sparse temporal coverage precludes the reconstruction of individual corrosion trajectories and directly motivates the modelling strategy described below.

Corrosion loss over time is modelled using the power law [Ca18; Me03; Me14]:

$$C(t) = A \cdot t^n, \quad (1)$$

where $C(t)$ is the cumulative corrosion loss in mm, A the initial corrosion rate, and n a dimensionless exponent describing the deceleration of corrosion over time. A global fit over the full dataset yields $C(t) = 0.0800 \cdot t^{0.7135}$, which serves as the reference curve for all subsequent analyses. Prior to fitting, measurement positions within each site are averaged to avoid treating repeated measurements at the same structure as statistically independent observations.

Since the majority of sites provide only a single observation, a station-level trajectory fit is not feasible. Instead, an age-adjusted residual C_{residual} is introduced as the target variable, defined as the deviation of the observed corrosion loss from the age-dependent expectation of the global fit:

$$C_{\text{residual},i} = C_{\text{obs},i} - \hat{C}(t_i), \quad (2)$$

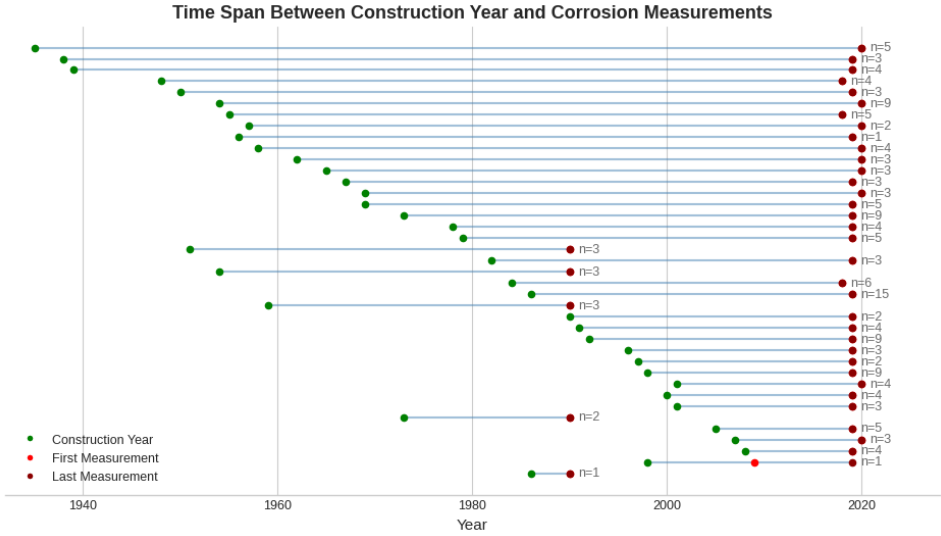


Fig. 1: Temporal coverage of corrosion measurements per site. Green markers indicate the year of construction, orange markers the first measurement, and red markers the last available measurement. The annotation gives the number of sheet piles per site for which both a construction year and at least one measurement are available. The majority of sites have only a single measurement point.

where $C_{obs,i}$ is the observed corrosion loss, t_i the age of site i at the time of measurement, and $\hat{C}(t_i) = A_{fit} \cdot t_i^{n_{fit}}$ the global power law evaluated at age t_i . Positive values indicate corrosion in excess of the global expectation; negative values indicate below-average corrosion. Because the subtracted term $\hat{C}(t_i)$ varies with t_i , $C_{residual}$ is station-individual and is largely decoupled from the global age trend.

Water quality time series from monitoring stations within the harbour basin are used to characterise local chemical and physical exposure conditions. The dataset covers water temperature, dissolved oxygen, oxygen saturation, chloride concentration, dissolved organic carbon (DOC), pH, and electrical conductivity.

Figure 2 shows the temporal coverage and measurement frequency for each parameter.

A central methodological challenge is the spatial assignment between water quality monitoring stations and corrosion measurement sites. As monitoring stations are distributed at a coarser spatial resolution than the corrosion sites, a one-to-one assignment is not feasible. Instead, the assignment is performed via structural groups, each comprising multiple corrosion parents that are jointly assigned to a single water quality station. These groups were defined by a hydraulic engineering expert based on flow behaviour and spatial proximity, and represent hydrologically homogeneous zones. As a consequence, all corrosion sites within the same group receive identical water quality values. The dataset was further restricted to sites for which all seven water quality parameters were actually recorded at the assigned

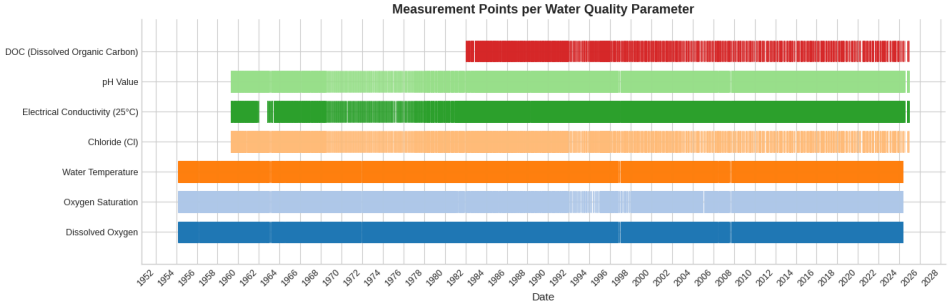


Fig. 2: Temporal coverage and measurement frequency of the seven water quality parameters at the monitoring stations in the Port of Hamburg. Regular measurements are available for most parameters from the 1950s onwards, substantially exceeding the temporal density of the corrosion dataset.

monitoring station, yielding five monitoring stations assigned to 39 corrosion parents and 153 individual measurement sites.

The feature matrix for each corrosion site comprises the temporal mean of each of the seven water quality parameters over the full available measurement period of the assigned monitoring station, the structure age at the last measurement (Years Since Construction), and the assignment group encoded as a one-hot variable. Since most sites provide only a single observation, the structure age corresponds to the age at that observation. Temporal averaging of water quality data represents a first exploratory approximation, assuming that long-term mean values adequately represent the cumulative chemical exposure over the structure's lifetime.

The assignment group is deliberately retained as a feature despite its close relationship with the water quality values. This enables the contribution of the group structure to be made explicit in the model and quantified through ablation experiments comparing the full model against a model excluding the group variable. The target variable throughout is C_{residual} as defined in Equation 2.

A Random Forest regression model is trained on C_{residual} to identify relevant drivers of corrosion behaviour. The model is evaluated using 5-fold cross-validation, supplemented by hold-out validation to assess generalisation. Feature contributions are assessed using permutation importance, computed on the hold-out set to avoid overfitting bias, and complemented by a SHAP-based analysis to examine contributions at the level of individual sites. A binary classification model is additionally trained, with the target variable indicating whether $C_{\text{residual}} > 0$, i.e. whether a site corrodes more than globally expected.

4 Preliminary Results

A Random Forest regression model trained on all features achieves a cross-validated R^2 of 0.719 ± 0.087 on C_{residual} , substantially more stable than a model using water quality features alone ($R^2 = 0.388 \pm 0.083$). The cross-validation estimate is confirmed by a hold-out

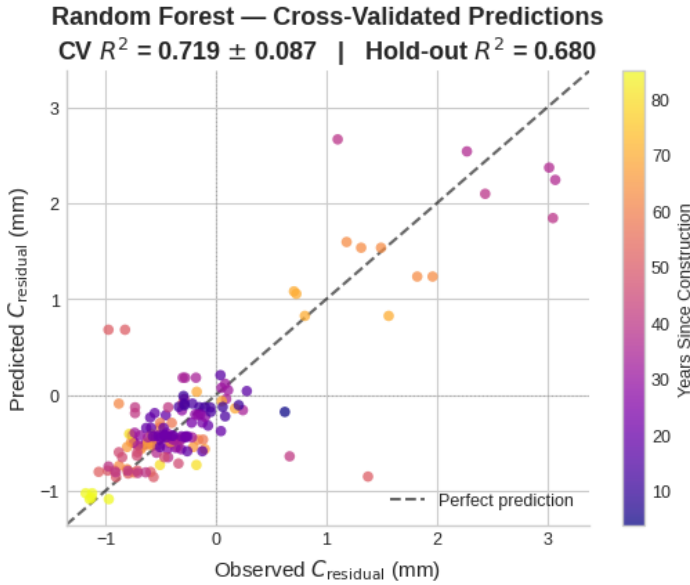


Fig. 3: Cross-validated predictions of the Random Forest regression model against observed C_{residual} values ($CV R^2 = 0.719 \pm 0.087$, hold-out $R^2 = 0.680$). The dashed line indicates perfect agreement. Colour encodes structure age; sites with high positive residuals are predominantly older. Predictions are well-calibrated in the central range but systematically underestimate strongly corroding sites.

R^2 of 0.680. Figure 3 shows the cross-validated predictions against observed C_{residual} values. Predictions are well-calibrated in the central range; sites with strongly positive residuals, indicating corrosion substantially above the global expectation, are systematically underestimated, consistent with their underrepresentation in the training data.

Permutation importance on the hold-out set attributes predictive performance predominantly to three features: structure age (Years Since Construction, 0.475), mean dissolved oxygen (0.204), and a single assignment group (wq_group_3, 0.157), while all remaining water quality features contribute negligibly. The moderate negative association between structure age and C_{residual} suggests that the global fit tends to overestimate corrosion loss at older structures.

A direct comparison between the full model ($R^2 = 0.719 \pm 0.087$) and a model excluding the explicit group variable ($R^2 = 0.718 \pm 0.087$) reveals virtually identical performance. This reflects the hydrological consistency of the spatial assignment: the groups were defined by a hydraulic engineering expert based on flow behaviour and spatial proximity, and represent hydrologically homogeneous zones. Within each zone, water quality values are by construction constant, so the water quality means already implicitly encode the group structure. The practical consequence is that site-specific corrosion differences within a hydrologically homogeneous zone cannot be fully explained by water quality data alone.

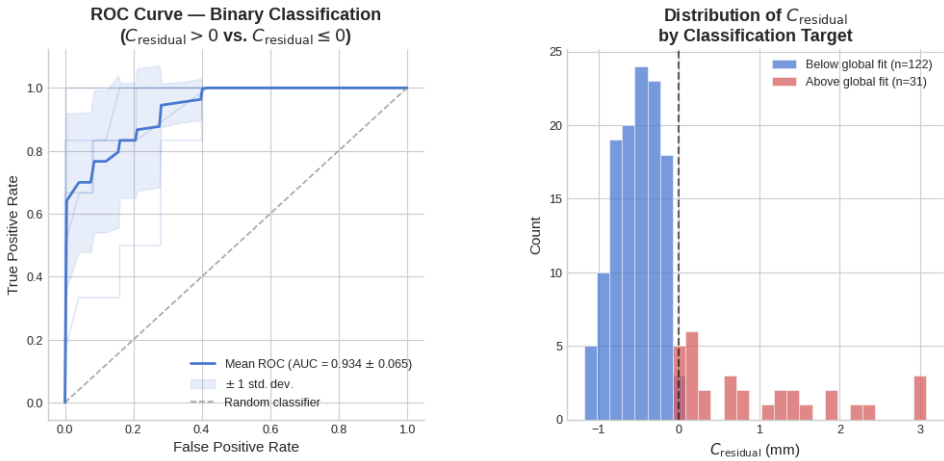


Fig. 4: Left: ROC curve of the binary classification model ($C_{\text{residual}} > 0$ vs. $C_{\text{residual}} \leq 0$) across five cross-validation folds (AUC = 0.934 ± 0.065). Right: distribution of C_{residual} by class ($n = 122$ below vs. $n = 31$ above the global fit).

As a preliminary robustness check, the model was additionally evaluated under leave-one-group-out cross-validation, holding out one of the five monitoring groups per fold. The pooled out-of-group R^2 falls below the global-mean baseline, indicating that the random cross-validated performance partly reflects group identity shared between training and test sites. A larger dataset with more monitoring groups would be required to quantify generalisation to unseen locations reliably.

A binary classification model predicting whether $C_{\text{residual}} > 0$ achieves a cross-validated ROC-AUC of 0.934 ± 0.065 , indicating that sites corroding above the global expectation are reliably distinguished from those below. Figure 4 shows the ROC curve and the distribution of C_{residual} by class. Of the 153 sites, 122 lie below and 31 above the global fit. The two classes show little overlap in the residual space, and the pronounced class imbalance should be considered when interpreting the AUC, as structure age and group membership contribute substantially to discriminative performance.

SHAP values were computed for the Random Forest model to assess feature contributions at the site level. Figure 5 shows the beeswarm plot excluding group variables. Structure age dominates with a wide, non-monotone SHAP spread (-0.75 to $+1.10$). Among water quality parameters, only mean dissolved oxygen (DO) shows a noteworthy contribution (-0.25 to $+0.60$), with low DO values tending to be associated with elevated C_{residual} , consistent with the role of oxygen availability in aquatic corrosion processes. All remaining parameters contribute negligibly.

Closer inspection confirms that structure age acts as the primary differentiating feature within hydrologically homogeneous monitoring regions. Sites sharing the same parent and age receive identical SHAP values, indicating that age captures location identity rather

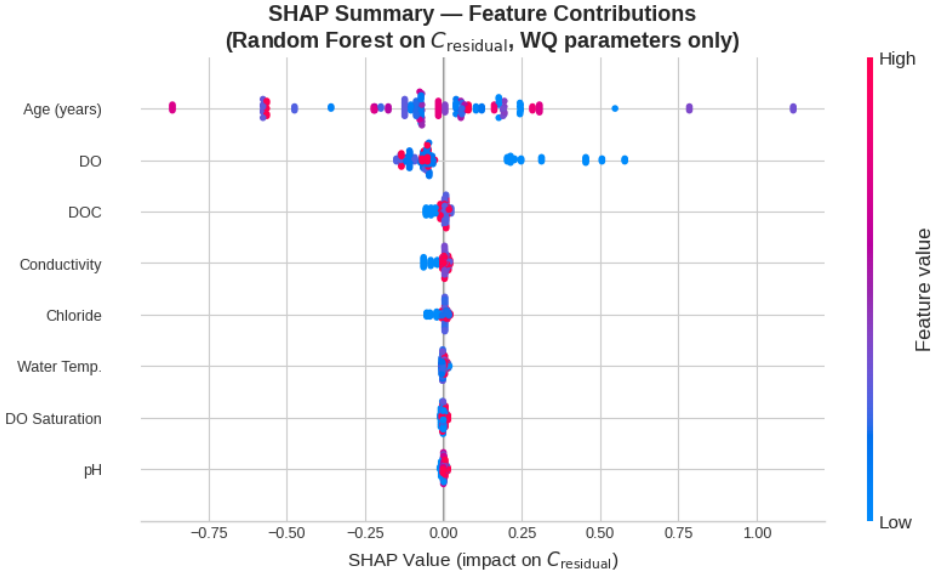


Fig. 5: SHAP beeswarm plot for the Random Forest model on C_{residual} (group variables excluded). Each point represents one site; colour encodes the feature value (red: high, blue: low). Structure age dominates with a wide, non-monotone SHAP spread. Among water quality parameters, only mean dissolved oxygen (DO) shows notable variation; low DO values are associated with elevated C_{residual} . All other parameters contribute negligibly.

than a genuine ageing effect. Within the same monitoring region, sites of different ages exhibit opposing SHAP signs, confirming that structure age captures location-specific differences rather than a direct ageing effect, as it remains the only feature able to account for within-group variation in C_{residual} where water quality is constant.

5 Discussion

The Random Forest model achieves a cross-validated R^2 of 0.719 on the age-adjusted target variable C_{residual} , confirmed by a hold-out R^2 of 0.680, indicating stable generalisation. Since test sites were selected randomly during cross-validation, the reported R^2 likely overestimates generalisation to entirely unseen locations. A preliminary leave-one-group-out check supports this, though a larger dataset is needed for a robust assessment.

The near-identical performance of the full model ($R^2 = 0.719$) and a model excluding the explicit group variable ($R^2 = 0.718$) reflects the hydrological consistency of the spatial assignment: water quality means implicitly encode group structure, so site-specific differences within a group cannot be explained by water quality data alone.

The SHAP analysis confirms this. Structure age dominates predictions not through a genuine

ageing effect but as the only feature differentiating sites within a group. Among water quality parameters, only mean dissolved oxygen shows a noteworthy association with C_{residual} , with lower values tending to correspond to higher corrosion residuals, consistent with the role of oxygen in aquatic corrosion. Whether this reflects a genuine chemical signal cannot be resolved from the available data. The spatial resolution of the monitoring network is too coarse to disentangle chemical signal from location effects, and 99.3 % of sites provide only a single measurement, precluding individual trajectory reconstruction.

section Conclusion and Outlook Water quality data reflect corrosion behaviour at the level of hydrologically homogeneous monitoring groups, providing a consistent basis for group-level assessment of groups represented in the training data. A preliminary leave-one-group-out check suggests that prediction at unobserved locations is substantially harder. For site-specific prediction within groups, water quality data alone are insufficient, as finer-grained factors such as local flow conditions or structural properties govern the remaining variability. Future work should identify spatially finer-grained predictors, extend the feature set towards richer temporal statistics, and expand the corrosion dataset to multiple measurements per site. A variance-partitioning analysis over structure age, group membership, and water quality would further quantify how much predictive signal originates from the spatial group structure rather than the water quality parameters. The long-term objective is reliable corrosion risk assessment at sites with sparse inspection records, supporting more efficient maintenance planning along federal waterways.

References

- [Ab26] Ab Karim, M. A.; Wan Ismail, W. Z.; Mohd Shuib, F. M.; Ab Aziz, N. A.; Ghazali, A. K.: Water Quality Monitoring and Assessment Using Machine Learning: A Review of Formulation, Modeling Approaches, and Explainable Artificial Intelligence. *Environments* 13 (5), 2026, DOI: 10.3390/environments13050267.
- [BHZ22] Ben Seghier, M. E. A.; Höche, D.; Zheludkevich, M.: Prediction of the Internal Corrosion Rate for Oil and Gas Pipeline: Implementation of Ensemble Learning Techniques. *Journal of Natural Gas Science and Engineering* 99, 2022, DOI: 10.1016/j.jngse.2022.104425.
- [Bu22] Bundesanstalt für Wasserbau: BAW-Empfehlung: Spundwanddickenmessungen in Häfen und an Wasserstraßen: Grundlagen, Planung, Durchführung, Auswertung und Interpretation (ESM), BAW-Merkblatt, -Empfehlung und -Richtlinie, Karlsruhe: Bundesanstalt für Wasserbau (BAW), 2022, URL: https://izw.baw.de/publikationen/empfehlungen/0/BAWEmpfehlung_Spundwanddickenmessung_ESM_2022.pdf.
- [Ca18] Cai, Y.; Zhao, Y.; Ma, X.; Zhou, K.; Chen, Y.: Influence of Environmental Factors on Atmospheric Corrosion in Dynamic Environment. *Corrosion Science* 137, pp. 163–175, 2018, DOI: 10.1016/j.corsci.2018.03.042.
- [Ch17] Chico, B.; de la Fuente, D.; Díaz, I.; Simancas, J.; Morcillo, M.: Annual Atmospheric Corrosion of Carbon Steel Worldwide: An Integration of ISOCORRAG, ICP/UNECE and MICAT Databases. *Materials* 10 (6), 2017, DOI: 10.3390/ma10060601.
- [Co22] Coelho, L. B.; Zhang, D.; Van Ingelgem, Y.; Steckelmacher, D.; Nowé, A.; Terryn, H.: Reviewing Machine Learning of Corrosion Prediction in a Data-Oriented Perspective. *npj Materials Degradation* 6 (1), p. 8, 2022, DOI: 10.1038/s41529-022-00218-4.

