Volker Wohlgemuth, Stefan Naumann, Grit Behrens,
Anna Zagorski, Maximilian Höb
(Editors)

EnviroInfo 2025
Short-/Work in Progress-Papers

September 17-19, 2025
Potsdam - Germany

Gesellschaft für Informatik e.V. (GI)

Lecture Notes in Informatics (LNI) - Proceedings
Series of the Gesellschaft für Informatik (GI)

Volume P-372

ISSN 2944-7682

Volume Editors

Volker Wohlgemuth	HTW Berlin, University of Applied Sciences, Berlin, Germany
Prof. Stefan Naumann	Umwelt-Campus Birkenfeld – Trier University of Applied Sciences, Germany
Prof. Grit Behrens	University of Applied Sciences Bielefeld, Germany
Anna Zagorski	German Environment Agency (Umweltbundesamt), Germany
Maximilian Höb	Leibniz Supercomputing Centre (LRZ), Bavarian Academy of Sciences, Germany

Preface

This book contains the short and work-in-progress contributions of the 39th International and Interdisciplinary Conference on Environmental Informatics – EnviroInfo 2025.

The conference took place from 17 to 19 September 2025 at the University of Potsdam, Campus III – Griebnitzsee and is organised by technical committee of Environmental Informativs of the Gesellschaft für Informatik e.V. (GI). The University of Potsdam, together with the Hasso Plattner Institute (HPI), acts as an important host partner.

For almost four decades, the EnviroInfo series has been bringing together national and international players who are involved in applied computer science and sustainability – with the aim of making concrete contributions to a sustainable society through information and communication technologies.
The motto of EnviroInfo 2025 "Open Data, Open Science and Open Society for a Sustainable Future" illustrates the claim to understand openness as a central component of sustainable development. Open data creates transparency and forms the basis for informed decisions, open science promotes collaboration and accelerates innovation, and an open society strengthens participation and ensures that digital solutions are firmly anchored in society.

With this volume, we would like to present contributions that are still in development or show initial results, as well as innovative approaches that combine interdisciplinarity and practical relevance. They take up the motto of the conference by showing how openness in data, science and society drives the further development of methods and technologies in environmental informatics. The variety ranges from green coding and sustainable digital systems to new forms of data infrastructure and practical applications that connect science, politics and society.

Our special thanks go to all authors for their valuable contributions and to the members of the programme and organisation committee for their careful review. We would also like to thank our partners, especially the University of Potsdam and the Hasso Plattner Institute (HPI) as host institutions, the Gesellschaft für Informatik e.V., as well as our sponsors Disy Informationssysteme GmbH and Hewlett Packard Enterprise, who contributed significantly to the success of the conference with their commitment.

Potsdam, September 2025

On behalf of the organization Committee
Volker Wohlgemuth, Conference President

Sponsors

We would like to thank the following companies for supporting the conference.

Disy Informationssysteme GmbH

Hewlett Packard Enterprise

German Informatics Society

General Chair

Volker Wohlgemuth	HTW Berlin University of Applied Sciences, Berlin, Germany
Stefan Naumann	Umwelt-Campus Birkenfeld – Trier University of Applied Sciences, Germany
Grit Behrens	University of Applied Sciences Bielefeld, Germany
Anna Zagorski	German Environment Agency (Umweltbundesamt), Germany
Maximilian Höb	Leibniz Supercomputing Centre (LRZ) der Bayrischen Akademie der Wissenschaften, Germany

Programme Committee

Achim Guldner, Trier University of Applied Sciences, Germany
Amna Ramzy, German University in Cairo (GUC), Egypt
Andreas Abecker, Disy Informationssysteme GmbH, Germany
Andreas Kamilaris, CYENS Center of Excellence, Cyprus
Andreas Möller, Leuphana University Lüneburg, Germany
Andreas Winter, University of Oldenburg, 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
Dieter Kranzlmüller, Leibniz Supercomputing Centre (LRZ), Germany
Frank Fuchs-Kittowski, HTW Berlin University of Applied Sciences, Berlin, Germany
Friedhelm Hosenfeld, DigSyLand, Germany
Hans-Knud Arndt, Otto-von-Guericke University Magdeburg, Germany
Heiko Thimm, Hochschule Pforzheim, Germany
Horst Kremers, UNDRR SEM - RIMMA CoE Community, Germany
Ioannis Athanasiadis, Wageningen University & Research, The Netherlands
Jochen Wittmann, HTW Berlin University of Applied Sciences, Berlin, Germany Jorge Marx Gomez, Carl von Ossietzky University of Oldenburg, Germany
Klaus Greve, University of Bonn, Germany
Kostas Karatzas, Aristotle University of Thessaloniki, Greece
Kristina Voigt, Formerly Helmholtz Zentrum München, Germany
Margaret M. MacDonell, Argonne National Laboratory, USA
Maggie Mashaly, German University in Cairo (GUC), Egypt
Martina Willenbacher, HTW Berlin University of Applied Sciences, Berlin, Germany
Malgorzata Burchard-Dziubinska, University of Lodz, Poland
Mikko Kolehmainen, University of Eastern Finland (UEF), Finland
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, Germany
Stefan Naumann, Trier University of Applied Sciences, Germany
Sven Schade, European Commission - Joint Research Centre, Italy
Thomas Rose, Fraunhofer Institute for Applied Information Technology (FIT), Germany

Organising Committee

Anna Zagorski	German Environment Agency (Umweltbundesamt), Germany
Dennis Junger	HTW Berlin University of Applied Sciences, Berlin, Germany
Grit Behrens	University of Applied Sciences Bielefeld, Germany
Max Westing	Umwelt-Campus Birkenfeld – Trier University of Applied Sciences, Germany
Maximilian Höb	Leibniz Supercomputing Centre (LRZ), Bavarian Academy of Sciences, Germany
Stefan Naumann	Umwelt-Campus Birkenfeld – Trier University of Applied Sciences, Germany
Volker Wohlgemuth	HTW Berlin University of Applied Sciences, Berlin, Germany

Contents

Work-in-Progress Papers / Short Papers

Artificial Intelligence and Sustainability

Work-in-Progress Papers / Short Papers

Artificial Intelligence and Sustainability

SkyFusion: A Physics-Informed GAN-Transformer Framework for Intra-Hour Sky Forecasting and PV Prediction

Naoufal El Atmioui[1], Grit Behrens[2], Andreas Boschert[3], Michael Zehner[4]

Abstract: Reliable integration of renewable energy sources into electrical grids hinges on accurate forecasting of sky conditions and photovoltaic (PV) power generation [Nie23]. However, existing methods, whether traditional numerical weather prediction or current deep learning models, have faced challenges in reliably generating realistic intra-hour sky conditions (e.g., at 15-minute horizons) and associated PV power forecasts, primarily due to the complex and stochastic dynamics of cloud motion [Nie23]. In this work, we address this gap, demonstrating successful probabilistic sky forecasting and associated PV predictions at a 15-minute horizon. To address prior limitations, we introduce **SkyFusion**, a novel, modular pipeline that synergistically combines physics-informed latent-space modeling (PhyCell) [LG20], transformer-based autoregression, and a multi-stage generative adversarial network (GAN) approach. Although GAN components faced initial stability challenges, clear strategies for future integration are outlined, paving the way towards high-fidelity, physically plausible intra-hour sky sequences and robust probabilistic PV power prediction.

Keywords: Solar Forecasting; PV Power; Cloud Motion; Physics-Informed NN; Generative Adversarial Networks; Transformer; ConvLSTM; VQ-VAE.

1 Introduction

The integration of solar energy into modern power grids requires accurate, short-term forecasting of photovoltaic (PV) output, especially at minute-level resolutions [Nie23, LG20]. Rapid cloud dynamics disrupt grid stability and challenge both energy trading and load balancing. Traditional NWP models lack temporal granularity, while purely data-driven methods offer limited physical interpretability and often yield deterministic outputs [Nie23, LG20]. SkyFusion addresses this by combining physics-informed modeling with deep generative learning to enable robust, probabilistic 15-minute forecasts. Chapter 5 details the ConvLSTM-UNet baseline, which forms the foundation of the forecasting pipeline.