- [DI10] DIN: DIN EN 1993-5:2010-12: Eurocode 3: Bemessung und Konstruktion von Stahlbauten – Teil 5: Pfähle und Spundwände; Deutsche Fassung EN 1993-5:2007 + AC:2009, 2010, doi: <https://dx.doi.org/10.31030/1723936>.
- [Fr21] Fritsch, U. et al.: Klimawirkungs- und Risikoanalyse für Deutschland 2021: Teilbericht 3: Klimarisiken im Cluster Wasser, Climate Change 22/2021, Umweltbundesamt, 2021.
- [HAH07] Houyoux, C.; Alberts, D.; Heeling, A.: Design Method for Steel Structures in Marine Environment Including the Corrosion Behaviour, Directorate-General for Research and Innovation, European Commission, 2007, URL: <https://op.europa.eu/publication-detail/-/publication/d742313f-f26f-4171-8f05-e5f78f2e4c6f>.
- [IJM24] Imran, M. M. H.; Jamaludin, S.; Mohamad Ayob, A. F.: A Critical Review of Machine Learning Algorithms in Maritime, Offshore, and Oil & Gas Corrosion Research: A Comprehensive Analysis of ANN and RF Models. Ocean Engineering 295, p. 116796, 2024, doi: 10.1016/j.oceaneng.2024.116796.
- [LL17] Lundberg, S. M.; Lee, S.-I.: A Unified Approach to Interpreting Model Predictions. In: Advances in Neural Information Processing Systems. Vol. 30, pp. 4765–4774, 2017, doi: <https://doi.org/10.48550/arXiv.1705.07874>.
- [Me03] Melchers, R. E.: Mathematical modelling of the diffusion controlled phase in marine immersion corrosion of mild steel. Corrosion Science 45 (5), pp. 923–940, 2003, doi: [https://doi.org/10.1016/S0010-938X\(02\)00208-1](https://doi.org/10.1016/S0010-938X(02)00208-1).
- [Me06] Melchers, R. E.: Modelling Immersion Corrosion of Structural Steels in Natural Fresh and Brackish Waters. Corrosion Science 48 (12), pp. 4174–4201, 2006, doi: 10.1016/j.corsci.2006.04.012.
- [Me14] Melchers, R. E.: Modelling Long Term Corrosion of Steel Infrastructure in Natural Marine Environments. In (Melchers, R. E.; Liengen, T.; Féron, D.; Basséguy, R.; Beech, I. B., eds.): Understanding Biocorrosion. Vol. 66, Woodhead Publishing Series in Materials, Woodhead Publishing, Amsterdam, pp. 213–241, 2014.
- [Me25] Melchers, R. E.; Jeffrey, R.; Chaves, I. A.; Petersen, R. B.: Predicting Corrosion for Life Estimation of Ocean and Coastal Steel Infrastructure. Materials and Corrosion 76 (6), pp. 776–789, 2025, doi: 10.1002/maco.202314201.
- [PI22] PIANC: Managing Climate Change Uncertainties in Selecting, Designing and Evaluating Options for Resilient Navigation Infrastructure: Technical Note No. 1, Permanent Task Group for Climate Change, 2022, URL: <https://pianc.org.au/wp-content/uploads/2022/04/PTGCC-Technical-Note-1-Managing-Climate-Change-Uncertainties-April-2022.pdf>, visited on: 02/24/2026.
- [Sa25] Salih, M. et al.: A Perspective on Explainable Artificial Intelligence Methods: SHAP and LIME. Advanced Intelligent Systems 7 (1), p. 2400304, 2025, doi: 10.1002/aisy.202400304.
- [Ya20] Yan, L.; Diao, Y.; Lang, Z.; Gao, K.: Corrosion Rate Prediction and Influencing Factors Evaluation of Low-Alloy Steels in Marine Atmosphere Using Machine Learning Approach. Science and Technology of Advanced Materials 21 (1), pp. 359–370, 2020, doi: 10.1080/14686996.2020.1746196.
- [Ya24] Yang, J.; Zou, D.; Zhang, M.; Que, Z.; Liu, T.; Zhou, A.; Li, Y.: Marine Steel Corrosion Prediction and Zonation Using Feature Extraction and Machine Learning in the Seas Around China. Ocean Engineering 314, p. 119649, 2024, doi: 10.1016/j.oceaneng.2024.119649.

Environmental Management Software

Air Quality Knowledge as a Driver of Computing Competences: An International Project-Based Learning Project

Lukasz Szalata ¹, Grit Behrens ², Cezary Orłowski ³, Rafał Grzeszczyk ⁴

Abstract: This paper presents key recommendations developed based on the analysis of an international didactic project conducted jointly by two universities. The initiative aimed to enhance students' competencies in data processing systems, software engineering, and the applications of computer science in environmental engineering. Utilizing real-world air quality data and modern technologies (Node-RED, RabbitMQ, Docker), the project provided students with hands-on experience in near-production environments. An analysis of the course outcomes and performance led to the formulation of targeted recommendations across three core pillars: Software Engineering: Emphasizes deeper integration of the "learning by doing" approach alongside formal engineering practices (documentation, automated testing, code quality control) to effectively balance rapid prototyping with professional industry standards. Environmental Application Design: Highlights the necessity of embedding the environmental context as the foundation of the design process, recommending the development of user-engaging features (gamification, mobile sensor data) to foster citizen science. Inter-University Cooperation: Advocates for a hybrid teaching model to maximize interaction, alongside the standardization of evaluation criteria and collaborative tools (e.g., Git) to better prepare students for a global professional environment. These guidelines offer a strategic framework for optimizing future educational initiatives that bridge theory, practice, and international academic synergy.

Keywords: Air quality, computing competence, system development, Software Engineering, Environmental Application Design.

¹ Faculty of Environmental Engineering - Wrocław University of Science and Technology, PL. Grunwaldzki 9, 50-377 Wrocław, Poland, Lukasz.Szalata@pwr.edu.pl, <https://orcid.org/0000-0001-8307-6331>

² Applied Computer Science, Bielefeld University of Applied Sciences, Minden Campus, Artilleriestraße 9, 32427 Minden, Grit.Behrens@hsbi.de, <https://orcid.org/0009-0009-0247-8204>

³ IBM Centre for Advanced Studies, Faculty of Computer Science and New Technologies, WSB Merito University, Aleja Grunwaldzka 238 A, 80-266 Gdansk, Poland, Cezary.Orlowski@gdansk.merito.pl, <https://orcid.org/0000-0001-7716-5522>

⁴ IBM Centre for Advanced Studies, Faculty of Computer Science and New Technologies, WSB Merito University, Aleja Grunwaldzka 238 A, 80-266 Gdansk, Poland, Rafal.Grzeszczyk@gdansk.merito.pl, <https://orcid.org/0000-0001-7716-5522>

1 Introduction

In the era of progressive climate change and its compounding negative impacts, life in urban agglomerations is becoming increasingly challenging. Urban populations, alongside local flora and fauna, are highly vulnerable to the amplified frequency of extreme weather events and severe thermal stress. Anthropogenic climate models predict that the global average maximum temperature in European, South American, and African cities will escalate by 2 to 8°C over the coming decades [Cl26]. This systemic temperature increase, combined with shifting precipitation regimes, triggers a dangerous synergy with localized air pollution, drastically reducing the environmental carrying capacity of modern cities [Wi16].

The World Health Organization (WHO) classifies ambient air pollution as one of the most critical environmental hazards to human health. Prolonged exposure to elevated concentrations of atmospheric pollutants induces severe systemic health failures, predominantly manifesting as chronic respiratory diseases, acute asthma exacerbations, and ischemic strokes. Highlighting the scale of this crisis, the European Environment Agency (EEA) attributes substantial premature mortality rates directly to poor air quality, citing annual losses of 253,000 lives due to fine particulate matter (["PM"] _2.5), 52,000 due to nitrogen dioxide (["NO"] _2), and 22,000 due to ground-level ozone (["O"] _3) (klimaslo, 2025).

This crisis is heavily compounded by rapid global demographic shifts. Urbanization acts as a major driver of local microclimatic modification, often intensifying the Urban Heat Island (UHI) effect. Szałata et al. (2024) identify expanding urban structures as complex, concentrated sources of diverse pollutants, ranging from vehicular exhaust fumes and industrial emissions to heavy construction dust.

Addressing these interconnected challenges requires a fundamental realignment of municipal policies toward sustainable development goals [Gi22]. Derlukiewicz et al. (2023) argue that the most effective response to these urban vulnerabilities is the systemic implementation of low-carbon solutions and the structural transition toward "low-carbon city" frameworks. Crucially, as Fu et al. (2010) emphasize[FLW10], low-carbon urban planning is a multi-objective optimization process. It does not merely seek the reduction of greenhouse gas emissions, but simultaneously aims to secure decoupling-based economic growth and elevate the socioeconomic quality of life for citizens. According to contemporary frameworks, low-carbon initiatives represent an intersection of advanced technologies, mitigation strategies, and green practices designed to capture or reduce atmospheric carbon dioxide (["CO"] _2) while fostering long-term urban resilience (klima.praha, 2025). Consequently, modern municipalities are increasingly forced to re-engineer their infrastructure—ranging from energy grids to real-time environmental monitoring networks—to mitigate emissions and safeguard public health.

Armed with knowledge of air quality and its impact on health, the authors of this publication attempted to develop an application solution aimed at implementing an app

that would enable navigation around the Campus and its surrounding areas in Gdańsk and Minden, in areas with the best air quality. The work aimed to monitor the most optimal cycling routes and implement them into the operating system.

The authors of the publication, together with students, identified the most frequently used access roads to the Campus to verify them in terms of air quality.

The jointly developed application for air quality assessment, designed to integrate with publicly available applications, determines optimal access routes based on the environmental parameter of air quality. According to the authors, who represent an interdisciplinary perspective, this is a novelty that can be confidently implemented into publicly available solutions in the future, constituting a new area of research activity.

This publication presents key project assumptions, highlighting the potential of AI tools for environmental protection and building useful applications. It also outlines the research methodology and expands the chapter on Integration of Web Engineering and Environmental Education. Importantly, it presents recommendations for implementing the developed solutions and key conclusions drawn from the international project.

2 Student education model

Developing internationally coordinated, interdisciplinary student projects bridging hardware, server infrastructure, and artificial intelligence (AI) requires a multifaceted pedagogical approach. Establishing this framework relies on literature covering modern Internet of Things (IoT) pedagogy, distributed project management, AI-assisted coding, and machine learning (ML) for environmental data analysis [GA23], [Ka23].

IoT pedagogy has shifted from lectures to active, Project-Based Learning (PBL) and STEM Maker education. Integrating iterative engineering design—proposing questions, planning, implementing, and evaluating—bridges theoretical knowledge with physical outputs [Ch20]. Building tangible environmental monitors connects students directly with data collection, making abstract computational concepts accessible and boosting engagement [SCS25].

Managing complex, multi-university projects requires robust technological and team management frameworks. Distributed IoT systems provide real-time insights into technical bottlenecks and student performance, allowing managers to adapt teaching methodologies and align distributed teams [GA23].

Simultaneously, Large Language Models (LLMs) and AI coding assistants (e.g., GitHub Copilot, ChatGPT) have transformed software engineering. In education, AI tools enhance programming efficiency, though students often struggle with over-reliance and evaluating code correctness [Ni24]. In professional settings, AI productivity relies on suggestion quality rather than time-to-completion, using telemetry models to suppress low-quality

suggestions and minimize cognitive fatigue [Mo24]. However, commercial deployment still requires human oversight to meet production standards like accessibility [Mo25].

For environmental applications, low-cost IoT sensors face hardware noise and calibration drift. Server-side ML algorithms (e.g., Random Forest, XGBoost, Support Vector Regression) correct these errors, enabling real-time Air Quality Index (AQI) predictions and making student-built sensors viable for scalable monitoring [Ka23].

While existing literature addresses IoT pedagogy, project management, AI coding, and ML analysis separately, a research gap remains in synthesizing them into a cohesive higher-education curriculum. Combining tangible IoT development with advanced AI analysis provides a unified framework to teach web engineering, modern AI coding, and sustainable environmental applications within a globally coordinated initiative [Zi22].

To address this gap and evaluate the effectiveness of the proposed international didactic framework, this study formulates three primary research questions:

- RQ1 (Software Engineering & Pedagogy): How does integrating real-world environmental streaming data (e.g., Airly API, mobile sensors) into an international Project-Based Learning (PBL) framework affect students' acquisition of software engineering competencies compared to traditional simulated coding tasks?
- RQ2 (Didactics & AI/IoT Technologies): In what ways can flow-based data modeling tools (Node-RED) combined with message brokers (RabbitMQ) and AI-assisted workflows balance rapid prototyping with the enforcement of production-grade engineering standards in distributed student teams?
- RQ3 (Environmental Engagement & Citizen Science): To what extent does embedding domain-specific air quality knowledge (e.g., AQI, PM_{2.5}/PM₁₀, route exposure) and gamification mechanisms into web application design enhance both user engagement and student environmental literacy?

3 Research methodology

In the winter semester of the 2025/2026 academic year, Bielefeld University of Applied Sciences and Arts and WSB Merito University in Gdańsk executed a joint educational project. Supervised by Prof. Dr. Ing. Grit Behrens and Prof. dr hab. inż. Cezary Orłowski, the initiative aimed to build practical competencies in distributed data processing systems using real-world air quality data.

Environmental measurements—including NO, CO, NO₂, O₃, SO₂, PM_{2.5}, PM₁₀, and NH₃—were sourced via the Airly API (**Fig. 1**). Working with authentic data provided

students with a production-like operating environment.

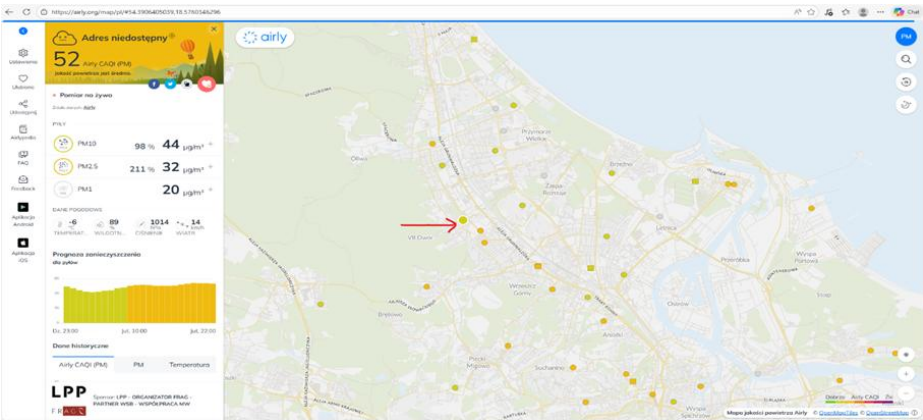


Fig. 1 Airly API: Our Data Source

Twelve students from WSB Merito University in Gdańsk implemented nine applications, each dedicated to a specific local air quality monitoring station. The course combined theoretical foundations delivered via Moodle with hands-on, experiential learning. Following an introduction to project objectives and distributed systems concepts, students entered the implementation phase using Node-RED for flow-based data pipeline modeling and RabbitMQ for asynchronous message queuing and routing.

The detailed data processing pipeline is presented in Figure 2 (Fig. 2 – Data Flow Steps), which illustrates the successive stages of data handling—from data acquisition via the API, through preprocessing and classification, to storage in structures enabling further analysis. This approach allowed students to gain a comprehensive understanding not only of data processing itself but also of asynchronous communication mechanisms and the scalability characteristics of distributed systems.

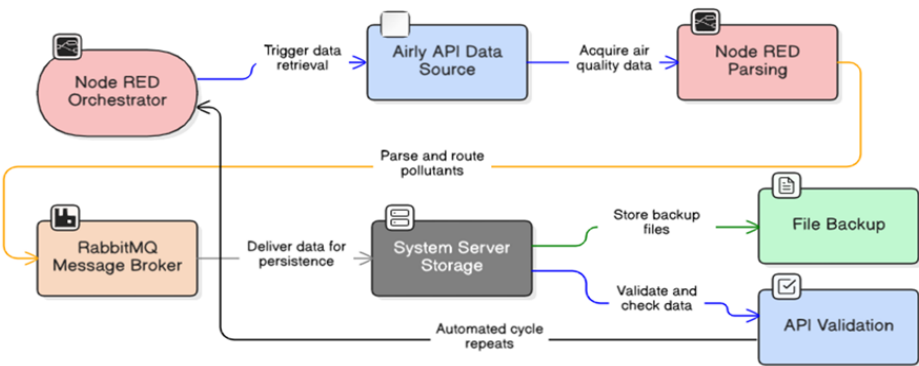


Fig. 2 Data Flow Steps

A key project aspect was preparing data for artificial intelligence systems, specifically Retrieval-Augmented Generation (RAG) paradigms. Processed RabbitMQ data were converted into structured textual documents to create knowledge bases for Large Language Models (LLMs). Students managed the complete solution lifecycle, including Docker containerization (docker-compose), Node-RED pipeline development, system monitoring, and technical presentations covering architecture and data flows.

The outcomes were presented in the form of technical presentations covering system architecture, component configuration, and data flow analysis. Monitoring mechanisms and potential use cases of the collected data are depicted in Figure 3 (Fig. 3 – Monitoring and Using the Data Collected), which highlights both technical and application-oriented perspectives.



Fig. 3 – Monitoring and Using the Data Collected

The project yielded nine functional applications and several key pedagogical insights:

- **Tooling Synergy:** Flow-based tools (Node-RED) combined with message brokers (RabbitMQ) simplified distributed system design and supported rapid prototyping ("vibe coding"), particularly for automated container configuration (e.g., docker-compose).
- **Online Limitations:** Fully remote instruction restricted communication and teamwork development, indicating that hybrid or in-person formats are better suited for collaborative skills.
- **Engineering Standards:** While individual projects accelerated implementation, they lacked formal software lifecycle rigor. Future iterations should incorporate testing, version control, and documentation

In conclusion, the project represents a valuable example of integrating academic education with practical applications of modern data processing technologies and artificial intelligence. The incorporation of real-world data sources, industry-relevant tools, and international collaboration significantly enhanced the quality of the learning process and contributed to better preparing students for the challenges of the contemporary labor market.

4 Integration of Web Engineering and Environmental Education

During the 2025/2026 winter semester at Hochschule Bielefeld (Campus Minden), the Web Engineering course applied a Project-Based Learning (PBL) approach connecting software engineering with healthy air initiatives. Students developed web-based components for the SmartMonitoring Airquality system, addressing urban air quality data collection, visualization, and analysis.

The methodology combined theoretical instruction, agile workflows, and functional prototyping. Following an international exchange lecture by Dr. Łukasz Szalata (Wrocław University of Science and Technology) on air quality metrics (PM_{2.5}, PM₁₀, NO_x, O₃, temperature), students translated environmental domain knowledge into technical requirements, database models, APIs, and web applications.

The course utilized real-world stationary data from Gdansk (Airly API) alongside stationary and mobile bicycle-based sensor data from Minden. Incorporating cycling added a citizen-science dimension, allowing students to analyze spatiotemporal pollution patterns, compare routes, and identify high exposure areas. Development was divided into specialized strands, including a Researcher Interface featuring map visualizations, route management, sensor overviews, statistical summaries, and configurable plots for decision-makers.

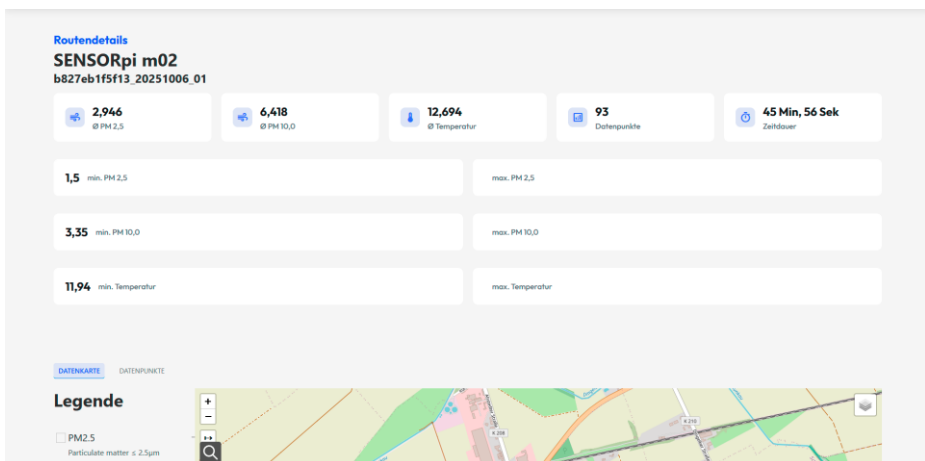


Fig. 4 - Web page for route details of mobile bicycle sensor with values for mean PM2.5 of 2.9, mean PM10 of 6.418, mean temperature, number of measurements along the route, and duration of the route.

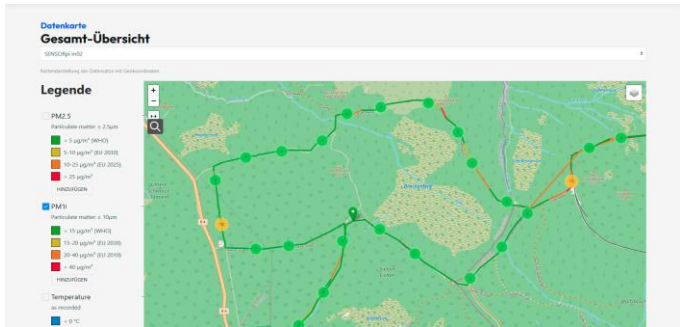


Fig.5 - Visualization of a bicycle route in a green area around mountain Brocken mountain in Germany in students' web application

Another group developed a Progressive Web Application (PWA) with push notifications and gamification mechanisms—including achievements, streaks, and leaderboards—to incentivize regular user participation in environmental data collection.



Fig. 6. Achievements designed by the student team for motivation for air-quality measurements and used in the gamified web application prototype.

From a software engineering perspective, students applied full-lifecycle practices—including requirements analysis, architecture and database modeling, REST APIs, Git-based workflows, issue tracking, and iterative development. The technological stack comprised the SmartData API, Smart Web Application Components, PostgreSQL/PostGIS, Payara, Swagger/OpenAPI, and PWA capabilities (Service Workers, Push API).

This approach integrated three core learning dimensions:

- **Technical Mastery:** Hands-on competencies in web engineering and distributed systems.
- **Domain Context:** Deeper understanding of urban air quality issues.
- **Environmental Literacy:** Leveraging informatics to transform raw data into visible, actionable insights for sustainability and healthier mobility choices.

5 Recommendations

Based on the evaluation of this joint international initiative, three key sets of recommendations are proposed to optimize future educational programs integrating software engineering, distributed systems, and environmental analytics:

1. **Software Engineering Didactics**
 - a) **Balance Prototyping with Rigor:** While tools like Node-RED, RabbitMQ, and Docker accelerate architectural understanding through "learning by doing", future iterations must strictly integrate full software lifecycle practices—specifically automated testing, systematic documentation, version control, and code quality assurance.
 - b) **Near-Production Environments:** Practical projects should continuously rely on authentic data streams to bridge the gap between academic theory and industrial software deployment.
2. **Application Design in Environmental Engineering**
 - a) **Domain-Centric Design:** Environmental context must serve as the structural foundation of the system (dictating requirements, data models, and UI) rather than a superficial dataset.
 - b) **Citizen Science & Engagement:** Integrating gamification, push notifications, and mobile sensor data (e.g., bicycle-based measurements) enhances spatiotemporal analysis, encourages user participation, and drives interdisciplinary learning.
3. **Multi-University Collaborative Frameworks**
 - a) **Hybrid Instruction Model:** Purely remote delivery hinders teamwork and communication. A hybrid approach—combining online collaboration with targeted in-person sessions—is essential for building soft competencies.
 - b) **Standardized Infrastructure:** Joint courses require unified work environments (e.g., shared Git platforms, project management tools), clear role allocation, and aligned evaluation standards to facilitate effective cross-institutional knowledge transfer.

6 Discussion and conclusions

The outcomes of this joint international project provide direct insights into the three formulated research questions. Regarding RQ1, the exposure to authentic, noisy environmental data streams significantly elevated students' engagement and forced them to address real-world system edge cases, which simulated coding environments fail to replicate. Addressing RQ2, while the combination of Node-RED, RabbitMQ, and Docker facilitated rapid architectural prototyping ("vibe coding"), it highlighted the vital necessity of enforcing strict software lifecycle practices—such as automated testing and continuous documentation—to maintain production quality. Finally, concerning RQ3, placing air quality domain knowledge at the core of the application design (e.g., eco-routing, gamification, and mobile sensor tracking) successfully fostered citizen science principles and elevated students' broader environmental literacy.

The authors clearly emphasize that international cooperation forms the cornerstone of effective research and educational initiatives. By combining the expertise of academics and students across universities around a shared objective—the development and implementation of the "Air" project—this initiative established an innovative international framework for data processing and forecasting systems.

The core strength of the project lies in placing the environmental context at the very center of the engineering lifecycle: air quality parameters directly shaped the requirements analysis, database models, and user interfaces. This domain-driven approach allowed students to develop a highly functional tool for data processing and air-quality-aware route optimization, while substantially strengthening their practical IT competencies.

To ensure the success and sustainability of future cross-border teaching initiatives, the authors highlight several structural insights:

- **Incentive Frameworks for Online Collaboration:** Managing remote, multi-institutional teams requires significant effort. To boost engagement, academic curricula and examination regulations should be adapted to award extra ECTS credits to participating students and credit additional teaching hours to instructors.
- **Coherent Project Structure:** Joint university projects demand clearly defined goals, clear role distribution, and intensive student-instructor interaction to overcome the inherent challenges of remote delivery.
- **Multi-Format Educational Model:** A multi-format teaching approach that systematically reinforces formal software engineering practices across consecutive semesters provides high educational value, preparing students for real-world labor market challenges.

Ultimately, this project demonstrates how merging computer science education with environmental awareness creates a valuable, replicable model for international academic collaboration.

7 Bibliography

- [Ch20] Chen, R., Zheng, Y., Xu, X., Zhao, H., Ren, J., & Tan, H.-Z. STEM Teaching for the Internet of Things Maker Course: A Teaching Model Based on the Iterative Loop. *Sustainability*, 12(14), 5758, 2020
- [FLW10] Fu, Y., Liu, Y. J., & Wang, Y. L., Evaluation method and supporting system of low carbon cities. *China Population, Resources and Environment*, 20(8), 44–47, 2010
- [GA23] Ghashim, I. A., & Arshad, M. Internet of Things (IoT)-Based Teaching and Learning: Modern Trends and Open Challenges. *Sustainability*, 15(21), 15656, 2023
- [Gi22] Giuliano, A., Errico, M., Salehi, H., & Avino, P., Environmental impact assessment by green processes. *International Journal of Environmental Research and Public Health*, 19(23), 15575, 2022
- [Cl26] Climate Action Network International, https://climatenetwork.org/ipcc_6ar/, 19.06.2026
- [Ka23] Karnati, H. IoT-Based Air Quality Monitoring System with Machine Learning for Accurate and Real-time Data Analysis. *arXiv*, 2023
- [Mo25] Mowar, P., Peng, Y.-H., Wu, J., Steinfeld, A., & Bigham, J. P., CodeA11y: Making AI Coding Assistants Useful for Accessible Web Development. *arXiv. Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 2025
- [Mo24] Mozannar, H., Bansal, G., Fourney, A., & Horvitz, E., When to Show a Suggestion? Integrating Human Feedback in AI-Assisted Programming. *Proceedings of the AAAI Conference on Artificial Intelligence*, 38(9), 10137–10144, 2024
- [Ni24] Nizamudeen, F., Gatti, L., Bouali, N., & Ahmed, F. Investigating the Impact of Code Generation Tools (ChatGPT & GitHub Copilot) on Programming Education. *Proceedings of the 16th International Conference on Computer Supported Education*, 221–229, 2024
- [SÇS25] Spaho, E., Çiço, B., & Shabani, I., IoT Integration Approaches into Personalized Online Learning: Systematic Review. *Computers*, 14(2), 63, 2025
- [Sz24] Szałata, Ł., Derlukiewicz, N., Mańkowska, D., & Zwoździak, J., Urban challenges and air quality in selected European cities. *Environment Protection Engineering*, 50(3), 61–75, 2024
- [Wi16] Wis-Bielewicz, J., Serre, C., Śniegocki, A., Jackl, F., Kassenberg, A., & Szymalski, W., The benefits of a low-carbon economy in cities. *Warszawa/Berlin: adelphi/WiseEuropa/Institute for Sustainable Development*, 2016
- [Ze26] ZELENÍ Praha, <https://praha.zeleni.cz/praha-ma-plan-jak-na-klima.>, 19.06.2026
- [Zi22] Ziegler, A., Kalliamvakou, E., Li, X. A., Rice, A., Rifkin, D., Simister, S., Sittampalam, G., & Aftandilian, E., Productivity assessment of neural code completion. *Proceedings of the 6th ACM SIGPLAN International Symposium on Machine Programming*, 2022

Green Coding

Windows Integration into the Green Metrics Tool: Energy Measurement of Native Windows Applications via RAPL

Maximilian Jahns¹, Geerd-Dietger Hoffmann², Dennis Junger¹, and Volker Wohlgemuth¹

Abstract: The Green Metrics Tool (GMT) profiles energy and performance for Linux and macOS, but previously had no Windows support. This work integrates Intel RAPL measurement into GMT for native Windows applications, comparing two hosting designs that leave GMT's core components unmodified: a WSL2-bridged prototype in which the GMT runner executes inside WSL2, and a native Windows design orchestrating measurement containers via Docker Desktop. Both are validated with ten-run reproducibility studies of a Blender 5.1.1 (*Lone Monk*, Cycles CPU, 1440 × 1080) benchmark. Total-energy coefficients of variation stay at or below 2.7 % in every domain: the native design yields 1.272 ± 0.028 Wh for `cpu_package` (CV = 2.2 %) and the WSL2-bridged design 1.166 ± 0.019 Wh (CV = 1.6 %); a second acquisition path (EMI) reproduces the native figure to within 0.1 %. The identical Windows workload differs by 9.1 % in energy and 10.9 % in render duration between the two hosting designs – reported as observed in the tested configuration and not separable from a time-of-day confound. The MSR-reading driver layer is adopted from the Scaphandre project; the GMT integration and reproducibility evaluation are original to this work, which supports UN Sustainable Development Goals 9 and 12 by making the energy footprint of Windows software measurable within a certification-anchored framework.

Keywords: RAPL, Windows, Green Metrics Tool, energy measurement, sustainable software, reproducibility

1 Introduction

Environmental management information systems (EMIS) help organizations systematically record, evaluate, and translate the environmental impact of their processes into decisions [Wo05], increasingly including the energy consumption of software and IT infrastructure [Na11]. Crucially, an EMIS is not only a lens onto other software – it is itself software, operated and maintained, and should therefore be subject to the same evaluation standards it demands of other digital systems. A heterogeneous tool ecosystem has established itself for software-related energy measurement, but this diversity makes comparability across systems and platforms harder: tools differ in resolution, instrumentation level, and methodological validity, and many deliver only retrospective figures without robust comparisons [OBR20; RT24].

This is precisely where the Green Metrics Tool (GMT) comes in. Rather than providing yet another isolated measurement tool, GMT anchors its methodology in a recognized

¹ HTW Berlin – University of Applied Sciences, Environmental Informatics Unit, Wilhelminenhofstraße 75A, 12459 Berlin, Germany,

² Green Coding Solutions GmbH, Berlin, Germany,

certification framework – the Blue Angel eco-label for resource- and energy-conscious software products. Through defined product groups, audited award criteria, and standardized test scenarios, it creates a basis for comparison that individual measurement tools cannot offer on their own: what matters is not an isolated benchmark figure, but the reproducibility and transferability of measurements across software, versions, and operating environments [NGK20].

That robustness, however, has so far been confined to Linux and macOS, while Windows-based workstations and engineering applications retain high practical relevance in corporate IT landscapes. The obstacle is technical: although Intel Running Average Power Limit (RAPL) is available at the processor level independent of the operating system, Windows does not expose the underlying Model Specific Registers (MSRs) to applications, and access requires a signed kernel driver [Sc19]. Existing tools such as Scaphandre [Sc24] enable RAPL access under Windows in principle, but offer no interface to GMT’s metric-provider architecture. This work closes that gap and makes the following contributions:

- Two architectures for RAPL energy measurement on Windows within the GMT framework – which had no such capability before – are presented and compared: a WSL2-bridged design and a native Windows design that orchestrates measurement containers via Docker Desktop, neither modifying GMT’s core components. The MSR-reading kernel driver and the core acquisition loop are adopted from the open-source Scaphandre project [Sc24]; the contribution of this work is the integration of that access path into GMT.
- A GMT metric provider built around this driver, extended with environment-variable-based configuration and integrated driver-health monitoring.
- A reproducibility evaluation ($n = 10$) of both designs, cross-checked against a second acquisition path (EMI).

2 Background and Related Work

Technical measurement approaches build on Intel RAPL interfaces, ACPI, hardware counters, kernel instrumentation, or container-related telemetry [RT24]. Several tools address software-related energy measurement without being tied to a certification methodology such as the Blue Angel. PowerAPI [Fi24] provides a middleware layer for process-level RAPL measurement via the MSR kernel module, with limited Windows support. Scaphandre [Sc24] is an agent-based tool with RAPL support on Linux and Windows, using a kernel driver (ScaphandreDrv) for MSR access and exporting to Prometheus, but without GMT integration. EnergiBridge [SCD24] wraps cross-platform measurement around arbitrary commands, including a Windows backend, but operates standalone. GMT [Ho25] closes this methodological gap with standardized metrics tied directly to the Blue Angel methodology; it orchestrates Docker containers, structures runs into phases, and stores data from a configurable set of providers in a PostgreSQL database, with a frontend for

visualization and impact conversion [Gr24]. GMT contains providers for Linux RAPL (`CpuEnergyRaplMsrComponentProvider`), machine-wide PSU measurement, and model-based estimation, but no Windows-native RAPL provider existed prior to this work. To the best of our knowledge, no prior work integrates Windows RAPL measurements natively into a Docker-orchestrated energy benchmarking framework.

Intel RAPL, introduced with Sandy Bridge, provides cumulative energy counters via MSRs for the domains `cpu_package`, `cpu_cores`, `cpu_gpu`, `dram`, and `psys` (platform energy, from Skylake onward), each incrementing in units defined by MSR `0x606` (typically $\approx 15.3 \mu\text{J}$) [RT24]. Under Linux these are reachable via the `Powercap` class or directly via `MSR`; under Windows no equivalent user-space interface exists, so access requires a kernel driver.

3 Architecture

Closing this gap required bridging Windows’ kernel-level MSR access with GMT’s Python-based metric-provider framework. Two designs were explored, both introduced by this work and summarized in Table 1. They share the same driver and acquisition-binary layers (Section 3.1) and differ only in how the GMT runner is hosted; both are evaluated with identical $n = 10$ studies, enabling direct comparison (Section 4.4).

Tab. 1: Design 1 (WSL2-bridged) vs. Design 2 (native Windows)

	Design 1 (WSL2)	Design 2 (native)
GMT runtime hosted in	WSL2/Ubuntu (Linux)	Windows user space
Reaches acquisition binary via	Interop. bridge (Unix socket)	Direct subprocess call
Workload (<code>blender.exe</code>) runs as	Windows host process	Windows host process
GMT runner requires WSL2	Yes	No
Evaluated in this work	$n = 10$ (Table 3)	$n = 10$ (Table 4)

3.1 Implementation

The Windows RAPL integration is structured into four layers: (1) the `ScaphandreDrv.sys` kernel driver, (2) the `rapl_reader.exe` acquisition binary, (3) GMT’s `provider.py` metric provider, and (4) `runner.py` with the PostgreSQL database, handling phase control, orchestration, and storage. Layers 1–2 are adopted from the open-source Windows RAPL driver maintained by the Scaphandre project [Sc24]; the contribution of this work begins at Layer 3.

The `ScaphandreDrv.sys` kernel driver registers a Windows device object under `\Device\ScaphandreDriver`, exposing a single IOCTL code that accepts an MSR register address and CPU index and returns the 64-bit raw MSR value. The second layer,

`rapl_reader.exe`, is a C application for Windows x64 that opens this device handle and continuously reads RAPL values at configurable intervals. Energy is obtained by differencing the cumulative RAPL counters between consecutive reads, with counter wraparound handled explicitly, so the reported energies are independent of sampling jitter. Four mechanisms ensure measurement quality: `QueryPerformanceCounter()` anchors a wall-clock Unix timestamp at start, producing monotonic microsecond timestamps compatible with GMT’s timing mechanism; `timeBeginPeriod(1)` sets the Windows timer resolution to 1 ms so that `Sleep(99)` sleeps ≈ 99 ms instead of ≈ 15.6 ms, with a deadline-based mechanism compensating measurement time per round; a detection pass at startup disables unsupported domains; and the PKG domain serves as a driver-health canary, exiting after 5 consecutive IOCTL failures – which did not occur in any reported run.

The next layer integrates GMT’s metric provider: `rapl_reader.exe` is started as a native Windows subprocess, with stdout redirected to a file that GMT’s `BaseProvider` parses after the run. In Design 1 that subprocess is launched from within WSL2 through the interoperability layer and its output consumed by the WSL2-hosted runner; in Design 2 it is launched directly from `provider.py`. The new provider extends `BaseMetricProvider`, overriding `check_system()` (verifies driver availability) and `start_profiling()` (constructs and starts the Windows command with output redirection).