[1]University of Applied Sciences Bielefeld, Solar Computing Lab, Artilleriestraße 9, 32427 Minden, naoufal.el_atmioui@hsbi.de
[2]University of Applied Sciences Bielefeld, Solar Computing Lab, Artilleriestraße 9, 32427 Minden, grit.behrens@hsbi.de, https://orcid.org/0009-0009-0247-8204
[3] Technical University of Applied Sciences Rosenheim, Laboratory for Solar Energy Systems, Hochschulstraße 1, 83024 Rosenheim, andreas.boschert@th-rosenheim.de, https://orcid.org/0000-0002-1540-8098
[4] Technical University of Applied Sciences Rosenheim, Laboratory for Solar Energy Systems, Hochschulstraße 1, 83024 Rosenheim, michael.zehner@th-rosenheim.de, https://orcid.org/0000-0001-9642-0496

doi:10.18420/env2025_01

2 Objectives and Scope

This paper focuses on **Milestone 1**, which covers data preparation, sun-position masking (via PVLib), and the implementation of a ConvLSTM-UNet baseline that forecasts the next 15 minutes at 30-second intervals. GAN components—specifically PatchGAN—were successfully integrated into the baseline to enhance the visual quality of predictions. Their positive effect in reducing blurriness is demonstrated through side-by-side comparisons. A roadmap for future GAN extensions (e.g., TemporalGAN) is outlined to further improve realism and temporal coherence.

The structure of the paper has been revised for clarity and scientific rigor. Section 3 reviews related work; Section 4 outlines the SkyFusion pipeline; Section 5 details the ConvLSTM-UNet implementation and validation; Section 6 presents benchmark results; Section 7 discusses future work, including GAN stabilization and scalability; Section 8 concludes.

3 Relevant Research

Recent advances in sky image forecasting and PV power prediction have leveraged deep learning to address the challenges of spatiotemporal dynamics and uncertainty.

SkyGPT [Nie23] introduced a two-stage pipeline using a VQ-VAE and a physics-constrained transformer (Phy-Transformer) for probabilistic sky video generation, followed by PV power nowcasting. It employs L2 and moment losses but lacks adversarial or perceptual objectives, resulting in blurred textures. SkyFusion builds on this by integrating GAN-based perceptual training (planned in future stages) to enhance visual sharpness and extend forecasting beyond 15 minutes.

PhyDNet [LG20] proposed PhyCell, a physics-informed recurrent unit that disentangles motion from appearance, improving the physical consistency of video predictions. However, it lacks adversarial refinement, often yielding blurred outputs. SkyFusion adopts similar physics-based constraints while planning GAN integration to address visual fidelity.

ConvLSTM [SH15] remains a standard for spatiotemporal video prediction but suffers from blurriness and temporal degradation in long-range forecasting. SkyFusion uses a ConvLSTM-UNet baseline enhanced with multi-scale SSIM loss to address these limitations.

GANs [IS17] improve image realism in video forecasting by enforcing local (PatchGAN) and temporal (TemporalGAN) coherence. While unstable during training, their integration with structured models can significantly enhance visual quality. SkyFusion's GAN module shows a finite resolution in new predicted pictures (see Section 6).

4 Project Overview

Accurate intra-hour PV power forecasting is vital for grid stability under high renewable penetration. **SkyFusion** addresses this by combining four modular deep learning stages that forecast sky conditions 15 minutes ahead and convert them into probabilistic PV power estimates.

4.1 Stage1 ConvLSTM-UNet Baseline Generation (SkyBuilder)

Captures short-term cloud motion and forecasts 30 frames (15 minutes at 30s intervals) using L1 and SSIM losses. Inputs are downsampled to 128×128 for efficiency. PatchGAN is applied for texture enhancement.

4.2 Stage2 Latent Tokenisation with VQ-VAE (SkyCompressor)

Compresses sky video clips into discrete 16×16×T latent tokens using a shared codebook ($K \approx 1024$).

4.3 Stage3 Transformer Forecasting with Physics Constraints (SkyForecaster)

Uses a GPT-style transformer with PhyCell to forecast latent tokens autoregressively. PatchGAN and optional TemporalGAN aim to improve spatial and temporal coherence.

4.4 Stage4 PV NowCasting and Energy Forecasting (SkyNowcaster)

Maps the predicted sky frame at t+15 to a probabilistic PV power output using a CNN regressor trained with RMSE and CRPS losses. This supports uncertainty-aware grid decision-making.

Each module contributes toward generating physically consistent, visually realistic sky videos and linking them to operational PV forecasting.

5 Milestone 1: Data Preparation & Baseline

This section details the **first milestone** of the SkyFusion project, focused on preparing a high-quality, physics-consistent dataset and implementing a functional **ConvLSTM-UNet baseline** for intra-hour sky video forecasting. This foundational step establishes a reproducible benchmark to evaluate our data pipeline, modeling infrastructure, and early forecasting capabilities.

5.1 Dataset Cleaning and Image Selection

The first milestone of SkyFusion focuses on building a high-quality, physics-consistent dataset and implementing a ConvLSTM-UNet baseline for intra-hour sky video forecasting. Starting with over 400,000 raw images from two ASI cameras (asi_16030 and asi_16515) at a German PV test field [AI24], we applied a rigorous preprocessing pipeline. This included removal of corrupted frames and calibration artifacts, while preserving minor imperfections like droplets to retain real-world variability. The original 15-second interval was downsampled to 30 seconds to reduce redundancy and improve temporal coherence. The resulting dataset of approximately 178,000 images serves as a robust foundation for model training and reproducible benchmarking.

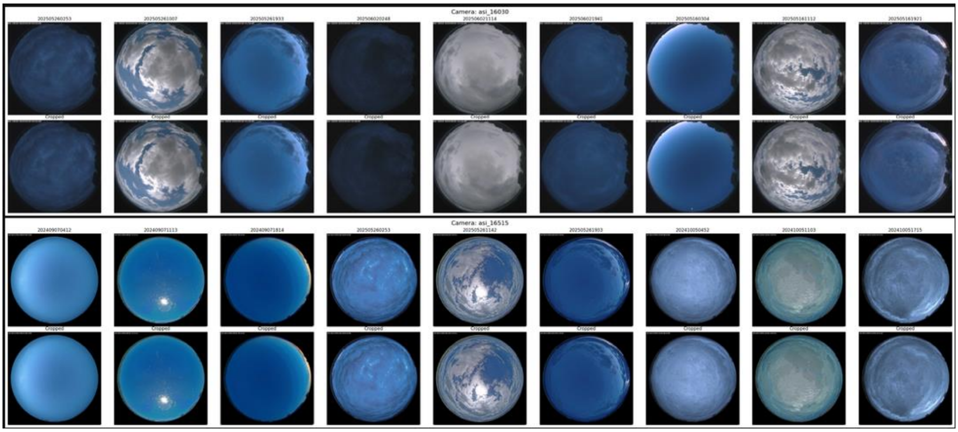


Fig. 1: Sample raw sky images from the asi_16030 (top rows) and asi_16515 (bottom rows) datasets, illustrating typical cloud cover and lighting conditions. The bottom image in each pair shows the cropped version used for processing.

5.2 Data Augmentation

To enhance generalization under real-world variability, we applied physics-plausible augmentations, including horizontal flips (p=0.3) to promote symmetry-invariant motion learning, Gaussian blur to simulate sensor imperfections, and ColorJitter to mimic natural sky brightness and color variations. Crucially, MixUp and CutMix were excluded, as they disrupt temporal and physical continuity key elements in structured video forecasting tasks like cloud evolution.

Original	Gaussian

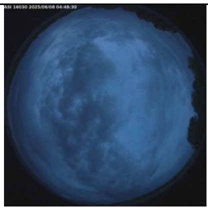
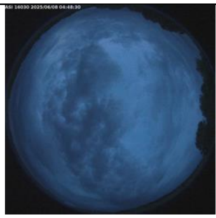
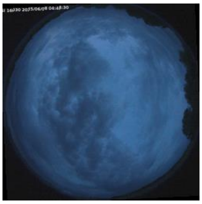
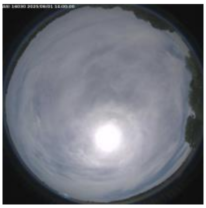
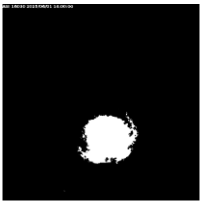
	
Small Rotation 	Color Jitter 

Fig. 2: Examples of physics-plausible augmentations applied to a sample sky image.

5.3 Sun Position Estimation and Masking

We used PVLib [Hol18] to compute solar azimuth and elevation per timestamp and location. A binary sun mask was generated as a white disk placed at the sun's location over a black background. This mask was added as a fourth input channel to the RGB image, forming a (4, 128, 128) tensor. It helps the model track sun position under partial cloud cover and improves physical consistency.