4 Evaluation – Blender Rendering

4.1 Experimental Setup

Blender 5.1.1 (build 2026-04-14) under Windows 11 was used as the test application: an open-source 3D creation suite producing reproducible, CPU-intensive workloads. The scene, *Lone Monk*, has complex geometry with several hundred objects. DRAM and CPU/GPU domains are not supported on the test system and are automatically disabled. In both designs the workload is the same `blender.exe` Windows binary executing as a host process; the designs differ only in whether it is invoked through the WSL2 interoperability layer (Design 1) or directly (Design 2). In both designs the RAPL values are acquired by `rapl_reader.exe` through the `ScaphandreDrv.sys` driver, so the two arms share an identical acquisition path. Table 2 shows the common configuration. GMT structures each run into phases: [BASELINE] (system at rest before any container starts, the reference idle level), [INSTALLATION], [BOOT], [IDLE] (containers up, no workload), [Blender Render], and [REMOVE]. Unless stated otherwise, reported energies refer to [Blender Render]; net energy subtracts a duration-scaled [BASELINE] idle baseline.

4.2 Design 1: WSL2-bridged

Design 1 was evaluated under an $n = 10$ repeated-run protocol. Table 3 shows the results. The CV of 1.6 % reflects a quiet execution environment: the idle baseline is typically ≈ 4.1 W

Tab. 2: Configuration parameters of the Blender benchmark experiment

Parameter	Value
Renderer	Cycles (CPU-only, no GPU rendering)
Scene	Lone Monk
Samples	64 spp (samples per pixel)
Resolution	1440 × 1080 px
Render duration	≈266 s (Design 1/WSL2), ≈295 s (Design 2/Windows)
Operating system	Windows 11; WSL2 Ubuntu 22.04 for Design 1, Docker Desktop for Design 2
CPU	Intel Core i7-1185G7, Turbo Boost enabled
RAM	16 GB DDR4
RAPL sampling rate	99 ms
Measured domains	cpu_package, cpu_cores, psys

Tab. 3: Design 1 – WSL2-bridged, $n = 10$ repeated runs (Scaphandre/MSR acquisition path). Render-phase energy, mean $\pm$ SD. Net-energy CVs are elevated by one anomalous run (476c3d76) in which the Windows host showed unusually high background activity during [BASELINE] (≈ 8.5 W vs. ≈ 4.1 W typical).

Domain	Total (Wh)	Net (Wh)	CV total	CV net
cpu_cores	0.890 ± 0.012	0.798 ± 0.070	1.3 %	8.8 %
cpu_package	1.166 ± 0.019	0.825 ± 0.102	1.6 %	12.4 %
psys	2.337 ± 0.044	1.018 ± 0.134	1.9 %	13.2 %

(mean ≈ 4.5 W including run 476c3d76) and the render duration highly consistent at 266 ± 2 s (CV = 0.8 %).

4.3 Design 2: Windows Native

Design 2 was evaluated under the same $n = 10$ protocol on the same machine. Table 4 shows total and net render-phase energy per domain. Total-energy CVs (1.4–2.7 %) are

Tab. 4: Design 2 – Windows native, $n = 10$ repeated runs. Render-phase energy, mean $\pm$ SD. Scaphandre/MSR acquisition path unless noted; the EMI row is a consistency check on cpu_package and shares the same baseline term in the net column.

Domain	Total (Wh)	Net (Wh)	CV total	CV net
cpu_cores (Scaphandre)	0.951 ± 0.013	0.766 ± 0.057	1.4 %	7.4 %
cpu_package (Scaphandre)	1.272 ± 0.028	0.797 ± 0.072	2.2 %	9.0 %
cpu_package (EMI)	1.271 ± 0.028	0.797 ± 0.072	2.2 %	9.0 %
psys (Scaphandre)	2.579 ± 0.070	0.992 ± 0.113	2.7 %	11.4 %

slightly higher than for Design 1, driven by run-to-run variation in [BASELINE] idle power

rather than in render-induced consumption itself: Windows idle power fluctuated between 4.9 and 7.9 W across the ten runs, consistent with background services varying their activity between sessions. Net-energy variability is higher (7.4–11.4 %) for the same reason, and the render duration is less stable at 295 ± 9 s ($CV = 3.0$ %) vs. 266 ± 2 s. The first run shows the lowest total energy across all domains, plausibly a Turbo-Boost cold-start effect. A second acquisition path – EMI (Energy Meter Interface, an inbox Windows 10+ driver interface) [msemi2024] – yields 1.271 ± 0.028 Wh for `cpu_package`, agreeing with Scaphandre/MSR to within 0.1 %. On this platform EMI is understood to surface the same RAPL package counter, so the agreement is a consistency check on the software pipeline, not an external validation of the counter; that the two channels are physically identical was not verified.

4.4 Comparison of the Two Hosting Designs

Figure 1 shows the mean power traces (`cpu_package`) for both designs across their respective $n = 10$ runs, aligned to render start. Figure 2 shows the mean energy per GMT phase. Together they reveal three systematic differences plus one methodological caveat. Because the workload binary, the acquisition path and the hardware are identical in both arms, these differences reflect the conditions under which the measurement run was carried out, not a property of the workload itself.

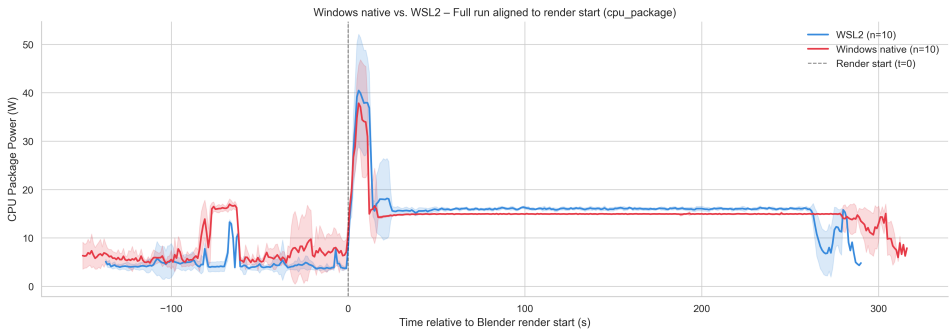


Fig. 1: Mean `cpu_package` power over time, aligned to render start – Design 1/WSL2 (blue, $n = 10$) vs. Design 2/Windows native (red, $n = 10$). Shading denotes ± 1 SD. The Windows BASELINE level is visibly higher, and the render phase extends further to the right.

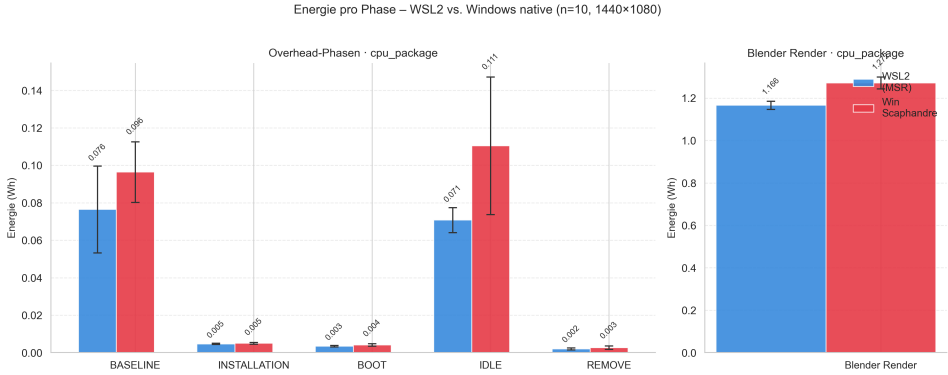


Fig. 2: Mean energy per GMT phase – Design 1/WSL2 (blue) vs. Design 2/Windows native (red), `cpu_package` ($n = 10$ each). Note the differing vertical scales: left panel, overhead phases; right panel, Blender Render phase. Error bars show ± 1 SD.

Baseline power. The [BASELINE] phase (60 s in both arms) consumes 0.096 Wh under Design 2 against 0.076 Wh under Design 1, i.e. 26 % more idle energy at equal duration (Figure 2); at ± 1 SD the two intervals overlap. On Windows 11 client systems Hyper-V uses the *root scheduler* by default, delegating all logical-processor scheduling to the Windows NT scheduler [Mi25], so background services compete for the same physical cores in both arms. The difference therefore reflects how much background activity each measurement session carried rather than a scheduling property of the designs; see Section 5.2.

Render duration. The render takes 29 ± 9 s (10.9 %) longer under Design 2 than under Design 1 for the same binary on the same physical CPU. Since rendering is CPU-bound, any CPU time taken by concurrent background processes directly prolongs it, which is consistent with the elevated baseline observed above.

Total energy. Design 2 consumes 1.272 ± 0.028 Wh for the Blender Render phase against 1.166 ± 0.019 Wh for Design 1 – a difference of +0.106 Wh (+9.1 %). Both arms show comparable average `cpu_package` power during the render (≈ 15.5 W and ≈ 15.7 W respectively), so the extra energy is explained almost entirely by the longer render duration (29 s at ≈ 15.5 W ≈ 0.12 Wh); the marginally higher Design 1 render power narrows the gap slightly.

Net energy is reported mainly as a caution: Design 2 (0.797 ± 0.072 Wh) is marginally *lower* than Design 1 (0.825 ± 0.102 Wh), because the higher idle level contributes a larger subtracted term (≈ 0.47 vs. ≈ 0.34 Wh). Net energy thus conflates workload energy with duration and baseline effects, making total phase energy the more interpretable metric here.

5 Discussion

5.1 Evaluation of the Results

Both designs produce stable, reproducible measurements, with total-energy CVs of at most 2.7 % in every domain. Design 1 shows lower variability for `cpu_package` (1.6 % vs. 2.2 %) and more stable render durations (0.8 % vs. 3.0 %). Net-energy CVs are elevated in both, driven by run-to-run variation in idle-phase power rather than measurement noise – consistent with the session-driven baseline variance reported by Ournani et al. [OBR20]. For Design 1 this is attributable almost entirely to run 476c3d76 (host baseline ≈ 8.5 W vs. ≈ 4.1 W typical); excluding it yields a net `cpu_package` CV of 0.9 %.

As an inbox interface requiring no custom kernel driver, EMI is the more deployment-friendly option where package-level granularity suffices, while Scaphandre/MSR remains necessary for `cpu_cores` and `psys`. Architecturally Design 2 is the more direct of the two; Design 1 shows lower variability, but its runs were also the quieter evening sessions (Section 5.2), so this cannot be attributed to the hosting design.

From an EMIS perspective, two findings are relevant. First, for CPU-bound workloads on Intel single-socket hardware the architecture delivers plausible, internally consistent energy data, extending GMT's comparability framework to Windows-based corporate workstations. Second, the measurement environment is not energy-neutral: the same application on the same hardware differed by 9.1 % in measured render energy depending only on how the run was orchestrated and how much background activity the host carried. For organizations benchmarking or eco-labelling Windows software, controlling that environment is therefore as relevant as the instrument.

5.2 Limitations

The Scaphandre kernel driver is unsigned and requires Windows test-signing mode, preventing production use without an Extended Validation certificate; EMI is a viable alternative at the package level. Only single-socket systems are supported (`cpuIndex = 0`). The effective sampling interval averages ≈ 103 ms against the 99 ms target, with outliers up to 130 ms; since energy is obtained by differencing cumulative counters, this affects the resolution of the power traces but not the reported energies. `psys` covers the entire platform including non-CPU components, making `cpu_package` and `cpu_cores` more informative for workload-specific analyses.

Design 2 removes the WSL2 dependency of the GMT runner, but Docker Desktop on Windows 11 may itself use a WSL2 backend, which was not recorded, so the container runtime may retain it.

No external ground-truth validation was performed: the agreement between Scaphandre/MSR and EMI to within 0.1 % shows that two independently implemented code paths read the

same hardware counter consistently, not that the counter reflects true energy draw. RAPL is well validated against physical power meters on Linux [Kh18], but this Windows code path has not, to our knowledge, been validated against an external reference.

The comparison in Section 4.4 is subject to a time-of-day confound: the Design 2 runs were conducted during morning hours, when background services are more active, and the Design 1 runs during evening hours. Since background activity is also the proposed explanation for the observed difference, the two cannot be separated from the present data, and the 9.1 % figure should be read as an observation in the tested configuration rather than as a property of either design. Disentangling them requires repeating both sets interleaved within a single session. Reported differences are point estimates without inferential statistics. Finally, all evaluations used a single CPU model (Intel Core i7-1185G7) and a sustained CPU-bound workload, so the findings are only partially generalizable to other CPU architectures or workload classes.

6 Conclusion and Outlook

This work integrates Intel RAPL energy measurement into GMT for native Windows applications – a capability GMT previously lacked – and compares a WSL2-bridged against a native Windows hosting design, without modifying core GMT components. Ten-run studies of both designs confirm stable, internally consistent measurements, with total-energy CVs of 1.6 % and 2.2 % for `cpu_package`, and a second acquisition path agreeing to within 0.1 %. Between the two designs the identical workload differed by 9.1 % in render energy; as the two sets of runs were collected at different times of day, this is reported as observed rather than as a property of either design. Either way it is a caution for cross-environment benchmarking: the conditions of measurement can account for a share of the energy attributed to a workload.

These results extend GMT’s applicability within corporate environmental information systems to Windows-based environments, addressing UN Sustainable Development Goals 9 and 12 by making the energy footprint of Windows software measurable within a certification-anchored framework. Future work includes signed driver distribution, multi-socket support, validation on further CPU architectures such as AMD, non-CPU-bound workload classes, external validation against a physical reference meter, and repeating both studies interleaved within a single session.

Availability and Acknowledgements

The GMT metric provider and the run configuration, with all component versions pinned, are available at `<repository URL>`. This work was funded by the Deutsche Bundesstiftung Umwelt (DBU) as part of the CASO project.

References

- [Fi24] Fieni, G. et al.: PowerAPI: A Python framework for building software-defined power meters. *Journal of Open Source Software* 9 (98), p. 6670, 2024.
- [Gr24] Green Coding Solutions: Introduction, 2024, <https://docs.green-coding.io/docs/prologue/introduction/>, accessed: 01/01/2025.
- [Ho25] Hoffmann, G.-D.: Green Metrics Tool: Measuring for fun and profit, 2025, <https://arxiv.org/html/2506.23967v1>, accessed: 01/01/2025.
- [Kh18] Khan, K. N. et al.: RAPL in Action: Experiences in Using RAPL for Power Measurements. *ACM Transactions on Modeling and Performance Evaluation of Computing Systems* 3 (2), 9:1–9:26, 2018.
- [Mi25] Microsoft Corporation: Manage Hyper-V Hypervisor Scheduler Types, 2025, <https://learn.microsoft.com/en-us/windows-server/virtualization/hyper-v/manage/manage-hyper-v-scheduler-types>, accessed: 08/08/2026.
- [Na11] Naumann, S. et al.: The GREENSOFT Model: A reference model for green and sustainable software and its engineering. *Sustainable Computing: Informatics and Systems* 1 (4), pp. 294–304, 2011.
- [NGK20] Naumann, S.; Guldner, A.; Kern, E.: The Eco-label Blue Angel for Software – Development and Components. In: *Advances and New Trends in Environmental Informatics – Digital Twins for Sustainability*. Progress in IS, Springer, Cham, pp. 79–89, 2020.
- [OBR20] Ournani, Z.; Belgaid, M. C.; Rouvoy, R.: Taming Energy Consumption Variations in Systems Benchmarking. In: *Proceedings of the 11th ACM/SPEC International Conference on Performance Engineering (ICPE)*. ACM, pp. 58–69, 2020.
- [RT24] Raffin, G.; Trystram, D.: Dissecting the software-based measurement of CPU energy consumption: a comparative analysis, 2024, <https://arxiv.org/abs/2401.15985>, accessed: 01/01/2025.
- [Sc19] Schöne, R. et al.: Energy Efficiency Features of the Intel Skylake-SP Processor and Their Impact on Performance. In: *2019 International Conference on High Performance Computing & Simulation (HPCS)*. IEEE, Dublin, pp. 399–406, 2019.
- [Sc24] Scaphandre Project: Scaphandre: A Metrology Agent for Energy Consumption, 2024, <https://github.com/hubblo-org/scaphandre>, accessed: 01/01/2025.
- [SCD24] Sallou, J.; Cruz, L.; Durieux, T.: EnergiBridge: Empowering Software Sustainability through Cross-Platform Energy Measurement. In: *Proceedings of the 46th International Conference on Software Engineering: Companion Proceedings (ICSE Companion)*. ACM, 2024.
- [Wo05] Wohlgemuth, V.: *Betriebliche Umweltinformationssysteme*. Springer, Berlin, Heidelberg, 2005.

Joint Evaluation of the Green Coding Roadmap Training

Unlocking the Potential of Green Coding into Practice – Evaluation of the BUIS Tage Workshop 2026 Outcomes

Nick Vicente Roland Kriesing Garcia ¹, Dennis M. Junger ¹, and Volker Wohlgemuth ¹

Abstract:

Sustainable software engineering requires learning resources that translate fragmented research findings into accessible and actionable guidance for practitioners. To address this need, we developed a Green Coding Roadmap that organises core concepts and practices across six phases: understanding the problem, learning the fundamentals, measuring and assessing impacts, writing greener code, optimising infrastructure, and embedding sustainability into processes and organisations. The roadmap combines theoretical foundations with practical tools, standards, and recommended actions in a roadmap-based format familiar to developers.

This paper reports on an exploratory participatory workshop study in which the roadmap was tested with twelve participants from academic and professional backgrounds at BUIS-Tage 2026. Through guided small-group work and plenary discussion, participants evaluated the roadmap with respect to content relevance, sequencing, clarity, and usability. The findings indicate that the roadmap is perceived as a promising framework for green coding education, while also revealing opportunities for improvement in narrative coherence, visual organisation, and the explicit treatment of measurement and trade-offs. The paper concludes with design implications for refining the roadmap and for developing workshop-based training formats in sustainable software engineering.

This paper is structured as follows: First, we introduce the Green Coding Roadmap and its development through Activities 3.1–3.4 of the Intervention and Training work package, including its six phases and key themes. Second, we present the participatory workshop study at BUIS-Tage 2026, detail the evaluation results across all roadmap chapters, and conclude with design implications for refining the roadmap and developing workshop-based training formats in sustainable software engineering.