Original 	Sun Mask 	Original + Sun Mask 
Original	Sun Mask	Original + Sun Mask

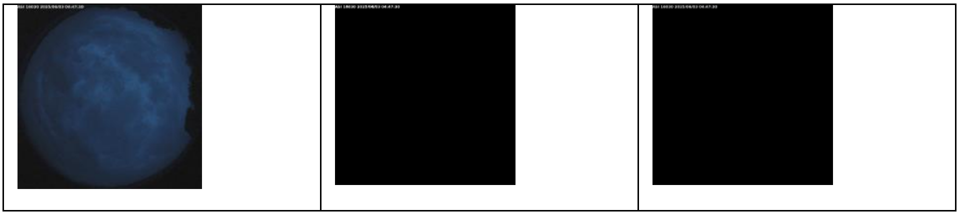


Fig. 3: Sun masking examples for clear (top) and overcast (bottom) conditions, showing original images, binary masks, and masked outputs.

5.4 Tensor Structure and Input Format

Each preprocessed sky image, including the sun mask, was resized to 128×128, forming a (Batch, 4, 128, 128) tensor combining RGB and sun position. Input sequences consisted of 12 frames at 30s intervals (6-minute lookback), while target sequences spanned the next 30 frames (15-minute forecast).

5.5 ConvLSTM-UNet Model Implementation

The ConvLSTM-UNet baseline combines ConvLSTM layers with a U-Net encoder-decoder architecture to model spatiotemporal dynamics and reconstruct future sky frames. The encoder downsamples inputs and captures temporal patterns via ConvLSTM units, while the decoder upsamples features with skip connections to restore fine details. Implemented in PyTorch Lightning, the model predicts 30 future frames from 12 input frames, ensuring modularity, GPU acceleration, and reproducibility.

5.6 Loss Functions and Metrics

To effectively guide the ConvLSTM-UNet model's training and evaluate its performance, a composite loss function and a set of standard image quality metrics were employed:

- L1 Loss: Mean Absolute Error (MAE) was used as a primary pixel-wise loss, encouraging accurate intensity prediction across the image.
- SSIM Loss: Structural Similarity Index Measure (SSIM) was integrated into the composite loss. This perceptual metric helps to prioritize structural fidelity and improve the visually perceived quality of the predicted frames, mitigating the common issue of blurry predictions often seen with purely pixel-wise losses.

5.7 Training Setup

The ConvLSTM-UNet baseline was trained using the Adam optimizer (lr = 1e-4) with a batch size of 32 for 100 epochs. Training was performed on an NVIDIA A100 (80GB) via a SLURM-managed cluster. Key metrics (loss, SSIM, PSNR) and predicted video samples were logged using Weights & Biases for detailed monitoring and visualization.

5.8 Optional GAN Refinement

An experimental **PatchGAN discriminator** was fully implemented and integrated with the ConvLSTM-UNet architecture to explore its potential in enhancing local texture sharpness and perceptual realism. We conducted ablation experiments **with and without PatchGAN**, confirming that the adversarially trained version produced noticeably sharper and more detailed predictions—a qualitative improvement, as illustrated in Section 6.

5.9 Observations: What Worked / What Didn't

The first milestone yielded key insights. PVLib-based sun masking aligned well with solar geometry. The ConvLSTM-UNet model was successfully integrated in PyTorch Lightning, and the pipeline trained reliably on the SLURM cluster. Using 30s intervals between frames helped capture cloud motion smoothly. However, limited data for rare cases (e.g., fog, sunrise, low light) may affect generalization in future stages.

6 Preliminary Results

The ConvLSTM-UNet model was trained for 100 epochs on sequences with a 15-minute prediction horizon (12 input and 30 target frames at 30s intervals). As shown in **Fig. 4**, the model without PatchGAN captures broad cloud dynamics but suffers from significant blur and lack of detail. In contrast, **Figs. 5 and 6** show predictions with PatchGAN, where visual sharpness noticeably improves, demonstrating enhanced texture and edge clarity while preserving motion dynamics. These results confirm that adversarial training reduces blurriness and improves realism.

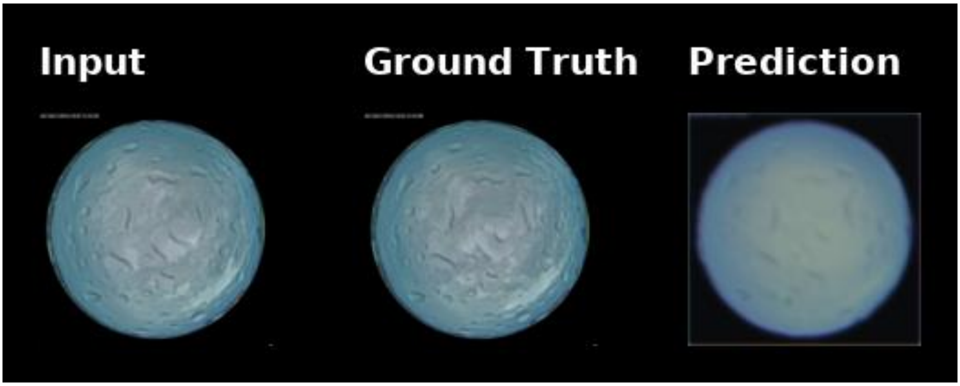


Fig. 4: Visual comparison of input, ground truth, and ConvLSTM-UNet predictions (no PatchGAN).

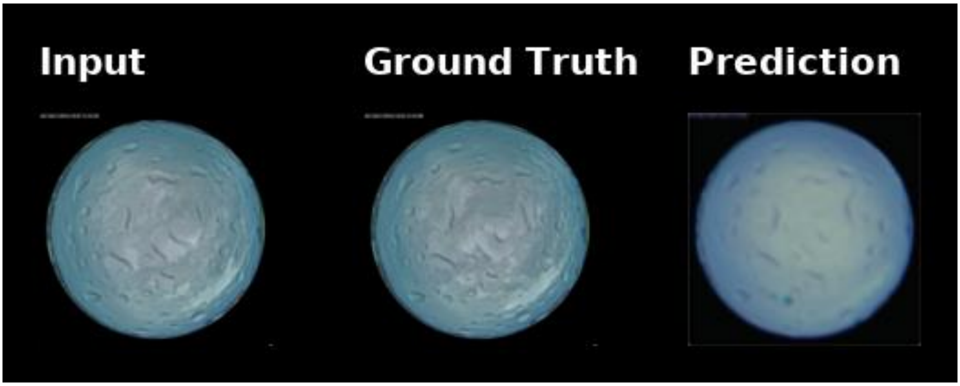


Fig. 5: Visual comparison of input, ground truth, and predicted frames.

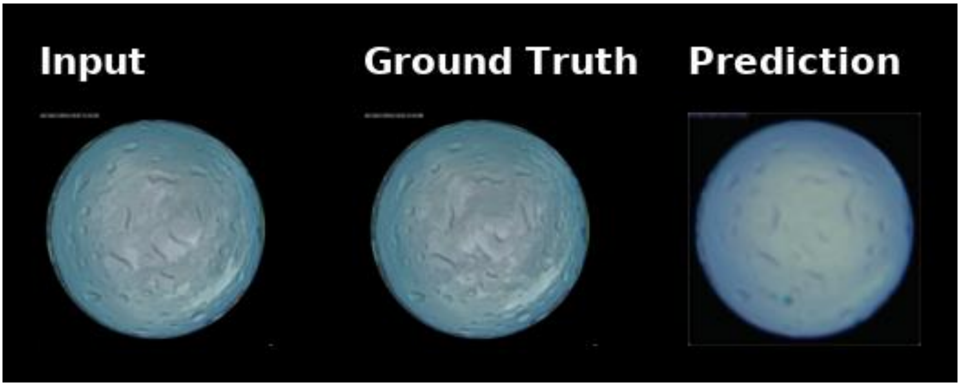


Fig. 6: Temporal evolution of predicted sky frames vs. ground truth.

Future work will expand evaluations under varied sky conditions and include per-frame error maps for detailed accuracy assessment.

7 Planned Next Steps

With the successful implementation of the ConvLSTM-UNet baseline in Milestone 1, the next steps for SkyFusion include expanding temporal forecasting, improving visual fidelity, and integrating PV power prediction. The upcoming **SkyCompressor** module will use a 3D VQ-VAE to compress sky videos into latent tokens, enabling longer forecasting while reducing computational load. The **SkyForecaster** will extend predictions with a physics-informed transformer combining GPT-style decoding and PhyCell, enhanced by PatchGAN and TemporalGAN for realism. The **SkyNowcaster** will map generated sky images into probabilistic PV estimates using CNN regressors trained with RMSE and CRPS losses. We will benchmark SkyFusion against SkyGPT, PhyDNet, and NWP baselines using both quantitative metrics (SSIM, MAE, CRPS) and visual outputs, and assess generalization across seasons, locations, and PV surface conditions.

8 Conclusion and Impact

SkyFusion offers a modular, physics-informed pipeline for intra-hour sky and PV forecasting, extending prior work like SkyGPT with enhanced interpretability, GAN-based realism, and transformer-driven latent forecasting. This milestone delivers a reproducible ConvLSTM-UNet baseline achieving 15-minute predictions, supported by a robust data pipeline and efficient training infrastructure. While GAN integration proved valuable for enhancing visual quality, SkyFusion ultimately aims to support uncertainty-aware decision-making for grid operators, bridging visual forecasting with operational PV power needs in a scalable and explainable framework.

Bibliography

[AI24] A. Boschert, R. Kessler, G. Behrens, A. Zeyher, and C. Hofmann, "Helios: Introducing a project for forecasting photovoltaic power generation using artificial intelligence," in *Proc. 37th Int. Conf. Informatics for Environmental Protection (EnviroInfo)*, Sep. 2024. [Online]. Available: https://dblp.org/rec/conf/enviroinfo/BoschertKBZH24.html

[IS17] P. Isola, J. Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2017, pp. 1125–1134.

[LG20] V. Le Guen and N. Thome, "Disentangling physical dynamics from unknown factors for unsupervised video prediction," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2020, pp. 1786–1795.

[Nie23] Y. Nie *et al.*, "SkyGPT: Probabilistic short-term solar forecasting using synthetic sky videos from physics-constrained VideoGPT," *arXiv preprint*, arXiv:2306.11682, Jun. 2023. [Online]. Available: https://arxiv.org/abs/2306.11682

[SH15] X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-C. Woo, "Convolutional LSTM network: A machine learning approach for precipitation nowcasting," in *Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 28, 2015, pp. 802–810.

[Ron15] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. (MICCAI)*, 2015, pp. 234–241.

[Hol18] W. Holmgren, C. Hansen, and D. P. Eldridge, "pvlib python: A Python package for modeling solar energy systems," *J. Open Source Softw.*, vol. 3, no. 29, p. 884, 2018. doi: 10.21105/joss.00884

Rotated Bitboards and Reinforcement Learning in Computer Chess and Beyond - Towards Absolute Sustainability and Progressive Computer Science

Johannes Buchner [1]

Abstract: We discuss the potential of using methods from AI (in a broad definition) to achieve absolute sustainability, with a focus on Reinforcement Learning (RL). To give some context, we explain how RL was successfully applied in computer chess and beyond (e.g. in AlphaZero / AlphaFold), and also describe a method called 'rotated bitboards' with potential applications in the context of sustainability. We discuss the computer-science related challenges of changing the economic paradigm towards (absolute) sustainability, and in how far what we call 'progressive computer science' needs to contribute. Those challenges include the closing of material loops in a circular economy in order to optimize for (absolute) sustainability, and we present some new ideas in this direction. Finally we discuss the potential of such methods for ecological economic planning by changing the current mostly 'linear' and profit-driven way of organizing the economy towards a more circular, democratic and sustainable mode of production.

Keywords: Reinforcement Learning, Circular Economy, Life Cycle Assessment, Absolute Sustainability, Artificial Intelligence, Computer Chess, Rotated Bitboards, Progressive Computer Science

1 Introduction

The aim of this paper is to describe some methods that were successful in computer chess such as (rotated) bitboards ([Ch25c]) and Reinforcement Learning ([Ch25b]), and to discuss their potential for solving problems in other domains. We note that since the introduction of AlphaZero ([Si17]) which was able to achieve superhuman performance in chess, trained only by reinforcement learning from games of self-play, similar approaches lead to spectacular success in other domains, e.g. the "other Alphas" (all created by DeepMind) like AlphaFold [De21], AlphaTensor [Fa22] and AlphaProof [De24]. In this spirit, we propose some new ideas how to tackle the problem of closing of material loops in a circular economy in order to optimize for sustainability, which is part of a research program that we call "Progressive Computer Science" in order to respect planetary boundaries (see [Ro09]) and avoid climate tipping points [Ar22]. Finally we elaborate on the potential of artificial intelligence methods for (ecological) economic planning in order to achieve economic democracy within planetary boundaries, as outlined e.g. in [Sa21] or [Sa24].

The two sections following the Introduction are essentially a compressed version of the (unpublished) paper [Bu25a]. In Section 2, our implementation of rotated bitboards in

[1] Berlin School of Economic and Law (HWR Berlin) , IPE - Institute for International Political Economy (associated member), Badensche Straße 52, 10825 Berlin, Germany, , https://orcid.org/0009-0002-6890-5298

doi:10.18420/env2025_02

FUSc# is discussed in detail. FUSc# is the chess program that was developed by the “AG Schachprogrammierung” at the Free University in Berlin ([Ch25a]). For more background of FUSC#, see [Bl05]. Section 3 deals with reinforcement learning in computer chess and beyond. After briefly discussing AlphaZero and its sample efficiency, we review some application of the ideas developed in AlphaZero in other domains. In addition, rotated binary neural networks are discussed briefly.

In the final part, we propose some new ideas how to tackle the problem of closing of material loops in a circular economy in order to optimize for sustainability. We also briefly discuss why this is urgently necessary to avoid climate tipping points [Ar22], and achieve absolute sustainability [HKR20]. According to a recent “Global Tipping Points Conference” ([Gl25]), there is growing scientific evidence that exceeding 1.5 °C global warming could trigger multiple climate tipping points. We are deeply convinced that these alarming facts from climate science necessarily demand research on what we call ’Progressive Computer Science’, as computer science (as well as many other fields) urgently need to contribute to fighting the ecological crisis.

2 Rotated Bitboards in FUSC#

The idea for the bitboard representation of the chessboard is based on the observation that modern CPUs are 64bit-processors, and those 64 bit correspond to the 64 squares on the chess board. Thus, very fast bit-wise operations are available. The bitboard method for representing a board game appears to have been invented in the mid-1950s, by Arthur Samuel and was used in his checkers program [Sa59]. A prominent chess program that used this technique successfully is the former computer chess champion “Cray Blitz”, written by Robert Hyatt, who continued to develop the program as an open-source project called “Crafty” ([Hy25]). DarkThought, developed at the university of Karlsruhe in the late 90s, is also using bitboards. Crafty and DarkThought were also the first programs that used an important refinement of the bitboard-representation called “rotated bitboards”. The author of DarkThought Ernst A. Heinz gives an overview of rotated bitboards as used in DarkThought (see [He00]), which inspired much of our implementation of rotated bitbords in FUSC#.

2.1 The bitboard-approach towards move-generation

The move-generation is used many times during the search-algorithms used in chess programs. Therefore, an efficient move-generation is needed. Based on the bitboard-approach, there exist different strategies for each of the piece-types in chess. One important concept is to compute bitboards of all possible moves (e.g. of a knight) from all the squares beforehand during the initialization of the program, and store this information in a data-structure that provides efficient access to these pre-computed moves during the move-generation. For non-sliding pieces, this approach works straightforward, but for sliding-pieces some more tricks are needed, because the possible moves for a sliding piece will depend on the configuration of the line/file/diagonal it is standing.

Therfore, the idea for bitboard-move-generation for sliding pieces is to compute all the possible moves for all squares **and** all configurations of the involved ranks/files/diagonals! For example, in FUSC#, the rank-moves for a rook standing on a1 on an otherwise empty chessboard are stored in "rank_moves[a1][00000001]", with the second index of the array being the configuration of the involved rank (i.e. 8 bits, with only "a1" being occupied as the rook is standing there itself). This works fine for rank-moves, because the necessary 8 bits for the respective rank can be easily obtained from the bitboard of the occupied pieces (this bitboard consists of 8 byte, and each of those corresponds to one rank). For file-moves of rooks and queens, and especially for the diagonal moves of bishops and queens, things turn out to be much more difficult: the necessary bits about the respective files/diagonals are spread all over the bitboard storing the "occupied" squares, they are not "in order", as they are for rank-moves. Here the idea of rotated bitboards helps out.

2.2 Rotated bitboards

The idea of rotated bitboards is to store the bitboards that represents the "occupied squares" not only in the "normal" way, but also in a "rotated" manner. Therefore, the necessary bits representing files/diagonals are "in-order" in those rotated bitboards, as needed by the move-generation (see previous section). The "rotated bitboards" are updated incrementally during the search, i.e. when a move is done or undone. The following bitboards are maintained in FUSC#:

- board.occ, which represents the occupied squares in the"normal" representation
- board.occ_l90, the board flipped by 90 °(for file moves)
- board.occ_a1h8, for diagonal moves in the direction of the a1h8-diagonal
- board.occ_a8h1, for diagonal moves in the direction of the a8h1-diagonal

The idea that such an "incremental update' of bitboards is possible whenever a move is made (during search) is essential, and this concept might also be useful when applying the idea of bitboards to other domains like sustainability (see below). For more details on our implementation of rotated bitboards in FUSC#, see [Bu25a].