Keywords: Green Coding, Sustainable Software Engineering, Green Software, Software Sustainability, Learning Roadmap, Participatory Design, Workshop Study, Sustainability Education, Green IT, Software Engineering Education

¹ HTW Berlin, University of Applied Sciences, Industrial Environmental Informatics Unit, Wilhelminenhofstr. 75A, 12459 Berlin, Germany,
s0587788@HTW-Berlin.de,  <https://orcid.org/0009-0007-6761-3169>;
dennis.junger@htw-berlin.de,  <https://orcid.org/0000-0002-9271-2261>;
volker.wohlgemuth@htw-berlin.de,  <https://orcid.org/0009-0000-3150-1746>

1 Introduction and Motivation

Sustainability concerns in computing have led to a growing body of work on Green IT and green coding, including guidelines, metrics, and tools to reduce the environmental impact of software systems. Research in green software engineering has led to a wide range of concepts, methods, and tools, covering areas from energy-efficient algorithms and architectures to carbon-aware scheduling and data centre optimisation. At the same time, educational efforts have explored curricula, courses, and guidelines to build competencies among students and practitioners. However, many of these contributions remain fragmented and are mainly presented in academic publications, which can make it difficult for practitioners to incorporate them into their day-to-day development workflows. There is still limited evidence on how developers perceive structured green coding roadmaps and which design choices support or hinder their adoption. After the project “Potentials of Greencoding” [Wo24a], in which the HTW Berlin team demonstrated that green coding can be integrated into curricula, the following project “Unlocking Potentials of Green Coding into Practice” [Wo24b] seeks to translate research on sustainable software development into a structured training programme. The work package “*Intervention and Training*” is structured into five activities. Activity 3.1 focuses on conducting structured expert interviews to identify key training needs. These were implemented through “Semi-structured expert interview” approach [Lo25]. An initial workshop held at EcoCompute 2024 ([Gr24a]), and published in [JW24] under the title “Bridging the Skills Gap: Insights from Professional Software Developers into the Challenges and Opportunities Faced by Young Career Starters Integrating Green Coding into Practice”, identified core requirements for the transition from academia to professional practice for junior developers. The findings highlight a clear need for competencies in change management, as well as organisational and process-related knowledge. These insights were presented, refined, and further developed during a subsequent workshop at the EnviroInfo 2025 conference ([Ge24]). They were then consolidated into a structured focus group study [Ki95], the results of which were published in [JW25] under the title “Requirements Study for a Professional Green Coding Course Format: A Focus Group Study to Explore Strategies for Implementing Green Coding Training in Software Development”. Subsequently, a preliminary academically oriented proposal was presented outlining how a Green Coding master’s degree program could be implemented academically through two master’s thesis projects and their respective content. These content elements should now be identified within Activity 3.3 and translated into an appropriate training concept. One issue that had already emerged during EnviroInfo 2025 [Ge24] is the implementation of a testing and certification system. This system is currently in the preparatory phase and will be implemented following the completion of the subsequent work packages and the integration of critical feedback. The current step of Activity 3.4 has produced the master training concepts: academically collected papers, enriched with established knowledge from years of experience and prior literature reviews [Ju23], presented in the form of a roadmap. Despite growing research on green software engineering, there is limited empirical evidence on how developers perceive structured, roadmap-based training and which design choices support its adoption; this paper addresses that gap through an exploratory participatory evaluation of a concrete

Green Coding Roadmap artefact. Our work builds on the popular open learning platform `roadmap.sh`, which serves several million users and has attracted hundreds of thousands of stars on GitHub [AC26]. We use this familiar roadmap metaphor to embed sustainability content as a first-class competency alongside other software engineering skills.

The present paper documents Activity 3.5 of the *Intervention and Training* phase, in which this pilot training is tested and evaluated. Through a participatory workshop at BUIS-Tage 2026 with twelve participants from academic and professional backgrounds, we systematically collected feedback on content relevance, sequencing, clarity, and usability. This represents the final step linking preceding design activities (3.1–3.4) with empirical evidence on the effectiveness and acceptance of the resulting training intervention.

2 The Green Coding Roadmap

Before detailing its structure, it is useful to situate the roadmap's first phase within established sustainability theory. Sustainability is commonly framed through concepts such as planetary boundaries [Ro09] and the triple bottom line of environmental, social, and economic dimensions [ER99], both of which predate and extend beyond computing. The Green Coding Roadmap grounds its "Understanding the Problem" phase in this broader lineage, connecting general sustainability thinking. For instance the Limits to Growth report already referenced in the roadmap content [MRA72], to ICT-specific framings such as GHG accounting scopes and the distinction between embodied and operational carbon. This linkage is intended to prevent green coding from being perceived as an isolated technical concern, and instead to position it as one instance of a wider sustainability discourse.

The Green Coding Roadmap is conceived as a structured, web-based learning path that translates green coding concepts into actionable steps for developers. It is organised into several major chapters that reflect different stages and dimensions of sustainable software engineering, including *Understanding the Problem*, *Learning the Fundamentals*, *Measuring and Assessing Impacts*, *Writing Greener Code*, *Optimising Infrastructure*, *Green Computing Toolkit*, and *Embedding Sustainability*.

Each chapter contains nodes that introduce key concepts, link to tools and resources, and suggest practical actions, with the overall roadmap providing milestones and a sense of progression. The design aims to offer not only technical guidance but also to incorporate organisational and cultural aspects, for instance, through chapters on governance, standards, and sustainability culture.

The roadmap is hosted on `roadmap.sh` so that it builds on an existing open-source infrastructure and user community [JKW26]. By adopting the familiar roadmap metaphor and visual style, we seek to lower entry barriers for developers and position sustainability as a first-class competency alongside other software engineering skills.

2.1 Content of the Roadmap

The *Green Coding Roadmap* organises theory and practical guidance into six phases: (1) understanding the problem, (2) learning the fundamentals, (3) measuring and assessing, (4) writing greener code, (5) optimising infrastructure, and (6) embedding sustainability into processes and organisations. Each phase combines theoretical foundations, practical methods, tools, standards, and measurement/certification approaches, guiding teams from sustainability awareness to concrete changes in software design, implementation, and operations.

2.2 Key Themes and Sections

The following list presents a representative excerpt of the implemented roadmap content for clarity; all comprehensive content is accessible via `roadmap.sh` [ro26].

- **Foundations and context:** History of sustainability [MRA72], ICT emissions [Fr21], GHG Scopes[RB04], carbon intensity [Gr24b].
- **Measurement and assessment:** Carbon footprint [RB04], SCI [Gr24c], Blue Angel [NGK21], cloud dashboards [Am26].
- **Coding and runtime practices:** Algorithmic complexity [Ca24], data structures [Ha18], language trade-offs [Ke25], benchmarking [Pe21].
- **Frontend and API efficiency:** Asset optimisation (images, fonts, scripts) [MD], SSR/SSG/CSR [Ad19], caching [Ka21a], API/database efficiency [Mi26].
- **AI/ML considerations:** Training vs. inference energy [SGM19], model compression [Da24], prompt efficiency [Pr26], token reduction [Ni24].
- **Infrastructure and operations:** PUE [Ka21b], renewable procurement [RE26], edge computing [An23], autoscaling [Ve25], carbon-aware scheduling [Gr24d].
- **Process and organizational adoption:** Green SDLC [Ru10], change management [CS22], GSMM [Gr25], ESG [Ga26], AI Act [EU24].
- **Tools, standards, and learning:** GREENSOFT [Na11], carbon-aware SDK [Gr26], community links [Bu25].

Overall, the roadmap offers a coherent and structured perspective on sustainable software development, integrating conceptual foundations, measurement approaches, implementation practices, infrastructure optimisation, and organisational adoption. It is particularly well suited for workshops, as it facilitates both technical learning and critical reflection on embedding sustainability into everyday software engineering practice.

The content of the roadmap was derived from multiple converging sources rather than compiled ad hoc: prior literature review work on green coding competencies [Ju23], and

established reference frameworks and standards in the field, for instance the GREENSOFT model [Na11], the Software Carbon Intensity specification [Gr24c], and the Blue Angel eco-label for software [NGK21].

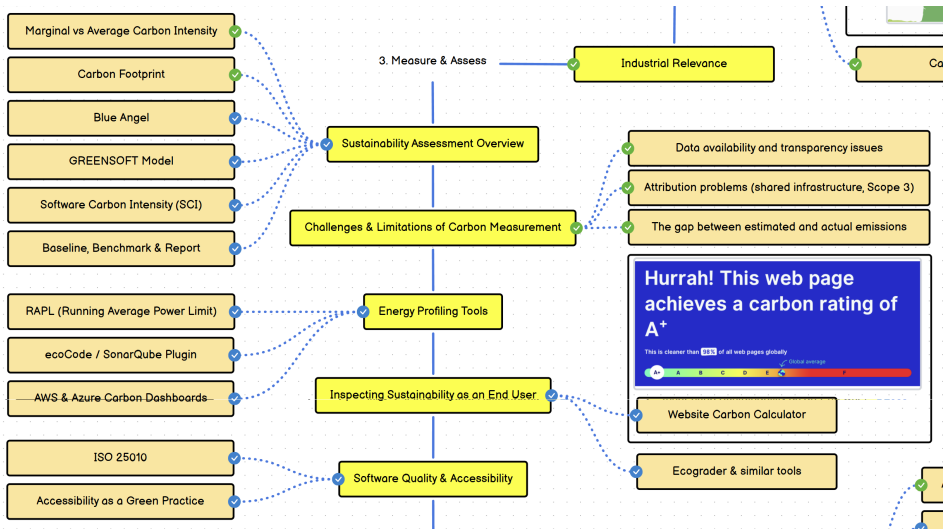


Fig. 1: Excerpt of the Green Coding Roadmap on roadmap.sh, showing the Measure & Assess phase, including the Software Carbon Intensity (SCI) node.

2.3 Sequencing Rationale

The ordering of the six phases follows a deliberate progression from conceptual grounding to operational change, corresponding broadly to the Revised Bloom’s Taxonomy of cognitive levels [Co05]. The early chapters, Understanding the Problem and Learning the Fundamentals, are positioned to build remembering and understanding, establishing the vocabulary and awareness, for instance, embodied versus operational carbon and GHG scopes, on which later technical decisions depend. Measuring and Assessing Impacts is placed next, moving towards understanding and analysing, on the assumption that developers require a baseline and a shared measurement vocabulary before the effects of coding or infrastructure choices can be meaningfully evaluated. Writing Greener Code and Optimising Infrastructure correspond to applying and analysing, while Embedding Sustainability also addresses these levels, treating organisational and cultural adoption as building on individual technical competency. This progression is intended as a general orientation rather than a strict rule that every node must follow.

Within chapters, the intended logic follows standards and norms first, then tools, then patterns and guidelines, and finally additional resources and communities, so that practitioners encounter what is expected before being directed toward how it can be achieved. As reported

in Section 4, this intra-chapter logic was not always transparent to workshop participants, which motivates the revisions to visual sequencing and intermediate nodes discussed in Section 6.

3 Workshop Design and Methods

3.1 Participants and Setting

The workshop was conducted as part of the BUIS-Tage 2026 at HTW Berlin on 2 June 2026 [Ge26]. It was designed as a participatory evaluation of the Green Coding Roadmap within the project “Unlocking Potentials of Green Coding 2024–2026”, which aims to connect sustainability research with software development practice [JKW26].

Twelve participants attended the workshop and were split into four groups, with two to four participants in each group. They came from a range of academic and professional backgrounds in computer science and related areas. Ten participants (83.3%) were enrolled in higher education programmes, while two participants (16.7%) came from professional practice in the transport and public sectors.

Among the students, four participants (33.3% of the total sample) were master’s students in business informatics, three (25.0%) were PhD candidates at the University of Oldenburg with IT-related specialisations, and three (25.0%) were students from universities of applied sciences or other IT-related programmes. This composition reflects the intended focus on prospective and early-career software professionals, complemented by selected practical perspectives. Given this composition, with the large majority of participants drawn from higher education and only two from professional practice, the study is best understood as an exploratory evaluation of the roadmap’s design and framing.

3.2 Materials

The workshop was supported by four complementary materials: a slide deck introducing green coding and the roadmap concept, printed posters with AI-generated imagery created using NotebookLM [Go25], the Green Coding Roadmap embedded via `roadmap.sh`, and digital worksheets for structured group feedback. NotebookLM was used to generate illustrative imagery from the workshop’s topic information; as with other generative AI tools, it can introduce unintended stylistic or representational bias. Its output was therefore treated as illustrative material, without influence on the roadmap’s conceptual or textual content, which was authored independently by the research team.

Each worksheet included guiding questions tailored to the assigned roadmap section(s), covering participants’ prior knowledge, their perceptions of clarity and completeness, the logical sequencing of the content, and possible improvements.

All workshop materials, including the AI-generated images and the participants' questions and answers, are available in our Zenodo dataset [JKW26]. Due to the GDPR (DSGVO), we removed all personal information of the participants.

3.3 Workshop Procedure

The workshop followed a three-phase format consisting of a plenary introduction, guided small-group work, and final plenary presentations.

In the introductory phase, the facilitators presented the motivation for green coding, the project context, and the concept of structuring knowledge through a roadmap hosted on `roadmap.sh`. Participants were then assigned to four groups using thematic posters and given access to specific roadmap sections.

After a brief exploration phase of about ten minutes, the groups completed five focused work sprints of approximately five minutes each. In each sprint, participants answered one guiding question from the worksheet and documented their responses as concise bullet points. The questions covered prior expectations, content quality, coherence and completeness, structural ordering, and overall suggestions for improvement.

In the final phase, each group presented its findings in the plenary. The presentations began with a summary of the assigned roadmap section(s), followed by the main observations and a specific recommendation for improvement. The presentations were followed by short discussions to clarify the feedback and connect insights across groups.

3.4 Data Collection and Analysis

The primary data sources consisted of the completed group worksheets and notes from the plenary presentations and discussions. One researcher conducted an initial inductive, open coding of this material, focusing on clarity, completeness, sequencing, and motivational aspects of the roadmap.

The resulting codes were then reviewed and iteratively grouped into higher-level themes through joint discussion, with disagreements or ambiguous codes resolved by consensus among the research team. This collaborative review step served to reduce the influence of any single researcher's interpretation on the final thematic groupings. The resulting themes capture the main strengths and weaknesses identified by participants and inform the concrete design implications for improving the roadmap presented in Section 5. The results are presented in the following section, structured by roadmap chapter.

4 Results

4.1 Understanding the Problem and Fundamentals

For the early chapters on *Understanding the Problem* and *Learning the Fundamentals*, participants emphasised that concepts such as embodied carbon, carbon intensity, and human-environment interaction should be further distinctly anchored in the problem framing. Some teams expected these topics to be introduced earlier and more prominently in the roadmap, as they form the conceptual basis for later strategies.

Participants also noted that transitions between nodes could be smoother and that the visual layout should support a clearer narrative from general sustainability concepts to their implications for software. They suggested increasing the number of intermediate nodes and improving the connections to better reflect the logical flow between topics.

4.2 Measure & Assess and Green Coding Toolkit

In the chapters on measuring and assessing impacts and on the Green Coding Toolkit, teams appreciated the concrete pointers to tools and software but found the overall structure and sequencing challenging. Several participants expected standards and norms to be introduced before tools, followed by patterns and guidelines, and only then additional resources and communities.

There was also confusion about the allocation of specific tools, for example, whether certain calculators and estimators should be classified as key tools or as additional resources. Participants argued that AI-based tools deserved a separate chapter due to their distinct characteristics and the rapid pace at which this area is developing. More generally, they requested richer descriptions of the main blocks, introductory sentences explaining the purpose and importance of each section, and more motivational language and calls to action.

4.3 Write Greener Code and Optimise Infrastructure

For the chapters on writing greener code and optimising infrastructure, participants largely agreed that the topics were relevant but raised critical questions about measurement and trade-offs. They pointed out that clients and client-side rendering were underrepresented and asked how energy consumption would be measured and compared when specific optimisation techniques were applied.

Some participants suggested that the term “optimising” might be misleading if measurement is not explicitly addressed, proposing alternatives such as “analysing” or “adapting” infrastructure. They also questioned whether a single metric such as Power Usage Effectiveness (PUE) would be sufficient to characterise data centre performance and called for a clearer representation of trade-offs and meta-level decisions.

4.4 Embedding Sustainability

In the *Embed Sustainability* chapter, participants welcomed the clear division into three sub-dimensions: planning, design and implementation; governance and ecological standards; and culture and people. They appreciated that this chapter goes beyond purely technical aspects and highlights policy and behavioural dimensions of sustainability.

At the same time, they reported initial confusion about the visual arrangement of governance and standards items directly under the main heading, and the lack of numbering at the beginning of subsequent chapters. Participants suggested reversing the sequence to start with awareness building, then standards, and then implementation, arguing that awareness is a prerequisite for meaningful adoption of standards and green coding practices. They also requested an abbreviation glossary, clearer visual separation between chapters, more technical examples (e.g., autoscaling based on current carbon intensity), and short introductory texts that situate each section within the overall context.

5 Discussion

The workshop results highlight that while participants overall found the content of the roadmap chapters pertinent and meaningful, they also identified several structural and presentational issues that hinder comprehension and motivation. In particular, the need for clearer storylines per chapter, more explicit sequencing of concepts and tools, and better optic cues emerged as recurrent themes.

From a design perspective, participants' feedback suggests that green coding roadmaps should not only list topics and tools but also articulate the problem they address, the learning goals, and the expected outcomes for each section. Participants' calls for introductory questions, explanatory sentences, and motivational phrasing indicate that, at least for this group, narrative framing was perceived as being as fundamental as technical correctness in encouraging engagement. A further implication concerns the explicit treatment of trade-offs and meta-level decisions, particularly in infrastructure optimisation. Participants asked how environmental, economic, and reliability objectives interact, and how developers can make informed decisions while considering constraints such as the political context, energy markets, and cybersecurity risks. Making such trade-offs more visible in the roadmap may help learners recognise sustainability as a systemic challenge rather than a purely technical add-on, though this remains a design hypothesis motivated by our findings rather than a tested effect.

A concrete mechanism for achieving this, motivated directly by participants' critique of Power Usage Effectiveness as a possibly insufficient single metric, would be to complement individual metric nodes with short comparative overviews. Similarly, scenario-style framings that connect a technical choice to its environmental, economic, and reliability implications could make explicit the kind of meta-level reasoning participants found underrepresented.

Developing and validating such comparative and scenario-based formats was beyond the scope of the present pilot study and workshop evaluation, but constitutes a concrete direction for the next iteration of the roadmap.

6 Conclusion and Future Work

This paper reports on a participatory workshop that tested an initial Green Coding Roadmap with professional and student teams, using structured group work and short presentations to elicit feedback on content, structure, and usability. The results of this exploratory pilot indicate that the roadmap is perceived as a valuable starting point for organising green coding knowledge but requires refinement in terms of narrative coherence, sequencing, and motivational framing. In future work, we plan to incorporate the feedback into a revised roadmap version with clearer chapter introductions, re-ordered sections, and dedicated treatment of AI-related tools and trade-offs. We further intend to trial the roadmap against concrete test cases spanning heterogeneous project types. We further aim to evaluate the updated roadmap with a broader participant group, including industry practitioners and members of the roadmap.sh community, and to explore interactive multiple-choice tests and certificates as assessment elements.

Acknowledgements

This work was funded by the Internet Society Foundation². We would like to thank the Fachgruppe Betriebliche Umweltinformationssysteme (FG BUIS)³ of the Gesellschaft für Informatik for providing the opportunity to host the workshop as part of the BUIS-Tage. We also extend our sincere thanks to all participants, whose contributions and engagement were essential for the validation of our research.

References

- [AC26] Ahmed, K.; Contributors: developer-roadmap: Interactive roadmaps, guides and other educational content to help developers grow in their careers, <https://github.com/nilbuild/developer-roadmap>, GitHub repository, accessed 2026-06-08, 2026.
- [Ad19] Addy Osmani and Jason Miller: Rendering on the Web, Last updated: 2026-01-05, web.dev, 2019, <https://web.dev/articles/rendering-on-the-web>, accessed: 06/11/2026.
- [Am26] Amazon Web Services, Inc.: AWS Sustainability Console, Amazon Web Services, 2026, <https://aws.amazon.com/sustainability/tools/console/>, accessed: 06/11/2026.
- [An23] Ananth, K. et al.: Green IoT Edge Computing Towards Sustainable and Distributed Data Processing. In: 2023 10th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON). Vol. 10, pp. 1644–1649, 2023.

² <https://www.isocfoundation.org/> [2026-06-06]

³ <https://fa-ui.gi.de/fachausschuss/fachgruppen/fg-buis/> [2026-06-06]

- [Bu25] Bundesverband Green Software e.V.: Green Software Tools Landscape, Bundesverband Green Software e.V., 2025, <https://landscape.bundesverband-green-software.de/>, accessed: 06/11/2026.
- [Ca24] Carter, K. et al.: Energy and Time Complexity for Sorting Algorithms in Java, 2024, arXiv: 2311.07298 [cs.SE], <https://arxiv.org/abs/2311.07298>.
- [Co05] Conklin, J.: A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives. *Educational Horizons* 83 (3), pp. 154–159, 2005, <http://www.jstor.org/stable/42926529>, accessed: 08/18/2026.
- [CS22] Chowdhury, A.; Shil, N.: Understanding change management in organizational context: revisiting literature. *Management and Entrepreneurship: Trends of Development* 1, pp. 28–43, 2022.
- [Da24] Dantas, P. et al.: A comprehensive review of model compression techniques in machine learning. *Applied Intelligence* 54, pp. 11804–11844, 2024.
- [ER99] Elkington, J.; Rowlands, I. H.: Cannibals with forks: The triple bottom line of 21st century business. *Alternatives Journal* 25 (4), p. 42, 1999.
- [EU24] EU AI Act Explorer: EU Artificial Intelligence Act (AI Act), 2024, <https://artificialintelligenceact.eu/>, accessed: 06/11/2026.
- [Fr21] Freitag, C. et al.: The real climate and transformative impact of ICT: A critique of estimates, trends, and regulations. *Patterns* 2 (9), p. 100340, 2021, <https://www.sciencedirect.com/science/article/pii/S2666389921001884>.
- [Ga26] Gartner: ESG Management and Reporting Software Reviews, Gartner Peer Insights, 2026, <https://www.gartner.com/reviews/market/esg-management-and-reporting-software>, accessed: 06/11/2026.
- [Ge24] German University in Cairo: EnviroInfo 2025- Transboundary Environmental Challenges: Digital Inclusion for Sustainable Development, 2024, <https://enviroinfo.org/>, accessed: 06/07/2025.
- [Ge26] Gesellschaft für Informatik (GI): 15. BUIS-Tage 2026: Nachhaltigkeit und Digitalisierung, <https://buis-tage.gi.de/>, Zugriff am 09.06.2026, 2026.
- [Go25] Google: NotebookLM, 2025, <https://notebooklm.google.com/>, accessed: 06/05/2026.
- [Gr24a] Green Coding Solutions GmbH: EcoCompute - The first engineering conference on sustainability in hardware & software, 2024, <https://www.eco-compute.io/>, accessed: 06/07/2024.
- [Gr24b] Green Software Foundation: Carbon Intensity, Green Software Foundation, 2024, <https://learn.greensoftware.foundation/carbon-awareness>, accessed: 06/11/2026.
- [Gr24c] Green Software Foundation: Software Carbon Intensity (SCI) Specification, version 1.1.0, 2024, <https://sci.greensoftware.foundation/>, accessed: 06/11/2026.
- [Gr24d] Green Software Foundation: Temporal shifting, Green Software Foundation, 2024, <https://learn.greensoftware.foundation/carbon-awareness/%5C#temporal-shifting>, accessed: 06/11/2026.
- [Gr25] Green Software Foundation: Green Software Maturity Matrix, Green Software Foundation, 2025, <https://maturity-matrix.greensoftware.foundation/>, accessed: 06/11/2026.
- [Gr26] Green Software Foundation: Carbon Aware SDK, Green Software Foundation, 2026, <https://github.com/Green-Software-Foundation/carbon-aware-sdk>, accessed: 06/11/2026.
- [Ha18] Ha, P. H. et al.: D2.1 Models for energy consumption of data structures and algorithms, 2018, arXiv: 1801.09992 [cs.DC], <https://arxiv.org/abs/1801.09992>.

- [JKW26] Junger, D. M.; Kriesing Garcia, N. V. R.; Wohlgemuth, V.: Workshop Green Coding – Outcomes (BUIS-Tage 2026, HTW Berlin), Zenodo, 2026, <https://doi.org/10.5281/zenodo.20626106>.
- [Ju23] Junger, D. et al.: Potentials of Green Coding - Findings and Recommendations for Industry, Education and Science. In: INFORMATIK 2023 - Designing Futures: Zukünfte gestalten. Gesellschaft für Informatik e.V., Bonn, pp. 1289–1299, 2023.
- [JW24] Junger, D.; Wohlgemuth, V.: Bridging the Skills Gap: Insights from Professional Software Developers into the Challenges and Opportunities Faced by Young Career Starters Integrating Green Coding into Practice, Bonn, 2024.
- [JW25] Junger, D. M.; Wohlgemuth, V.: Requirements Study for a Professional Green Coding Course Format: A Focus Group Study to Explore Strategies for Implementing Green Coding Training in Software Development. In: EnviroInfo 2025. Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Potsdam, Germany, pp. 121–130, 2025, https://www.researchgate.net/publication/395716093_Requirements_Study_for_a_Professional_Green_Coding_Course_Format_A_Focus_Group_Study_to_Explore_Strategies_for_Implementing_Green_Coding_Training_in_Software_Development.
- [Ka21a] Kabir, A. et al.: Energy-aware caching and collaboration for green communication systems. *Acta Montanistica Slovaca*, pp. 47–59, 2021.
- [Ka21b] Kat Henshaw: PUE, CUE and WUE: What do these three metrics represent and which one is the most important?, Blog post on data center sustainability metrics, Submer, 2021, <https://submer.com/blog/pue-cue-and-wue-what-do-these-three-metrics-represent-and-which-is-one-is-the-most-important/>, accessed: 06/11/2026.
- [Ke25] van Kempen, N. et al.: It's Not Easy Being Green: On the Energy Efficiency of Programming Languages, 2025, arXiv: 2410.05460 [cs.PL], <https://arxiv.org/abs/2410.05460>.
- [Ki95] Kitzinger, J.: Qualitative Research: Introducing focus groups. *BMJ* 311 (7000), pp. 299–302, 1995, eprint: <https://www.bmj.com/content>, <https://www.bmj.com/content/311/7000/299>.
- [Lo25] Louklander, M.: Just Asking: Semi-Structured Expert Interviews as a Method in Legal Research. *Verfassungsblog*, 2025, <https://verfassungsblog.de/just-asking/>.
- [MD] MDN Web Docs Contributors: Multimedia: Images and Video (Web Performance), Mozilla Developer Network, https://developer.mozilla.org/en-US/docs/Learn%5C_web%5C_development/Extensions/Performance/Multimedia, accessed: 06/11/2026.
- [Mi26] Microsoft: API design best practices, Microsoft Learn, 2026, <https://learn.microsoft.com/en-us/azure/architecture/best-practices/api-design>, accessed: 06/11/2026.
- [MRA72] Meadows, D.; of Rome, C.; Associates, P.: The Limits to Growth: A Report for the Club of Rome's Project on the Predicament of Mankind. Universe Books, 1972.
- [Na11] Naumann, S. et al.: The GREENSOFT Model: A reference model for green and sustainable software and its engineering. *Sustainable Computing: Informatics and Systems* 1 (4), pp. 294–304, 2011, <https://www.sciencedirect.com/science/article/pii/S2210537911000473>.
- [NGK21] Naumann, S.; Guldner, A.; Kern, E.: The Eco-label Blue Angel for Software—Development and Components. In (Kamilaris, A. et al., eds.): *Advances and New Trends in Environmental Informatics*. Springer International Publishing, Cham, pp. 79–89, 2021.

- [Ni24] Nicoomanesh, A.: Token Efficiency and Compression Techniques in Large Language Models: Navigating Context-Length Limits, Medium article on LLM token compression techniques, 2024, <https://medium.com/@anicomanesh/token-efficiency-and-compression-techniques-in-large-language-models-navigating-context-length-05a61283412b>, accessed: 06/11/2026.
- [Pe21] Pereira, R. et al.: Ranking programming languages by energy efficiency. *Science of Computer Programming* 205, p. 102609, 2021, <https://www.sciencedirect.com/science/article/pii/S0167642321000022>.
- [Pr26] PromptingGuide.ai: Prompt Engineering Guide, Educational resource on prompt engineering techniques, 2026, <https://www.promptingguide.ai/>, accessed: 06/11/2026.
- [RB04] Ranganathan, J.; Bhatia, P.: *The Greenhouse Gas Protocol: a Corporate Accounting and Reporting Standard*, Revised Edition. 2004.
- [RE26] RECS Energy Certificate Association: RECS Energy Certificate Association, 2026, <https://recs.org/>, accessed: 06/11/2026.
- [Ro09] Rockström, J. et al.: A safe operating space for humanity. *nature* 461 (7263), pp. 472–475, 2009.
- [ro26] roadmap.sh: Green Coding, <https://roadmap.sh/r/green-coding-eqprq>, Accessed 2026-06-11, 2026.
- [Ru10] Ruparelia, N. B.: Software development lifecycle models. *SIGSOFT Softw. Eng. Notes* 35 (3), pp. 8–13, 2010, <https://doi.org/10.1145/1764810.1764814>.
- [SGM19] Strubell, E.; Ganesh, A.; McCallum, A.: Energy and Policy Considerations for Deep Learning in NLP, 2019, arXiv: 1906.02243 [cs.CL], <https://arxiv.org/abs/1906.02243>.
- [Ve25] Vercel: Concurrency Scaling for Vercel Functions, Documentation on automatic scaling and concurrency limits for Vercel Functions, 2025, <https://vercel.com/docs/functions/concurrency-scaling>, accessed: 06/11/2026.
- [Wo24a] Wohlgemuth, V.: Potentials of Green Coding, Abgerufen am 06.10.2025, 2024, <https://wohlgemuth.f2.htw-berlin.de/forschung/potentials-of-green-coding/>.
- [Wo24b] Wohlgemuth, V.: Unlocking the Potential of Green Coding into Practice, Forschungsprojekt; Projektlaufzeit: 01.10.2024–30.09.2026; Mittelgeber: Internet Society Foundation; Kooperationspartner: German Informatics Society, Umwelt-Campus Birkenfeld, Hochschule für Technik und Wirtschaft Berlin (HTW Berlin), 2024, <https://wohlgemuth.f2.htw-berlin.de/forschung/unlocking-the-potential-of-green-coding-into-practice/>.

Toward a Framework for Performance and Carbon Footprint Assessment of Payment Channel Network Solutions

Danila Valko ^{1,2} and Jorge Marx Gómez ¹

Abstract: Payment Channel Networks (PCNs) have emerged as a critical scalability layer for blockchain infrastructures, most prominently exemplified by the Lightning Network built atop Bitcoin. While existing research on PCNs predominantly focuses on performance optimization, sustainability-related metrics such as energy consumption and carbon footprint remain largely under-reported.

This paper introduces a framework on performance and carbon footprint evaluation for PCN solutions and provides actionable work-in-progress recommendations for researchers and developers grounded in the FAIR+S framework, which extends FAIR by explicitly incorporating sustainability principles. Methodologically, we adopt a systematic literature review combined with metric mining across 313 PCN publications to identify, classify, and assess reported performance indicators. Our analysis revealed that the only 134 publications have reported quantitative performance metrics: 89 (66.4%) of them proposed algorithmic or topological performance improvements, 25 (18.7%) disclosed sufficient hardware context, and only 3 (2.2%) explicitly addressed energy efficiency or sustainability. These 134 publications constitute a performance-oriented subset that served as the source for the underlying metric mining and taxonomy construction.

Based on these, we propose a structured metrics taxonomy that links conventional performance dimensions, including time, computation, communication overhead, and success rate, to energy- and carbon-relevant proxies using community-accepted tools and specifications. The resulting FAIR+S-aligned evaluation framework promotes transparent and environmentally accountable PCN research.

Keywords: Performance, Carbon footprint, Payment Channel Networks, FAIR, FAIR+S

1 Introduction

Blockchain scalability solutions have driven rapid growth in Payment Channel Networks (PCNs) since their introduction in 2016[PD16]. PCNs enable off-chain transaction routing to reduce congestion and fees in distributed ledgers. However, the environmental debate surrounding blockchain systems, initially focused on proof-of-work mining, has expanded toward infrastructure layers such as routing nodes and channel management[KS23; VK24].

At the same time, research governance frameworks have advanced reproducibility and openness. The FAIR principles emphasize findability, accessibility, interoperability, and

¹ Carl von Ossietzky Universität Oldenburg, Ammerländer Heerstraße, 114-118, 26129 Oldenburg, Germany, danila.valko@offis.de,  <https://orcid.org/0000-0002-8058-7539>;
jorge.marx.gomez@uol.de,  <https://orcid.org/0000-0002-7833-7549>

² OFFIS – Institute for Information Technology, Escherweg 2, 26121 Oldenburg, Germany, danila.valko@offis.de,  <https://orcid.org/0000-0002-8058-7539>

reusability of research artifacts, later extended to software and data through FAIR4RS[Ba22; Wi16]. Despite their success, these frameworks omit sustainability-related dimensions such as energy efficiency or carbon emissions[Va26]. This omission is critical in software development and computing domains where performance improvements may inadvertently increase energy demand[Be15; Du20]. Performance without sustainability transparency risks shifting environmental costs rather than reducing them, which is especially significant in the context of blockchain-based payment networks.

This paper addresses this gap by:

- Utilizing a systematic literature review and exploring the state of FAIR principles adoption in PCN research.
- Mining and classifying reported performance metrics, and mapping them to energy and carbon footprint proxy measurements.
- Proposing a FAIR+Sustainability-aligned performance and carbon footprint assessment framework for PCN research, equipped with recommendations and community-accepted tools.

2 Methodology

2.1 Systematic Literature Review

A systematic literature review was conducted in July 2025 following the PRISMA methodology, using the Scopus platform and integrated preprint repositories with the query: 'Payment Channel Network' OR 'Lightning Network'. The initial and supplementary searches identified 2,939 records (2,716 articles and 223 preprints). Following relevance screening, 494 records remained; deduplication removed 117 entries, leaving 377 records for further assessment. After final screening, 313 publications were retained, comprising 266 articles and 47 preprints, see [VM25].

For the purpose of this work³, publications proposing PCN performance-related solutions or reporting at least one quantitative performance metric were selected, yielding 134 (42.8%) papers for further analysis. Here, the performance category encompasses works aimed at improving network operational conditions or performance itself, typically using quantitative metrics such as success rate, latency, or communication overhead.

2.2 Metric Mining Procedure

To support reproducibility of the metric-mining process, the analysis should be understood as consisting of three traceable artefacts: (i) the publication corpus and eligibility information,

³ The systematic literature review is fully reported in a separate publication, please, see [VM25]. The current analysis, however, was conducted specifically for this work.

(ii) the extracted raw metric terms, and (iii) the normalized metric taxonomy and category assignments. For each publication, the coding records the reported metric terminology and its assignment to the standardized categories defined in this study. The metrics were categorized into five groups: *Time*, *Computation*, *Communication*, *Payment measurements*, such as payment hop count and payment success, and *Carbon Footprint*. This taxonomy provided a structured basis for assessing implicit energy relevance and identifying gaps in FAIR+S compliance.

The analysis was performed manually using predefined category definitions. To reduce ambiguity, normalization decisions were based on the semantic meaning of the reported metric rather than on terminology alone. For example, “execution time”, “runtime”, and “running time” were mapped to the *Time* category when they represented the duration of computation or routing.

To improve transparency, the publication-level corpus, together with the extracted terms, will be made available alongside the systematic review repository; see [VM25].

2.3 FAIR+S as Methodological Foundation

The FAIR+S or FAIR+Sustainability framework was recently introduced in order to extend FAIR and FAIR4RS by integrating five sustainability dimensions[Va26]:

- *S1 – Energy Efficiency Attributes*: Digital artefacts must include runtime energy measurements or validated estimators.
- *S2 – Sustainability Benchmarks*: Energy metrics should rely on standardized tools and documented assumptions.
- *S3 – Alignment with Sustainability Frameworks*: Standardized carbon intensity reporting ensures comparability across studies.
- *S4 – Carbon Transparency and Accountability*: Reports must disclose uncertainty, geographic energy mix, and measurement tool versions.
- *S5 – Life Cycle Sustainability*: Long-term node operation, maintenance, and data storage must be considered in total footprint calculations.

These combined FAIR+S principles serve here not only as methodological criteria for evaluating research artefacts and extracted performance metrics, but also as a foundation for developing recommendations for the PCN research and developer community.

3 Results

3.1 FAIR in PCN Research and Performance Metrics Landscape

3.1.1 FAIR Principles Adoption and Reproducibility Gaps

In this part of the study, we examined whether PCN research is generally compliant with Open Science and FAIR principles, specifically whether publications transparently report digital artefacts, provide supplementary repositories to ensure reproducibility, and disclose relevant technical and experimental details.

To this end, we reviewed all 134 performance-oriented papers and found that:

- 89 (66.4%) actually proposed algorithmic or topological performance improvements.
- 25 (18.7%) papers disclosed sufficient hardware context (e.g., CPU/GPU specifications, memory, operating system version).
- 14 (10.4%) provided repository links that remain accessible at the time of analysis.
- 5 (3.7%) reported a complete reproducibility context, including both repository access and hardware details.

These findings demonstrate that, even within a relatively new research field, fundamental reproducibility requirements are frequently unmet, and FAIR compliance gaps remain substantial. The results underscore the urgent need to strengthen the adoption and enforcement of FAIR (including FAIR4RS) principles across research and developer communities. They also raise important questions regarding the extent to which publicly funded research adheres to transparency and reproducibility standards. Importantly, without explicit hardware and runtime context, converting execution time or computational metrics into reliable energy consumption estimates is infeasible. Therefore, a critical next step is the systematic reporting of hardware and software environments, particularly in light of growing concerns about computational energy efficiency.

3.1.2 Distribution of Metrics Categories

In this part of the study, we examined which categories of metrics, indicators, or quantitative criteria were used to evaluate the performance of PCN solutions. We were also interested in whether performance-oriented approaches implicitly address energy efficiency or the environmental impact of the proposed solutions.

As described in the Methodology section, we systematically coded and classified all performance metrics reported across the reviewed publications. We found that 134 papers reported at least one quantitative performance metric. Fig. 1 presents a word cloud illustrating the most frequently occurring raw metrics.



Fig. 1: A word cloud illustrating the most frequently occurring performance metrics

Based on this initial set of terms, we conducted a normalization process in which metrics with identical or closely related meanings were consolidated. This procedure resulted in five main categories: *Time*, *Computation*, *Communication*, and *Payment Measurements* (see Tab. 1).