2.3 Rotated Binary Neural Networks

Here, we shortly discuss Rotated Binary Neural Networks (see [Li20]) . Similar to rotated bitboards, the idea is a rotation (here of the full precision weight vector of a neural network) can significantly improve the performance of an algorithm, although of course the details of the implementation are very different (again bit-wise operations increase the speed). This encourages us that the idea of (rotated) bitboards is applicable to other domains such as sustainability that we discuss below.

Following [Li20], we observe that Binary Neural Networks (BNN) reduce the complexity of deep neural network, but there is a severe performance degradation. One of the major problems is the large quantization error between the full-precision weight vector and its binary vector. One idea is to compensate for the norm gap (this was done e.g. by [Qi20]) , but the angular bias was hardly touched. In the paper [Li20], the influence of angular bias on the quantization error is discussed and then a Rotated Binary Neural Network (RBNN) is introduced, which considers the angle alignment between the full-precision weight vector and its binarized version.

3 Reinforcement Learning in Computer Chess and Beyond

3.1 The Paper "Reinforcement Learning in Chess Engines" (2008)

We quote from the introduction of the paper "Reinforcement Learning in Chess Engines" ([Bl08]) from 2008, where the state of the art at that time was discussed, based on experiments with FUSC#: "The method presented in this paper optimizes the evaluation functions and its coefficients by automating the use of temporal differences and thereby increasing it's own understanding of chess after each game." Back then, the idea was that "the main problem lies in the correct tuning of the coefficients" of the evaluation functions, and that only there Reinforcement Learning could play a useful role in chess engines, not during the search. In the "related work"-section of the same paper [Bl08], a paragraph on "NeuroChess" is interesting to read from the perspective of today: It states that, based on the experience with "NeuroChess", experiments "showed that a learning strategy based on playing against oneself, does not yield satisfying results", which was the common sense in 2008, but changed dramatically with AlphaZero, less that 10 years later.

3.2 The Approach of AlphaZero

In December 2017, a paper was uploaded to arxiv with the title "Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm" [Si17]. Its reinforcement learning algorithm, "starting from random play, and given no domain knowledge except the game rules, achieved within 24 hours a superhuman level of play". So the game of chess was mastered "by tabula rasa reinforcement learning from games of self-play", in contrast to what seemed possible a decade ago. The approach of AlphaZero is very different to classical chess programs and FUSc#: Instead of an "alpha-beta search with domain-specific enhancements", it uses "a general-purpose Monte-Carlo tree search (MCTS) algorithm". Each search consists of a series of simulated games of self-play that traverse a tree" from root to leaf (compare p.3 of [Si17]).

But to achieve this impressive result, millions of self-play training games were used for the training: About 20 million games (approx. 4 hours of training time) were necessary to achieve super-human performance, and about 44 million games (approx. 9 hours of training time) were necessary to beat Stockfish, the best available computer chess program at the time ([Si17]). We now collect some ideas how to improve “sample efficiency” (i.e. to reduce the number of training games necessary). In the “Conclusion”-Section of the 2008-paper, it is stated that FUSc# considerably improved performance only after 119 (!!) games, and also “we estimate that FUSc# requires more than 50.000 training games ...” (p.9 of [Bl08]) - so in a way, there was a much higher “sample efficiency” (of course with limitations w.r.t the performance/chess playing strength, which peaked only at about 2000 ELO, which is ëxpert"level, but significantly below "grandmaster"). A promising strategy improve sample efficiency is to relate the “piece evaluation heuristics” to “position type” (e.g. opening, middle game, endgame), because the former will greatly vary with the latter (e.g. “king safety” in the middle game vs. “active king play” in the endgame). At first such a structure could be be “given”, but the aim is that a “FUSC#-Zero” will “re-discover” such “position-types” itself in a second step, maybe with even more “position types” that just 3 (according to [Bl08], in FUSc# 33 were used).

3.3 The “other Alphas”

We review the application of the ideas developed in AlphaZero in other domains, i.e. the “other Alphas” (all created by DeepMind) like AlphaFold [De21], AlphaTensor [Fa22] and AlphaProof. [De24]. AlphaFold [De21] had spectacular success applying the ideas of AlphaZero to predictions of protein structure. The most recent AlphaFold 3 was announced in May 2024, and it can predict the structure of complexes created by proteins with DNA, RNA, various ligands, and ions. Demis Hassabis and John Jumper of Google DeepMind shared one half of the 2024 Nobel Prize in Chemistry, awarded "for protein structure prediction” with AlphaFold (for more details see e.g. [De21]). AlphaTensor [Fa22] was developed to shed “light on a 50-year-old open question in mathematics about finding the fastest way to multiply two matrices”, and discovered new, faster algorithms to do so. AlphaProof and AlphaGeometry2 [De24] solved four out of six problems from the 2024 International Mathematical Olympiad (IMO), achieving the same level as a silver medalist.

4 Absolute Sustainability, LCA and Progressive Computer Science

In this final part of the paper, we propose some new ideas how to tackle the problem of closing material loops in a circular economy in order to optimize for sustainability. To the best of our knowledge, we are the first to propose using methods that originate in computer game research (described in the previous section) to tackle such questions. We were heavily inspired by the work of Spyros Samothrakis (see [Sa21] and [Sa24]).

We start by briefly discussing why this is urgently necessary to avoid climate tipping points [Ar22], and achieve absolute sustainability [HKR20]. According to the “Global Tipping Points Conference” ([Gl25]), there is growing scientific evidence that exceeding 1.5 °C global warming could trigger multiple climate tipping points (see also [Ar22]).

4.1 Absolute Sustainability and Climate Tipping Points

Absolute sustainability aims to achieve a development that stays within the Earth’s environmental limits, ensuring both present and future generations can meet their needs while respecting the planet’s carrying capacity, according to [HKR20]. This means that the environmental impacts of our activities must be assessed (with “Life Cycle Assessment”, see below) against established thresholds, such as planetary boundaries, to determine if they are truly sustainable, in an “absolute” sense (in contrast to “relative” sustainability where the question is e.g. “if product A is more sustainable than product B”, without reference to an “absolute” reference frame). Absolute sustainability thus focuses on staying within the safe operating space defined by planetary boundaries, which represent the limits of the Earth’s systems (see [Ro09]).

According to [Ar22], exceeding 1.5 °C global warming could trigger multiple Climate Tipping Points (CTPs). As the authors explain, “climate tipping points are conditions beyond which changes in a part of the climate system become self-perpetuating. These changes may lead to abrupt, irreversible, and dangerous impacts with serious implications for humanity.”. The paper synthesizes evidence for 16 core and regional-impact tipping elements and finds that exceeding 1.5 °C global warming could trigger multiple climate tipping points. As the authors conclude, assessment “ provides strong scientific evidence for urgent action to mitigate climate change. We show that even the Paris Agreement goal of limiting warming to well below 2 °C and preferably 1.5 °C is not safe as 1.5 °C and above risks crossing multiple tipping points”.

4.2 Life Cycle Assessment and Circular Economy

The environmental impact of products can be traced with “Life Cycle Assessment” (LCA). One example is openLCA (see openlca.org), a free and open source software for modeling the life cycle of products and sustainability. For a broader discussion of Life Cycle (Sustainability) Assessment in the context of Societal-Ecological Transformation, see [ZBT23]. Clearly, a sophisticated LCA is necessary for absolute sustainability and for the closing of material loops in order to optimize for sustainability in a Circular Economy.

In order to promote a Circular Economy in Germany, the “National Circular Economy Strategy” (NKWS) was adopted in December 2024 ([Bu24]). In that context, the recently published survey paper “AI for the Circular Economy - a tool for sustainable transformation?” ([Le25]) underlines the important role Artificial Intelligence (AI) can play for such a transformation, if employed in the right places and within the proper context. However, it also states clearly that the creation of a sustainable economic system is not merely a “technical” question, but a political one, and that currently, the way AI is used has in many cases very bad consequences for the environment, e.g. related to energy use. As mentioned in the introduction, we propose a “strategic triangle of AI for sustainability” in order to capture those ambivalence, see also [Bu25b].

4.2.1 Research Challenge

The fundamental problem of trying to understand how to transform the economy to be more “circular” requires moving from individual production processes to a more global view of the economy. We look at products and their pre-products, from an overall perspective in the economy (like with LCA-data), and observe that there are multiple ways of creating the same product. The aim is to choose the most “circular” and “sustainable” one in order to achieve absolute sustainability, which is a “global” question for the whole economy, as the choices on pre-products and production processes heavily depend on each other, and from a sustainability perspective, those choices will have economy-wide consequences. As a new approach, we propose to use state of the art methods from AI, Reinforcement Learning and Game Tree Search (e.g. Monte Carlo Tree Search that was successfully used in “AlphaZero”, see above) to investigate such questions. As a first step, we introduce some definitions in order to illustrate this idea and to describe the arising challenges more precisely.

4.2.2 Some definitions of products and life cycle assessment

Definition 1 *(Product) A product p is a triple* $p = (l, n, \{a_i\}_{i=1}^n)$*, where* $l \in \mathbb{N}$ *stands for the “level” of the product,* $n \in \mathbb{N}$ *stands for the number of other pre-products it is composed of, and there is a set consisting of* a_i *defining those pre-products and the associated quantities:*

$a_i = (q_i, p_i)$*, where* $i \in \mathbb{N}$ *and* $q_i \in \mathbb{R}$ *stands for the quantity of product* p_i *that is needed to produce product* p*.*

Remark 1 *Note that this is a recursive definition, as the product p appears “on both sides”. Thus, we define “raw materials” as the seed:*

Definition 2 *(Raw Material) Let there be* $\tilde{r} \in \mathbb{N}$ *raw materials in the economy.*

A “raw material product” is a product of “level 0”, i.e $raw = p = (0, 1, (1, r))$ *with* $r \in \{1 \ldots \tilde{r}\}$

Remark 2 *By this definition, all products and intermediate "pre-products" can be tracked back to which raw materials they contain.*

Definition 3 *(Life Cycle Assessment, LCA)*

Let there be $m \in \mathbb{N}$ different indicators for LCA measuring the impact of a product on the environment (and society).

For a product p, define its Environmental Impact over its lifecycle, as a tuple

$lca(p) = (b_j)_{j=1}^m$. *For simplicity, and following [HZ24], we assume that only 3 indicators are measured for each product*

- labor time t, i.e. the time necessary to produce p
- "climate cost" c, i.e. the amount of greenhouse-gases (e.g. CO_2-equivalents) the production produces
- raw materials $r = (r_1, ..., r_{\tilde{r}})$ necessary to produce p

Remark 3 *In this simplified setting, it holds that for each product p_i we have*

$lca(p_i) = (t_i, c_i, r_i)$

Definition 4 *(Recursive Definition of LCA for a product)*

Let p be a product, with $p = (l, n, \{a_i\}_{i=1}^n)$ and $a_i = (q_i, p_i)$, i.e. p is composed of n pre-products p_i with $i = 1...n$, as above. Then define the "environmental and societal impact" / "life cycle assessment" as

$lca(p) = \sum_{i=1}^n q_i * lca(p_i) = \sum_{i=1}^n (q_i t_i, q_i c_i, q_i r_i)$, *with the last step following from our "simplified setting" with only 3 indicators.*

4.2.3 Research Ideas

Now we are in a position to formulate concrete challenges on how to close material loops in a circular economy and to find the most sustainable way to produce a product. In order to do so, we assume the following (making these assumptions more precise is part of the challenge to formulate the problem in a way that modern AI-methods can be implemented):

- Assume different "production processes" are defined, which result in products with the same "features", but different environmental impact...
- ... which also means that products with the same "features" are composed by different pre-products, and the challenge is to find the most "sustainable" and "circular" way to

produce all goods in the economy (the latter meaning that material loops are closed where possible, i.e. that the unused output/"waste" from one production process is used as "input" for as many production processes as possible).

Now compute the LCA of a product, depending on the production process / the pre-products.

Remark 4 *For illustration, consider the following basic example:*
Let p be a product as above, and let $x := p_1$ be a pre-product of p. Assume that a new production process allows us to replace x by a new pre-product y, and denote this new product involving y by $\tilde{p}$. Then the environmental impact of p changes in the following way:

$lca(\tilde{p}) = lca(p) - lca(x) + lca(y)$

Definition 5 *(Challenge: The Game of a Circular and Sustainable Economy)*

Observe that searching this "space of possibilities" for an economy with many such possible "pre-product replacements" has similarities with a search tree in computer games like chess, as discussed above. With this analogy in mind, define

- a "move" in the game of a "Circular and Sustainable Economy": $\tilde{p} = move(p, x, y)$, i.e. such a "pre-product replacement" is considered a move in the game[2].
- as "evaluation function", the environmental impact of all products produced in the economy needs to be calculated - and according to the principle of "absolute sustainability", it must be within the "planetary boundaries"

Remark 5 *Even "bitboards" that have been used successfully in many board games may play a role in this setting, because the "incremental update" of "sustainability bitboards" (which can be defined according to the concrete optimization question at hand) by bit-wise logical operations might save a lot of computing time, as with "classical" board games.*

4.3 Progressive Computer Science

There are arguably many possible definitions of "Progressive Computer Science", depending how "societal progress" is defined. Here, we put the focus on preserving the planet for future generations, as urgent action is necessary in order to avoid the climate tipping points and to respect the planetary boundaries, i.e. to reach absolute sustainability (see above).

Definition 6 *("Progressive Computer Science")*

[2] Note that this does not depend on any linear relationships, and that with this idea, also highly interdependent and non-linear structures in the economy can be captured.

We define "Progressive Computer Science" as an umbrella for research in computer science that contributes to transforming the economic system towards absolute sustainability. This involves changing the rules on how the economy works in a way that ensures that planetary boundaries are respected.

An example that falls in this category of "Progressive Computer Science" is research on ecological economic planning with computer science methods, which is the topic of the next section.

4.4 Ecological Economic Planning and AI

Recent developments in AI (e.g. in the field of reinforcement learning) have enabled economic planning at large scales e.g. in multinational companies (see e.g. Phillips and Rozworski (2019). In the EU, a "Circular economy Action Plan" [Eu20] was adopted, and one central part of it is to improve the availability of digital product data. The "EU digital product passport" (see e.g. [Eu19] or [Eu22]) is designed to provide information about each product's origin, materials, environmental impact, and disposal recommendations. With this European-wide product data infrastructure for all products sold and used in the EU, new forms of ecological economic planning will be become possible. For a recent discussion of such methods and institutions regarding Artificial Intelligence and economic planning, see [Sa21], for a broader perspective see [GS25] or [Ze24].

5 Conclusion and Discussion

We admit that currently our proposal is more an outline of a research program than a practical implementation with empirical data. More research is necessary to explore the possibilities that modern AI-methods provide to change the current mostly linear and profit-driven way of organizing the economy towards a more circular, democratic and sustainable mode of production (see e.g. [HZ24] or [Sa24]). First steps were discussed at the conference "Democratic Economic Planning for the Real World"(see e.g. [Bu25b; Ri25])

Acknowledgments

The section on the closing of material loops in a circular economy in order to optimize for sustainability by using ideas in the spirit of AlphaZero is based on common discussions with Spyros Samothrakis, e.g. at a research stay in Essex in September 2024. We are deeply indebted to him for this inspiration; without it, this paper could not have been written. Also thanks to Christoph Sorg, Jan Groos, Eric Meier and other participants in the INDEP-network for introducing me to the recent discussion on the topic of democratic planning, as well as to Walther Zeug and the other members of the organizing committee of the workshop "KIT-KREIS" at the INFORMATIK Festival 2025 for a fruitful exchange.

References

[Ar22] Armstrong McKay, D. I. et al.: Exceeding 1.5 °C global warming could trigger multiple climate tipping points. Science 377 (6611), 2022, https://www.science.org/doi/10.1126/science.abn7950.

[Bl05] Block, M. et al.: Das Schachprojekt FUSc#, Working Paper, Berlin: Freie Universität Berlin, 2005, https://refubium.fu-berlin.de/handle/fub188/18583.

[Bl08] Block, M. et al.: Using Reinforcement Learning in Chess Engines, tech. rep., Berlin: Freie Universität Berlin, 2008, https://page.mi.fu-berlin.de/block/concibe2008.pdf.

[Bu24] Bundesministerium für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz: Die Nationale Kreislaufwirtschaftsstrategie, Berlin, 2024, https://www.bmuv.de/download/nationale-kreislaufwirtschaftsstrategie-nkws.

[Bu25a] Buchner, J.: Rotated Bitboards in FUSc# and Reinforcement Learning in Computer Chess and Beyond. arXiv preprint, 2025, arXiv: 2503.10822, https://arxiv.org/abs/2503.10822v1.

[Bu25b] Buchner, J.: The Strategic Triangle of AI for Ecological Economic Planning in a Circular Economy, 2025, https://virtual.oxfordabstracts.com/event/37937/submission/1243.

[Ch25a] Chessprogramming Wiki: FUSc# Project, https://www.chessprogramming.org/FUSCsharp, Accessed 27 July 2025, 2025.

[Ch25b] Chessprogramming Wiki: Reinforcement Learning in Computer Chess, https://www.chessprogramming.org/Reinforcement_Learning, accessed 27 July 2025, 2025.

[Ch25c] Chessprogramming Wiki: Rotated Bitboards, https://www.chessprogramming.org/Rotated_Bitboards, Accessed 27 July 2025, 2025.