Category	% of mentions	Energy at-tributable?	Examples
Time	33.5	Yes (if hardware disclosed)	execution time, computation time, runtime, transaction time, routing time, payment time, transaction computation time, running time, time consumption, pathfinding time, ...
Computation	7.8	Yes (if standardized)	gas, FLOPs, hash computations, memory used, operations, gas costs, iterations, computation overhead, ...
Communication	10.8	Partially	communication overhead, message count, packet count, signatures broadcasted, communication cost, packet size, ...
Payment Measurements			
~ Payment Hops	10.5	Indirectly	path length, hop count, payment path length, routing steps, ...
~ Payment Success	37.4	No	success rate, payment attempts, successful payment count, successful transaction count, accept rate, success ratio, ...
Carbon Footprint	—	Yes	total carbon emissions, average carbon intensity of electricity

Tab. 1: Normalized frequency distribution of metrics categories

This taxonomy provides a structured basis for assessing implicit energy relevance. As shown

in the Tab. 1), several categories can, in principle, serve as proxies for energy consumption or marginal carbon footprint estimates, provided that the necessary hardware and software context is reported. Some of mentioned metrics can even be measured directly in controlled test-bed environments using appropriate tools. Note that percentages in Tab. 1) represent the proportion of all normalized metric mentions, rather than the proportion of publications.

We also found that only three papers (2.2%) directly address either energy efficiency or sustainability concerns (see [VK24; VK25a; VK25b]). Study [VK25b] employs a direct conversion of estimated energy consumption into total carbon emissions using the CodeCarbon tool[Lo19], while studies [VK24; VK25a] consider the geographical distribution of the network and use the average local carbon intensity of electricity data to derive a relative decrease in the infrastructure’s carbon footprint.

We therefore classify these approaches under a separate category, *Carbon Footprint*. In the next section, we argue that this category should be given high importance when evaluating new performance-oriented solutions.

3.2 Systematizing Performance and Carbon Metrics

3.2.1 From Performance to Energy and to Carbon

Conventional performance metrics can be translated into energy proxies using a simple relationship between power and execution time:

$$E = P \times t, \quad (1)$$

where E denotes energy consumption in watt-hours (Wh), P denotes the power draw in watts (W), and t denotes execution time in hours (h).

Average power draw can be measured following methodologies such as those implemented in CodeCarbon tool[Co25; Lo19]. On Intel processors, this is commonly achieved using the Running Average Power Limit (RAPL) interface, which exposes non-architectural model-specific registers that report power-related information about CPU subsystems. For machines equipped with NVIDIA GPUs, power consumption can additionally be obtained through the NVIDIA System Management Interface (NVIDIA-SMI) which allows users to query real-time GPU power usage. Combining these measurements provides an approximate estimate of the total operational power draw of the computational system during execution.

Carbon emissions or simply footprint can then be estimated as:

$$F = E \times CI, \quad (2)$$

where CI denotes the carbon intensity of the electricity supplying the computation, typically expressed in $\text{kgCO}_2\text{e/kWh}$.

Accurate estimation of the resulting carbon emissions requires information about the geographical location of the computing device, as electricity carbon intensity varies substantially across regions depending on the local energy mix. In practice, this location can be inferred automatically using the user's IP address or explicitly specified by the user. If no geographical information is available, global average carbon intensity values are typically used as a fallback.

This simplified formulation aligns with the methodology proposed by the Green Software Foundation (GSF) in the Software Carbon Intensity specification (SCI)[IS24], which provides a standardized approach to quantifying the carbon intensity of software systems. The SCI metric is defined as:

$$SCI = \frac{E \times CI + M}{R}, \quad (3)$$

where E denotes the operational energy consumption attributable to the software system, CI represents the carbon intensity of the electricity used, M captures embodied carbon emissions associated with the hardware infrastructure allocated to the software, and R represents a functional unit describing the useful output of the system.

Dividing by the functional unit R normalizes emissions per unit of useful work, thereby enabling meaningful comparisons across systems operating at different scales or workloads. In the context of payment channel networks (PCNs), a natural functional unit is the number of processed transactions. However, when comparing routing algorithms or protocol designs, a more granular unit, such as emissions per routing hop or per path step, may provide a even more meaningful basis for comparison.

In practice, estimating operational energy consumption from observable performance metrics requires translating application-level measurements into hardware-level energy use. This translation depends on several system- and context-specific parameters, including processor architecture and power characteristics, memory utilization patterns, network bandwidth and data transfer volume, and data center efficiency indicators such as Power Usage Effectiveness (PUE). In addition, the carbon intensity of the electricity supplying the infrastructure plays a decisive role in determining the resulting emissions.

Such metadata are rarely disclosed in benchmark reports or system descriptions. As a consequence, many analyses rely on coarse energy proxies or generalized hardware power models, which introduces substantial uncertainty into estimates of software-related carbon footprints.

To address this issue, we propose a FAIR+S-aligned performance and carbon metric framework organized into three conceptual layers (see, Fig. 2⁴). The first layer captures core performance metrics traditionally used in PCN research, including latency or runtime, hop count, and payment success rate.

⁴ Except for those provided by CodeCarbon and GSF, all other icons were sourced from Flaticon.com.

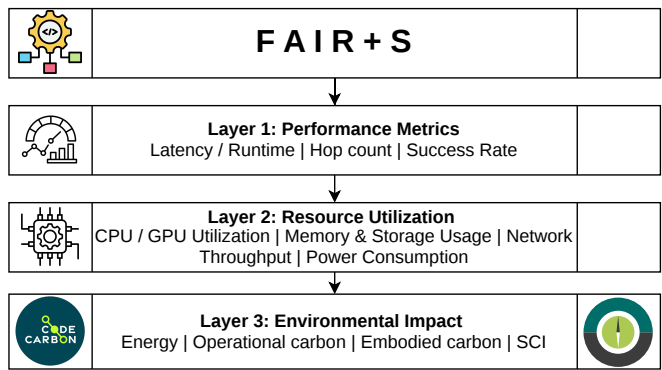


Fig. 2: PCN performance and carbon metrics framework

The second layer focuses on resource attribution and utilization, it describes how computational resources are consumed during execution. This layer includes metrics such as CPU and GPU utilization, memory usage, estimated power consumption, network throughput, communication overhead, and node availability characteristics (e.g., uptime).

The third layer introduces sustainability metrics that quantify the environmental implications of system operation. These include energy consumption measured in kilowatt-hours per transaction or per routing hop, carbon intensity expressed as grams of CO₂ equivalent per transaction or routing step, and life-cycle indicators such as embodied carbon emissions associated with the underlying hardware infrastructure.

3.2.2 Recommendations on Integrating FAIR+S into PCN Solution Evaluation

Our study reveals a structural disconnect between performance optimization and environmental accountability in PCN research. While many routing algorithms and protocol optimizations focus on improving latency or success rates, these improvements may come at the cost of increased computation, communication overhead, or resource utilization, which can ultimately lead to a higher carbon footprint.

We therefore propose that the FAIR+S-aligned performance and carbon metric framework should reframe the evaluation of PCN solutions by integrating sustainability considerations directly into performance assessment. Under such an approach, improvements in algorithmic performance must be evaluated jointly with their corresponding energy intensity and carbon emissions. This integrated perspective can help avoid rebound effects in which efficiency gains trigger increased computational demand, reduce redundant experimentation by improving reproducibility and transparency, support policy-aligned research funding decisions, and contribute to the development of more sustainable blockchain infrastructure.

Operationalizing this framework requires the use of community-recognized measurement tools, benchmarks, and development and reporting standards (e.g., W3C Sustainability Guidelines, Sustainability awareness framework for designing software systems [Du20], GHG protocol etc.). Measurement methodologies should therefore reference established tools and benchmarks such as CodeCarbon [Co25; Lo19], CarbonTracker [AKS21], the Software Carbon Intensity specification [IS24] etc.

At the institutional level, conferences and journals can further support this transition by incorporating sustainability-oriented reporting requirements into their evaluation processes. This includes mandating disclosure of hardware configurations used in experiments, requiring publication of source code and experimental repositories, encouraging reporting of SCI-aligned sustainability metrics, integrating sustainability checklists into the peer-review process, and promoting the use of standardized FAIR+S metadata schemas for experimental artefacts.

For an operational PCN assessment, we recommend distinguishing computational, memory, communication, and infrastructure-related components. Computational resources include CPU and GPU utilization and their associated power consumption. Memory should be considered where its utilization materially contributes to system energy demand. Network communication should be represented through traffic volume, packet or message counts, and, where measurable, the associated energy consumption. Storage should be included when experiments involve substantial persistent data access or storage operations. Finally, infrastructure overhead, such as data-centre PUE, should be incorporated when the experimental environment permits such information to be obtained. The appropriate system boundary should therefore be stated explicitly for every carbon assessment.

As this work is still largely in progress, we can currently provide only a draft checklist, which should be considered preliminary; see <https://github.com/ellariel/ln-fairs-checklist>. An appendix on benchmarking FAIR+S against currently available frameworks and the guidance provided by the GSF is also included.

4 Conclusion

This work systematizes performance and carbon footprint metrics for PCNs using FAIR+S as a methodological foundation. Through systematic literature review and metric mining, we demonstrate that sustainability reporting is largely absent despite extensive performance benchmarking.

We propose a structured, layered framework that integrates conventional PCN performance indicators with energy attribution and carbon estimation techniques aligned with community recognized sustainability standards and tools.

Future work may include the development of validated specifications and automated metadata pipelines for FAIR+S compliance.

References

- [AKS21] Anthony, L.; Kanding, B.; Selvan, R.: CarbonTracker: Tracking and Predicting the Carbon Footprint of Training Deep Learning Models, 2021, arXiv: 2007.03051 [cs.CY].
- [Ba22] Barker, M. et al.: Introducing the FAIR Principles for research software. *Scientific Data* 9 (1), p. 622, 2022.
- [Be15] Becker, C. et al.: Sustainability design and software: the karlskrona manifesto. In: *Proceedings of the 37th International Conference on Software Engineering - Volume 2. ICSE '15*, IEEE Press, Florence, Italy, pp. 467–476, 2015.
- [Co25] Courty, B. et al.: mlco2/codecarbon: v3.0.3, version v3.0.3, 2025.
- [Du20] Duboc, L. et al.: Requirements engineering for sustainability: an awareness framework for designing software systems for a better tomorrow. *Requirements Engineering* 25 (4), pp. 469–492, 2020.
- [IS24] ISO/IEC: Information technology – Software Carbon Intensity (SCI) specification, Retrieved March 11, 2026 from <https://www.iso.org/standard/86612.html>, International Organization for Standardization and International Electrotechnical Commission, 2024.
- [KS23] Khosravi, A.; Säämäki, F.: Beyond Bitcoin: Evaluating Energy Consumption and Environmental Impact across Cryptocurrency Projects. *Energies* 16 (18), 2023.
- [Lo19] Lottick, K. et al.: Energy Usage Reports: Environmental awareness as part of algorithmic accountability, arXiv, 2019.
- [PD16] Poon, J.; Dryja, T.: The Bitcoin Lightning Network: Scalable Off-Chain Instant Payments, Retrieved March 11, 2026 from <https://lightning.network/lightning-network-paper.pdf>, 2016.
- [Va26] Valko, D. et al.: FAIR+S: A validation study of a framework for sustainable research data and software, arXiv, 2026, arXiv: 2606.30663 [cs.CY].
- [VK24] Valko, D.; Kudenko, D.: Reducing CO2 emissions in a peer-to-peer distributed payment network: Does geography matter in the lightning network? *Computer Networks* 243, p. 110297, 2024.
- [VK25a] Valko, D.; Kudenko, D.: Hybrid pathfinding optimization for the lightning network with reinforcement learning. *Engineering Applications of Artificial Intelligence* 146, 2025.
- [VK25b] Valko, D.; Kudenko, D.: Increasing energy efficiency of bitcoin infrastructure with reinforcement learning and one-shot path planning for the lightning network. *Neural Computing and Applications* 37, pp. 18811–18821, 2025.
- [VM25] Valko, D.; Marx Gómez, J.: Recent advances in global payment channel networks: A systematic literature review, *Research Square*, 2025.
- [Wi16] Wilkinson, M. D. et al.: The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* 3 (1), p. 160018, 2016.

Towards a Taxonomy of ICT and Software Sustainability Assessment Instruments

Max Westing ¹ and Stefan Naumann ¹

Abstract: Buyers and regulators of ICT products face instruments that look alike but rest on different methodologies and make substantially different claims. We ask which properties of an instrument determine what it can validly claim, and where deployed instruments fail to realise what the underlying standards support. Following the method of Nickerson et al., we develop in three iterations two coordinated sub-taxonomies, one for the assessment object and one for the sustainability claim, each with five dimensions derived from the ISO 14000 family and from ICT-specific methodologies. Applied to 14 deployed instruments, the taxonomy surfaces a software labelling gap, a structural split between practice-maturity and emission-inventory assessment, a distinction between offset-backed and elimination-based conformance, and the absence of category-specific product category rules for ICT object classes.

Keywords: Sustainability assessment, Software sustainability, ICT environmental labels, Taxonomy development, Software Carbon Intensity

1 Introduction

Buyers, procurers, and regulators of ICT products routinely encounter sustainability instruments (SIs) - e.g. ecolabels, carbon dashboards and other green claims - that are superficially similar but differ in substance. They may all carry a green indicator and a number but rest on different methodologies. Green claims are often misleading and the underlying landscape of frameworks and standards is complex [Mc20], creating a need for a classification that shows what claim can be validly made. Existing reviews collect such instruments but rarely separate *what is being assessed* from *what is being claimed*, so that instruments sharing a label form are treated alike even when they address different objects [Mi20].

Two questions follow: **RQ1:** Which properties of an ICT SI determine which environmental claims it can validly make? **RQ2:** where does the layer of deployed instruments fail to realise claims that the underlying methodologies already support? To answer these questions we propose the development of a taxonomy as a tool to make visible which claims can be validly made about which objects and to navigate the complex field of ICT sustainability assessment as taxonomies are commonly used for this purpose in information systems research [NVM13] and provide an extensible artifact as opposed to capturing a snapshot of this evolving and regulation driven field.

¹ Hochschule Trier, Umwelt-Campus Birkenfeld, Institute for Software Systems, Campusallee 9925, 55768 Hoppstädten-Weiersbach, Germany, m.westing@umwelt-campus.de,  <https://orcid.org/0009-0001-5044-7578>; s.naumann@umwelt-campus.de,  <https://orcid.org/0000-0000-0000-0000>

We use the following terms: A *framework standard* specifies how an assessment is to be carried out, as ISO 14040/14044, ITU-T L.14xx, the Greenhouse Gas (GHG) Protocol ICT Sector Guidance etc. do. An *assessment instrument* is a deployed programme, label, certification, or tool that applies such a methodology and issues a result, and a *claim* is the statement about the assessed object that it makes or enables.

2 Background and Related Work

2.1 Framework standards

The ISO 14000 family, of which ISO 14001 is the most widely adopted member [Ka23], provides the conceptual frame for assessment, reporting, and certification. The ISO 14020 series defines three label types [Mi20]: Type I third-party certified conformance (ISO 14024), Type II self-declared claims (ISO 14021), and Type III quantified ecoprofiles under product category rules (PCR, ISO 14025). ISO 14040/14044 sets out life-cycle assessment principles and the functional-unit construct that enables rate-based accounting, ISO 14064-1 and ISO 14067 quantify emissions at organisational and product level, and ISO 14064-2 covers offset projects. The 2021 amendment to ISO 14021 extended self-declaration to carbon-neutral claims resting either on a zero footprint or on compensating offsets, acknowledging offset-backed substantiation as a distinct path at a time when a substantial literature questions offset integrity [Pr24; SP25].

The ITU-T L.14xx family covers product- and service-level life-cycle assessment (L.1410), organisational GHG and energy accounting (L.1420), and enabling, that is second-order, effects (L.1430), and ETSI ES 203 199 provides a related methodology for ICT goods, networks, and services. The GHG Protocol ICT Sector Guidance adds guidance on how to allocate emissions on shared resources [GH17], while ISO/IEC 21031:2024 specifies the Software Carbon Intensity (SCI) score as a rate of emissions per functional unit of software work, a design that responds to continuously deployed services without a fixed release state [Mo14]. The SDIA Digital Environmental Footprint adds a multi-criteria methodology for digital services that has not moved beyond pilot use.

2.2 Software and sustainability labels

Prior work approaches labelling from two directions. The software-engineering literature established sustainability as a software-quality concern and developed criteria and models for green software [CP15], and ecolabels specific to software have been investigated by Kern et al. [Ke15; Ke18], culminating in the Blue Angel ecolabel for software [NGK20]. The second direction comes from economics and communication studies, which analyse the logic of labelling claims [Bo03] and characterise environmental labels [Mi20].

3 Method

Choice of method and procedure. The taxonomy is developed with the seven-step method of Nickerson et al. [NVM13] which is selected due to its extensibility and the flexibility of offering approaches starting from a conceptual understanding as well as empirical observation. The iterative approach additionally makes it easier to approach a complex field that cannot be digested at a glance. The prior definition of a meta-characteristic provides a clear focus for the taxonomy and the termination criteria offer a mechanism to prevent the taxonomy becoming too fragmented. We retain the original step numbering so that Section 5 can be read against the method: steps 1 and 2 define meta-characteristics and ending conditions; step 3 chooses the approach for the iteration, conceptual-to-empirical (marked *c*) or empirical-to-conceptual (marked *e*); steps 4 to 6 follow that branch, conceiving dimensions (4*c*), examining objects for them (5*c*), and forming the revised taxonomy (6*c*), or selecting objects (4*e*), identifying common characteristics (5*e*), and grouping them into dimensions (6*e*); and step 7 testing the ending conditions.

Two coordinated meta-characteristics. The landscape is governed by four questions: what is assessed, how it is assessed, who asserts and verifies the claim, and for whom it is intended. The properties of the assessed object constrain which claims can validly be made about it and combining it with the assessment method would make it harder to ask which claims an object could support but does not receive, which is necessary to engage with RQ2. The analysis is therefore split into Taxonomy A, “the properties of the assessment object that determine which environmental claims can validly be made about it”, and Taxonomy B, “the nature of the claim and the communicative properties of claim that determine its validity, verifiability, and appropriate use context”.

The nine objective ending conditions of Nickerson et al. are adopted in full, and the five subjective ones are operationalised as targets: *concise* (7 ± 2 dimensions, 3–5 characteristics each), *robust*, *comprehensive*, *extensible*, and *explanatory* in the sense that at least one finding is surfaced.

4 Instrument Set and Coding

The SIs were identified between January and June 2026 based on prior domain knowledge, documentation of hosting and cloud providers and the standards catalogues of ITU-T, ETSI, ISO and GHG Protocol. An instrument was included if it was deployed and publicly accessible at that time, if its assessment object lies within ICT, and if it issues or directly enables an environmental claim about that object. Excluded were labels proposed only in academic work, general product labels without an ICT-specific object, programmes not yet operating, and the framework standards themselves, which prescribe how to assess rather than issue claims themselves. Each instrument was assigned one characteristic per dimension on the basis of its own primary documentation.

The procedure yields 14 instruments in five strata: the *software ecolabel* Blue Angel DE-UZ 215; the *hosting and data-centre programmes* Green Web Foundation (GWF) Hosting Directory, Green Web Check, the EU Code of Conduct for Data Centres [ABP25], and the Climate Neutral Data Centre Pact (CNDCP); the *SaaS and cloud customer-footprint tools* AWS Sustainability Console ², Google Cloud Carbon Footprint ³, Microsoft Emissions Impact Dashboard (EID) ⁴, and Neutrality CO₂Neutral Web Certified (CNWC); the *ICT hardware ecolabels* EPEAT, TCO Certified Gen. 10, and the EU Energy Label for servers; and the *practice-maturity certifications* GreenSeal GSP and Green DiSC ⁵. This set is a sample and not an exhaustive list.

5 Iterations

The notes on the full iterations are included in the supplementary material ⁶

Iteration 1 (conceptual-to-empirical). The goal is an initial dimension set derived from the framework standards. Step 4c produced five candidate dimensions per sub-taxonomy, among them *Locus of environmental impact* and *Lifecycle stance* on the object side. Step 5c classified four instruments (Blue Angel DE-UZ 215, EPEAT, SCI, and the GHG Protocol ICT Sector Guidance) and surfaced three problems: the first of those two dimensions produced multi-valued cells for Blue Angel, which makes an attributional and an enabling-effects claim at once; the second produced no distinguishing cells and overlapped with scope coverage and claim precision; the audience dimension was multi-valued for Blue Angel. In step 6c the first dimension was split into A3 *Scope coverage* and A4 *Enabling effects included*, the second was dropped as its content is carried by A3 and B2, and the audience conflict with the mutually exclusive and collectively exhaustive (MECE) requirement was deferred to Iteration 2, where the full set provides the evidence for a resolution rule.

Iteration 2 (empirical-to-conceptual). The goal is to test the working taxonomy against all 14 instruments and revise where they do not fit (steps 3 and 4e). Step 5e found two patterns that resist the taxonomy. The practice-maturity certifications GreenSeal GSP and Green DiSC assess no emissions of any artefact but certify organisational capability to engineer for sustainability, so forcing them into A1 *Organisational unit*, a characteristic derived for ISO 14064-1-style inventories, misrepresents the assessed object. The offset-backed claim of Neutrality CNWC fits neither B3 *Self-declared with substantiation*, which expects evidence about the product itself, nor *Third-party verified ecoprofile*, because

² <https://aws.amazon.com/sustainability/tools/console/>

³ <https://docs.cloud.google.com/carbon-footprint/docs/methodology>

⁴ <https://learn.microsoft.com/en-us/power-bi/connect-data/azure-emissions-calculation-methodology>

⁵ <https://www.software.ac.uk/GreenDiSC>

⁶ <https://doi.org/10.5281/zenodo.22211713>

its substantiation rests on the integrity of an offsetting scheme. Step 6e therefore adds *Organisational practice/capability maturity* to A1 and *Self-declared with offset substitution* to B3 without altering the dimension structure, and resolves the deferred MECE conflict by a precedence rule for B4: the audience an operator names explicitly takes precedence, and where none is named the form of the instrument decides, so that a public label indicates *Consumer-facing*, a purchasing directory *Procurement/B2B*, an internal dashboard *Engineering/operational*, and so on for the remaining two characteristics. Operators may also exclude an audience, as AWS does when advising against using reported consumption for optimisation⁷.

Iteration 3 (conceptual-to-empirical confirmation). The goal is to verify the derived dimensions. A1 *Organisational practice/capability maturity* is the assessment-object counterpart of ISO 14001, which certifies an environmental management system rather than an emission inventory: an organisation can hold ISO 14001 certification without producing an ISO 14064-1 inventory and vice versa. B3 *Self-declared with offset substitution* is anchored in ISO 14021/Amd 1:2021 on the use side and ISO 14064-2 on the supply side. Step 6c adds, merges, and splits nothing further, so all nine objective ending conditions are satisfied, and the subjective ones with two caveats: A1 and B3 carry six characteristics each, one above the preferred range, and several Taxonomy B characteristics have no instance in the set.

6 The Taxonomy

Table 1 states both sub-taxonomies with their characteristics and framework anchors. Three dimensions carry most of the analytical weight and are discussed here; the remainder follow directly from their anchors.

A2 *Substrate coupling* and A5 *Boundary determinacy* carry most of the analytical weight, and they are where software differs systematically from hardware. A2 records how tightly the object is bound to a physical substrate, from hardware that is its own object to software executable across heterogeneous environments whose claims hold only against a declared reference. A5 records whether the object's boundary is unambiguous, dependent on a declared allocation rule as in multi-tenant workloads, conditional on a reference configuration as in substrate-abstracted software, or open at organisational level. It belongs to the object rather than to the methodology: the methodology dictates how a contested boundary is resolved, whereas the object dictates whether it is contested at all [Mo14].

⁷ <https://docs.aws.amazon.com/sustainability/latest/userguide/energy-calculation.html>

Tab. 1: The two sub-taxonomies, their characteristics, and framework anchors. Codes in parentheses are used in Table 2. Characteristics added in Iteration 2 are marked ⁺.

Dimension	Characteristics	Anchor
<i>Taxonomy A — assessment object</i>		
A1 Object granularity	Component (Cp); Discrete artefact (Da); Service or deployment instance (Sd); Organisational unit (Ou); Organisational practice/capability maturity (Pm) ⁺ ; Sector or portfolio (Sp)	ISO 14040/44, 14064-1, 14067, 14001
A2 Substrate coupling	Substrate-identical (Si); Deployment-bound (Db); Substrate-abstracted (Sa); Substrate-aggregated (Sg)	GHG Protocol ICT
A3 Scope coverage	No GHG accounting (–); Operational, Scope 1+2 (Op); Operational + Embodied, Scope 1+2+3 (OE); nested	GHG Protocol scopes
A4 Enabling effects	None (–); Enabling effects claimed (En)	ITU-T L.1430
A5 Boundary determinacy	Determinate (De); Allocation-dependent (Ad); Conditional (Cd); Open or aggregated (Oa)	ISO 14044
<i>Taxonomy B — sustainability claim</i>		
B1 Claim purpose	Conformance (Cf); Performance (Pe); Disclosure (Ds); Pledge or target (Pl)	ISO 14020/24/25, 14064-1
B2 Claim precision	Qualitative (Ql); Ordinal (Or); Absolute quantified (Aq); Rate-based quantified (Rq); Multi-criteria eco-profile (Mq)	ISO 14025, 14067; ISO/IEC 21031
B3 Verification structure	Self-declared with substantiation (Ss); Self-declared with offset substitution (So) ⁺ ; Third-party periodic (Tp); Third-party verified eco-profile (Tv); Independent assurance (Ia); Unverified (Uv)	ISO 14020 types, 14064-3, 14021/Amd 1
B4 Audience and use context	Consumer-facing (Cn); Procurement/B2B (Pr); Engineering/operational (Eg); Regulatory/disclosure (Rg); Investor/ESG (Iv)	ISO 14020
B5 Comparability scope	Non-comparable (Nc); Within programme (Wp); Within PCR or functional-unit class (Wc); Cross-system (Cs); Absolute reference (Ab)	ISO 14025 PCR; functional unit

7 Classification and Gap Analysis

Table 2 classifies all 14 instruments using the codes of Table 1, separating what the standards support from what instruments actually issue.

Tab. 2: Classification of the 14 instruments. One code per dimension; see Table 1 for the code legend.

Instrument	A1	A2	A3	A4	A5	B1	B2	B3	B4	B5
<i>Software ecolabels</i>										
Blue Angel DE-UZ 215	Da	Sa	–	En	Cd	Cf	Ql	Tp	Cn	Wp
<i>Hosting and data-centre programmes</i>										
GWF Hosting Directory	Ou	Sg	Op	–	Oa	Cf	Ql	Ss	Pr	Wp
Green Web Check	Sd	Sg	Op	–	Oa	Cf	Ql	Ss	Cn	Wp
EU CoC for Data Centres	Ou	Si	Op	–	De	Cf	Ql	Ss	Rg	Wp
CNDCP	Ou	Si	Op	–	De	Pl	Aq	Ia	Rg	Wp
<i>SaaS and cloud customer-footprint tools</i>										
AWS Console	Ou	Sg	OE	–	Ad	Ds	Aq	Ia	Rg	Wc
Google Cloud CF	Ou	Sg	OE	–	Ad	Ds	Aq	Ss	Rg	Wc
Microsoft EID	Ou	Sg	OE	–	Ad	Ds	Aq	Ia	Rg	Wc
Neutrally CNWC	Sd	Db	Op	–	Cd	Cf	Ql	So	Cn	Wp
<i>ICT hardware ecolabels</i>										
EPEAT	Da	Si	OE	–	De	Cf	Mq	Tp	Pr	Wp
TCO Certified Gen. 10	Da	Si	OE	–	De	Cf	Mq	Tp	Pr	Wp
EU Energy Label	Da	Si	Op	–	De	Pe	Or	Tp	Pr	Wp
<i>Practice-maturity certifications</i>										
GreenSeal GSP	Pm	Sg	–	–	Oa	Cf	Or	Tp	Eg	Wp
Green DiSC	Pm	Sg	–	–	Oa	Cf	Or	Tp	Eg	Wp

7.1 Labelling-gap register

Three cells are supported by the framework standards but occupied by no instrument. The first is substrate-abstracted software with rate-based quantified disclosure, supported by SCI, which admits self-declared, third-party-verified, and independent-assurance application: no third-party-verified rate-based software disclosure exists. The second is substrate-abstracted software with a multi-criteria ecoprofile, supported by the SDIA methodology and reachable through ETSI ES 203 199 or ITU-T L.1410. The third is deployment-bound objects with *Operational + Embodied* disclosure for individual SaaS or cloud workloads, supported by ETSI ES 203 199, ITU-T L.1410, and SCI. All three describe the same asymmetry: the apparatus to measure software environmental impact exists and, for SCI, is reasonably mature, while the labelling layer that would convert it into market-visible claims is thin.

8 Discussion

Finding 1: Type I conformance labelling for software is rare and absent from service-instance objects. Blue Angel DE-UZ 215 is the only Type I software label in the set,

and it works because the programme fixes a reference system against which its criteria are measured [NGK20]. There is no Type I label that covers service-instance software although a programme could specify a reference substrate together with an audit-cycle and certify against them. The existing instruments designed for service-instance objects produce quantified disclosure instead.

Finding 2: The software layer shows a gap between framework support and deployed labelling. All three cells of the labelling-gap register are software-side, so the constraint on software sustainability claims lies in the labelling layer rather than in measurement methodology, which suggests that further methodological refinement is not where the next increment of practical value lies.

Finding 3: Practice-maturity assessment is structurally distinct from emission-inventory assessment. GreenSeal GSP and Green DiSC certify organisational capability to engineer for sustainability, while the vendor customer-footprint tools quantify organisational emissions. Both describe an organisation and are marketed as evidence of organisational sustainability, yet the former is anchored in ISO 14001 on management systems and the latter in ISO 14064-1 on inventories, so they are not interchangeable for compliance reporting such as CSRD and a procurement request should state which kind of evidence it requires.

Finding 4: Offset-backed conformance is structurally distinct from emission-elimination conformance. Neutrality CNWC issues an offset-backed neutrality claim and CNDP a pledge that uses offsets as part of its pathway, so their substantiation logic differs both from SCI-style elimination and from Type I conformance although all three read identically as “carbon neutral” in marketing, at a time when offset integrity is seriously questioned [Bj22]. The taxonomy reflects the distinction that ISO 14021/Amd 1:2021 recognises by placing Neutrality at *Self-declared with offset substitution* while CNDP retains *Independent assurance* with offsets recorded as a pledge-pathway annotation. Procurement and regulatory requirements should therefore ask for the substantiation pathway to be declared, not for the neutrality claim alone.

Finding 5: The taxonomy reveals an ICT-specific product category rule gap. ISO 14025 Type III declarations achieve comparability through product category rules that fix scope coverage, life-cycle stages, functional unit, and allocation for one category. The ICT sector has little equivalent infrastructure: ISO/IEC 21031 SCI is a partial analogue for software carbon intensity, the GHG Protocol ICT Sector Guidance is broader, voluntary, and methodology-only. There is no rule set at all for software ecolabels, cloud customer-footprint tools, or practice-maturity certifications. Rules for the principal ICT object classes could be a path towards establishing a way to compare the sustainability of software systems.

9 Limitations

Software-specific ecolabels have a narrow empirical base, with exactly one Type I software label and none for service-instance software, so evidence Findings 1 and 2 is weak. Coding was carried out by the author and no second-coder protocol has been executed, so inter-coder robustness remains unestablished. The set is not exhaustive, excluding emerging instruments such as EPEAT Renew, NADIKI Registrar⁸, and those implied by the first EU technical report on data centre energy performance [Di25]. Finally, the taxonomy preserves the MECE requirement on B4 and A3 only through workarounds: B4 relies on a procedural rule that resolves multi-audience instruments at coding time and thereby loses secondary audiences, and A3 uses a nested cumulative scale that cannot represent an instrument covering Scope 3 without Scope 2.

10 Conclusion and Future Work

This paper developed a taxonomy of ICT and software sustainability assessment instruments that separates the assessment object from the sustainability claim. Applied to 14 deployed instruments, RQ1 is engaged by defining ten dimensions that govern claim validity and RQ2 through a register of claims the framework supports but no instrument issues. The central observation is that the ICT sector lacks the category-specific infrastructure that physical-product sustainability has, a gap visible both in the labelling-gap register and in Finding 1, where the only Type I software label solves the problem by fixing a reference substrate at programme level. Future work should include: execution of the second-coder protocol, validation with practitioners in the Green IT community, and engagement with the feasibility of defining product category rules for software.

Funding. This work was funded by the Internet Society Foundation as part of the project “Unlocking the Potentials of Green Coding”. The authors are not aware of any conflict of interest.

References

- [ABP25] Acton, M.; Booth, J.; Paci, D.: 2025 Best Practice Guidelines for the EU Code of Conduct on Data Centre Energy Efficiency, JRC Publications Repository, 2025, <https://publications.jrc.ec.europa.eu/repository/handle/JRC141521>, accessed: 06/16/2026.
- [Bj22] Bjørn, A. et al.: Renewable Energy Certificates Threaten the Integrity of Corporate Science-Based Targets. *Nature Climate Change* 12 (6), pp. 539–546, 2022, <https://www.nature.com/articles/s41558-022-01379-5>, accessed: 04/27/2026.

⁸ <https://ided.digital/en/latest-news/publications/nadiki-registration-implementation-open-source>

- [Bo03] Boer, J. de: Sustainability Labelling Schemes: The Logic of Their Claims and Their Functions for Stakeholders. *Business Strategy and the Environment* 12 (4), pp. 254–264, 2003, <https://onlinelibrary.wiley.com/doi/10.1002/bse.362>, accessed: 05/11/2026.
- [CP15] Calero, C.; Piattini, M.: *Green in Software Engineering*. Springer International Publishing, 2015.
- [Di25] Directorate-General for Energy (European Commission) et al.: *Assessment of the Energy Performance and Sustainability of Data Centres in EU: First Technical Report*. Publications Office of the European Union, 2025.
- [GH17] GHG Protocol and GeSI and Carbon Trust: *GHG Protocol Product Life Cycle Accounting and Reporting Standard – ICT Sector Guidance*, Greenhouse Gas Protocol, 2017, [https://ghgprotocol.org/sites/default/files/2023-03/GHGP-ICTSG % 20 - %20ALL %20Chapters.pdf](https://ghgprotocol.org/sites/default/files/2023-03/GHGP-ICTSG%20-%20ALL%20Chapters.pdf), accessed: 07/20/2026.
- [Ka23] Karaeva, A. et al.: Family ISO 14000 Standards as a Tool of Achieving Environmental Sustainability of Enterprises. *IOP Conference Series: Earth and Environmental Science* 1126, p. 012036, 2023.
- [Ke15] Kern, E. et al.: Labelling Sustainable Software Products and Websites: Ideas, Approaches, and Challenges. In: *Advances in Computer Science Research*. Atlantis Press, 2015, <http://dx.doi.org/10.2991/ICT4S-ENV-15.2015.10>.
- [Ke18] Kern, E. et al.: Sustainable Software Products—Towards Assessment Criteria for Resource and Energy Efficiency. *Future Generation Computer Systems* 86, pp. 199–210, 2018, <http://dx.doi.org/10.1016/j.future.2018.02.044>.
- [Mc20] McGuinn, J. et al.: *Environmental Claims in the EU: Inventory and Reliability Assessment*, Prepared by Milieu Consulting SRL and IPSOS NV for the European Commission, Directorate-General for Environment, under contract 070201/2019/815369/SER/ENV.B.1, Luxembourg: Publications Office of the European Union, 2020, <https://op.europa.eu/en/publication-detail/-/publication/f7c4cb8b-f877-11ee-a251-01aa75ed71a1/language-en>, accessed: 08/31/2026.
- [Mi20] Minkov, N. et al.: Characterization of Environmental Labels beyond the Criteria of ISO 14020 Series. *The International Journal of Life Cycle Assessment* 25 (5), pp. 840–855, 2020, <https://doi.org/10.1007/s11367-019-01596-9>, accessed: 06/19/2026.
- [Mo14] Moghaddam, R. F. et al.: Challenges and Complexities in Application of LCA Approaches in the Case of ICT for a Sustainable Future. In: *ICT for Sustainability 2014 (ICT4S-14)*. Atlantis Press, pp. 155–164, 2014, <https://www.atlantis-press.com/proceedings/ict4s-14/13437>, accessed: 06/22/2026.
- [NGK20] Naumann, S.; Guldner, A.; Kern, E.: The Eco-label Blue Angel for Software—Development and Components. In: *Progress in IS*. Springer International Publishing, pp. 79–89, 2020, http://dx.doi.org/10.1007/978-3-030-61969-5_6.
- [NVM13] Nickerson, R. C.; Varshney, U.; Muntermann, J.: A Method for Taxonomy Development and Its Application in Information Systems. *European Journal of Information Systems* 22 (3), pp. 336–359, 2013, <https://doi.org/10.1057/ejis.2012.26>, accessed: 06/16/2026.
- [Pr24] Probst, B. S. et al.: Systematic Assessment of the Achieved Emission Reductions of Carbon Crediting Projects. *Nature Communications* 15 (1), p. 9562, 2024, <https://www.nature.com/articles/s41467-024-53645-z>, accessed: 04/27/2026.
- [SP25] Stolz, N.; Probst, B. S.: The Negligible Role of Carbon Offsetting in Corporate Climate Strategies. *Nature Communications* 16 (1), p. 7963, 2025, <https://www.nature.com/articles/s41467-025-62970-w>, accessed: 04/27/2026.

List of Authors

Bartoszczuk, Paweł, 47

Behrens, Grit, 89, 117

Bruck, Paul, 89

Burkard, Simon, 25

Deharde, Maximilian, 25

Elfeky, Ahmed, 101

Engstler, Martin, 57

Fellinger, Engelbert, 79

Flohr, Sarah, 89

Fuchs-Kittowski, Frank, 25

Garcia, Nick Kriesing, 141

Grzeszczyk, Rafał, 117

Gómez, Jorge Marx, 155

Hammann, Leonie, 67

Hoffmann, Geerd-Dietger, 131

Hosenfeld, Friedrich, 15

Jahns, Maximilian, 131

Junger, Dennis, 131, 141

Kahlfeld, Andreas, 89

Kleinmann, Jan, 89

Kobiela, Paweł, 47

Koravi, Sandhya, 35

Krüger, Kay, 15

Naumann, Stefan, 165

Orłowski, Cezary, 117

Poppel, Marius, 25

Souza, Larissa Koch de, 57

Starzl, David, 25

Stańczak, Jarosław, 47

Szałata, Łukasz, 117

Thimm, Heiko, 67

Valko, Danila, 155

Werner, Nicole, 67

Westing, Max, 165

Wohlgemuth, Volker, 131, 141