[De21] DeepMind: AlphaFold, Wikipedia, 2021, https://en.wikipedia.org/wiki/AlphaFold.

[De24] DeepMind: AI Achieves Silver-Medal Standard Solving International Mathematical Olympiad Problems, DeepMind Blog, 2024, https://deepmind.google/discover/blog/ai-solves-imo-problems-at-silver-medal-level/.

[Eu19] European Commission: The European Green Deal, tech. rep. COM(2019) 640 final, Brussels: European Commission, 2019, https://eur-lex.europa.eu/resource.html?uri=cellar:b828d165-27f3-4b9b-a5fa-4eae4978aa8a.0003.02/DOC_1&format=PDF.

[Eu20] European Commission: Circular Economy Action Plan: For a Cleaner and More Competitive Europe, tech. rep. COM(2020) 98 final, Brussels: European Commission, 2020, https://eur-lex.europa.eu/resource.html?uri=cellar:9903b325-6388-11ea-b735-01aa75ed71a1.0001.02/DOC_1&format=PDF.

[Eu22] European Commission: Proposal for Ecodesign for Sustainable Products Regulation, tech. rep. COM(2022) 142 final, Brussels: European Commission, 2022, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52022PC0142.

[Fa22] Fawzi, A. et al.: Discovering Novel Algorithms with AlphaTensor, DeepMind Blog, 2022, https://deepmind.google/discover/blog/discovering-novel-algorithms-with-alphatensor/.

[Gl25] Global Tipping Points Website: Global Tipping Points Confercence, 2025, https://global-tipping-points.org/.

[GS25] Groos, J.; Sorg, C., eds.: Creative Construction – Democratic Planning in the 21st Century and Beyond. Bristol University Press, Bristol, 2025.

[He00] Heinz, E. A.: How DarkThought Plays Chess. In (Heinz, E. A., ed.): Scalable Search in Computer Chess. Computational Intelligence, Vieweg+Teubner Verlag, Wiesbaden, pp. 185–198, 2000, https://doi.org/10.1007/978-3-322-90178-1_13.

[HKR20] Hauschild, M. Z.; Kara, S.; Røpke, I.: Absolute sustainability: Challenges to life cycle engineering. CIRP Annals 69 (2), pp. 533–553, 2020, https://www.sciencedirect.com/science/article/pii/S0007850620301372.

[Hy25] Hyatt, R. M.: Crafty Chess Engine Homepage, https://craftychess.com/hyatt/hyatt.html, Accessed 27 July 2025, 2025.

[HZ24] Heyer, J.; Zeug, W.: Ökobilanz und kybernetische Wirtschaftsplanung: Demokratisch geplante Wirtschaft zur Befriedigung gesellschaftlicher Bedürfnisse in planetaren Grenzen. PROKLA. Zeitschrift für kritische Sozialwissenschaft 54 (215), pp. 267–286, 2024, https://doi.org/10.32387/prokla.v54i215.2116.

[Le25] Leitl, M. et al.: Künstliche Intelligenz für die Circular Economy – Ein Werkzeug für die nachhaltige Transformation?, Prospektiven – Circular Economy News 2025/01, Bottrop: Prosperkolleg e.V., 2025, https://prosperkolleg.ruhr/wp-content/uploads/2025/02/PK_PROSPEKTIVEN_PK_PROSPEKTIVEN_KI-20251.pdf.

[Li20] Lin, M. et al.: Rotated Binary Neural Network. In: Advances in Neural Information Processing Systems. NeurIPS 2020, 2020, arXiv: 2009.13055, https://arxiv.org/abs/2009.13055.

[Qi20] Qin, H. et al.: Forward and Backward Information Retention for Accurate Binary Neural Networks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Computer Vision Foundation / IEEE, pp. 2247–2256, 2020, https://openaccess.thecvf.com/content_CVPR_2020/papers/Qin_Forward_and_Backward_Information_Retention_for_Accurate_Binary_Neural_Networks_CVPR_2020_paper.pdf.

[Ri25] Rikap, C.: From corporate and military planning to democratic planning for people and the planet: AI as a testbed, 2025, https://virtual.oxfordabstracts.com/event/37937/submission/1019.

[Ro09] Rockström, J. et al.: A safe operating space for humanity. Nature 461 (7263), pp. 472–475, 2009, https://doi.org/10.1038/461472a.

[Sa21] Samothrakis, S.: Artificial Intelligence Inspired Methods for the Allocation of Common Goods and Services. PLOS ONE 16 (9), pp. 1–16, 2021, https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0257399.

[Sa24] Samothrakis, S.: Artificial Intelligence and Modern Planned Economies: A Discussion on Methods and Institutions. AI & Society 39 (6), pp. 2961–2972, 2024, https://doi.org/10.1007/s00146-023-01826-7.

[Sa59] Samuel, A. L.: Some Studies in Machine Learning Using the Game of Checkers. IBM Journal of Research and Development 3 (3), pp. 210–229, 1959.

[Si17] Silver, D. et al.: Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm. arXiv preprint, 2017, arXiv: 1712.01815, https://arxiv.org/abs/1712.01815v1.

[ZBT23] Zeug, W.; Bezama, A.; Thrän, D.: Life Cycle Sustainability Assessment for Sustainable Bioeconomy, Societal-Ecological Transformation and Beyond. In: Progress in Life Cycle Assessment 2021. Sustainable Production, Life Cycle Engineering and Management, Springer, Cham, pp. 131–159, 2023, https://link.springer.com/chapter/10.1007/978-3-031-29294-1_8.

[Ze24] Zeitschrift Luxemburg Editors: Zukunft mit Plan, Issue 1/2024 of Zeitschrift «Luxemburg», Berlin, 2024, https://www.rosalux.de/publikation/id/50930/zukunft-mit-plan.

Integrated data structures in materials management as the foundation for circular economy in SMEs

Stefan Jordan [1], Volker Wohlgemuth

Abstract: SMEs are pivotal to Europe's circular transition but often lack capabilities to acquire and operationalize sustainability data. The contribution proposes an integrated approach that anchors triple-bottom-line attributes as data fields in material and supplier masters and acquires them automatically in the background. A crawl–validate–integrate workflow harvests public sources, i.e. LCI/LCA repositories, EPDs, aligns records via ontology-based data access, and assures fitness with a machine-learning quality filter and human-in-the-loop curation. The minimal, extensible model includes among others material-level CO_2e, recycled content, recyclability rate and supplier location emissions. Provenance, method, and validity are centrally tracked. A first validation against SAP Materials Management indicates that a small set of Z-fields and a lightweight architecture (OBDA, REST/GraphQL, low-code ETL) can make indicators available to purchasing, planning or inventory management supporting decisions in operational business management and with enough data collected for strategical decisions on business models.

Keywords: SMEs, ERP systems, materials management, sustainability data, artificial intelligence, twin transformation, data-driven sustainability transformation, circular economy, digital product passport, automatic data acquisition, triple bottom line

1 Introduction

Since 2019, the European Green Deal (EGD) has defined the guideline for a climate-neutral European Union by 2050 and, with an interim target of a 55 percent reduction in greenhouse gases by 2030 [Th19]. It has established a legally binding transformation goal that explicitly focuses on resource efficiency and the circular economy. In all three core areas of the EGD – circular economy, energy efficiency and emission reduction the use of data is essential for identifying areas for action, monitoring, controlling and implementing actions to adress them [Th19].The Circular Economy Action Plan (CEAP), which builds on this, specifies this requirement by highlighting digital product and process data as a key factor in making material flows transparent, closing resource loops and enabling new business models [Ci20].

Sustainability data according to CEAP can be defined as as all quantifiable economical

1 HTW-Berlin University of applied science, Fachbereich 2, Wilhelminenhofstr. 75A, Berin, 12459, jordans@htw-berlin.de, https://orcid.org/0009-0006-9153-886X

doi:10.18420/env2025_03

cost and value as well as ecological and social impact factors of a product or process over its life cycle [Ci20]. This definition is related to the triple-bottom-line approach [Je00] which is one of the key concepts connecting economical performance with environmental and social impacts. This definition is being referred to throughout the whole contribution when the "sustainability data" is being used. On the one hand, it includes compliance data that is required by law and can be found, for example, in the mandatory fields of the European Sustainability Reporting Standards (ESRS) [Eu23], and, on the other hand, performative data that goes beyond the minimum legal requirements and contributes directly to value creation, for example in the form of real-time footprints or risk-based supplier metrics [KM25].

Small and medium-sized enterprises (SMEs) face structural limitations in view of this data paradigm because they often lack the financial and human resources to collect sustainability data in the required depth, harmonise it semantically and integrate it seamlessly into existing information systems [PM24]. The lack of standardised data structures in enterprise resource planning systems [Wl04] [Al24b] not only makes it difficult to fulfil growing reporting obligations in a cost-efficient manner, but also prevents the use of the performative potential of circular business models , threatening to create an uneven playing field between large companies and SMEs [Pü24].

Against this backdrop, this paper examines how sustainability data can be automatically embedded in material and supplier master and transaction data so that SMEs can make both operational decisions in procurement and scheduling and strategic decisions towards circular value creation on a valid, integrative data basis.

2 State of research

2.1 Drivers for collection of sustainability data

Drivers that motivate companies to collect and use sustainability data started shifting recently [Pü24] from the rule-driven understanding that prevailed. The focus was on complying with legal requirements such as the reporting and due diligence obligations of the Corporate Sustainability Reporting Directive (CSRD) [Co22], and on maintaining legally compliant emissions and supply chain metrics through mechanisms such as the Carbon Border Adjustment Mechanism (CBAM) [Ca23] following the logic of the "licence to operate" paradigm. However, it this compliance is costly and often overwhelming for SMEs [PM24]. Recent political developments confirm this assessment: The EU's so-called Omnibus Initiative has reduced the scope of disclosure originally expected by the European Sustainability Reporting Standard by 68% of mandatory disclosure points in the short term and eliminated mandatory disclosure of listed SMEs starting from 2027 as SMEs in particular would otherwise have been unable to cope with the effort [EU25].

Parallel to this debate on relief, a survey of 1,600 corporate and financial executives conducted by the World Economic Forum (WEF) points to the growing strategic importance of integrated sustainability and financial data as performative data. The vast majority of executives asked see the linking of both data worlds as a substantial competitive advantage that goes far beyond mere compliance effects [KM25]. The survey emphasises that reliable Environmental, Social, Governance (ESG) indicators are now perceived as "strategic asset for agility and innovation" and are increasingly weighing as heavily as traditional cost factors such as raw materials or energy in investment and pricing decisions [KM25].

Sustainability data is no longer solely assigned to reporting systems, but identified as an element of performative corporate management that must be embedded in operational decisions if efficiencies are to be realised and new business models developed [PM24]. Integrated data structures that combine financial and sustainability-related information and can be implemented technically without prohibitive costs are considered a decisive lever in making this change accessible to SMEs as well. Only the combination of real-time production and supplier data with environmental and social indicators paves the way for data-driven optimisation and innovation processes [Pü24]. The semantic integration hurdles that arise specifically in heterogeneous IT landscapes, underscore the need for domain-specific data models and lightweight interfaces [PM24].

2.2 Processing of sustainability data in Enterprise resource planing software

Conventional enterprise resource planning (ERP) systems have been the basis of the IT infrastructure of many industrial companies for decades, but their primary strength lies in the procedural integration of finance, logistics and production, not in processing and analytics of sustainability information [Al25]. A case study on SAP S/4HANA shows that environmental indicators such as recycling rates cannot be recorded or evaluated in the standard model. Only after expanding the material master tables and adding additional business add-ins such indicators can be mapped at all [AG18]. Without expanding the material master, only a very limited ecological scope is available, making further analyses of product sustainability virtually impossible

This structural gap is particularly serious because ERP systems are widely used in the relevant industries: 95 percent of large manufacturing companies [Stat17]and 31% of all companies used in Germany an ERP solution [An24] . But even widely used products have so far only considered energy management to be a "marginal issue," so that despite high ERP penetration, there are hardly any integrated functions for environmental or resource data.

In the scientific debate, this gap is increasingly being described as an opportunity for innovation. Wunderlich et al. diagnose a clear "fit/gap" between the data requirements of a triple bottom line approach and the existing ERP data objects and call for domain-specific extensions and ontology tools to bring together material and sustainability data

[Wu18]. Putrama and Martinek classify these integration hurdles as typical for heterogeneous SME landscapes and see the core of the problem less in missing report exports than in a lack of semantic consistency, which prevents advanced analyses such as real-time CO_2 calculations [PM24]. Conceptual approaches to sustainable ERP, which since Chofreh et al. have been promoting an expanded range of functions along the three dimensions of sustainability, confirm the feasibility of such extensions, but point to the considerable investments in data architecture, project management and strategic anchoring that only large companies have been able to afford so far [CG14].

2.3 Standardized data exchange with the Digital Product Passport (DPP)

The DPP is regarded as a key instrument for the technical implementation of the transparency and circularity targets defined in the Ecodesign Regulation for Sustainable Products (ESPR). The latest CEN-CENELEC Workshop Agreement (CWA 18186:2025) explicitly refers to the DPP as an integrated information tool for the circular economy because it provides, for the first time, a structured, digitally accessible data core for all actors in a value chain [Gu25]. The European Commission expects that this will support data-based purchasing, usage and end-of-life decisions, thereby closing material cycles [Re24].

The EU project CIRPASS in particular clarifies in its system roadmap report from 2023, that the necessary DPP standards should be finalised "within the next five years" and that the first interoperable implementations should be available by 2027 for the priority product groups of batteries, textiles and electronics [Jo24].

While the data structures for most product groups are still pending, DIN SPEC 99100 provides a complete schema for the battery passport, which defines 130 mandatory attributes and a unique QR code and will be mandatory under the Battery Ordinance from 2027 [DI25]. But this content-fixed model will be of limited use without consistent semantic mappings in existing ERP environments because heterogeneous identifier systems, proprietary data formats and missing ontologies hinder automated data exchange [PM24]

The recently adopted European Data Act (EDA) can be seen as an catalyst, as it legally secures fair access to IoT and product data from 2025 onwards, thus creating the basis for open, cloud-based DPP platforms [Eu23b]. However SMEs, given their limited financial and IT resources, will hardly be in a position to implement complete DPP workflows and establish the necessary semantic data structures in the short term [Jo24] [PM24]. As SMEs account for the vast majority of companies, and a big share of gross domestic product in Germany [An25] [Ve25] but also other EU countries like Italy or Spain [Ma24] such a delay would significantly impair the ecologically and economically desirable broad impact of the circular economy.

2.4 Integration of heterogeneous sustainability data in business software

Traditional Extract-Transform-Load (ETL) processes, in which data is physically extracted, transformed and loaded into dedicated repositories, are considered costly and inflexible because each new data source requires extensive mapping and loading procedures [PM24]. As a resource-saving alternative virtual integration approaches in which federated queries or so-called data fabrics leave the data at the source and only link it at runtime can be used. A promising methodology is Ontology-Based Data Access (OBDA), which maps relational tables using a semantic intermediate schema, thus enabling technical queries across heterogeneous structures [Ca19]. Virtual integration patterns could serve as a resource-saving alternative that keep data at the source and link it only at query time, thus reducing duplication and improving timeliness [PM24].[Ca19]

Engines such as Ontop have matured to support SPARQL-to-SQL rewriting with soundness guarantees and industrial-scale performance [Ca19]. Application ontologies can act as pragmatic gateways into OBDA setups in manufacturing "Internet of Production" contexts, lowering adoption barriers by aligning operational schemas to domain vocabularies incrementally [Br23]. Nevertheless, even modular OBDA setups depend on sensible ontology scope and careful mapping governance; developing and maintaining such domain artefacts still requires expertise and time [PM24]

Complementary to semantics, machine learning (ML) is increasingly leveraged to reduce manual data wrangling. "Humans-in-the-loop" AutoML configurations are emphasised for schema/signal discovery under governance constraints, showing that active learning and curated feedback substantially improve label quality and mapping suggestions in real settings [CF21] [Br23].

From a data-engineering angle, automated feature extraction, anomaly detection and outlier handling are considered as critical building blocks in automated data processing pipelines, where data quality assurance is a prerequisite for trustworthy downstream analytics [MM25] [Br23]. Earlier semantic-integration studies report progress on (semi-)automatic mapping between ontologies and relational schemas which are useful for aligning sustainability attributes to ERP tables. But noisy or sparse input can degrade mapping fidelity without manual curation checkpoints [AG18] [WW17]

Overall virtual, semantics-based and ML-assisted integration promises considerable rationalisation potential for sustainability data, but their effectiveness depends significantly on the availability of consistent ontologies, meaningful training data and established data quality metrics.. SMEs in particular face the challenge of creating these infrastructural prerequisites with limited resources and no proficiency, making the development of lightweight, preconfigured integration solutions a priority research area [Br23] [Ca19] [PM24].

3 Need for action

3.1 Extending operational data structures in ERP systems

Materials management in enterprise systems rests on a small set of stable artefacts, most prominently the material master and the supplier master, around which operational functions such as purchasing, inventory management and inventory analysis are organised [Go22]. The material master provides the canonical key for planning, sourcing and valuation (for example item identifiers, units of measure, price or standard cost, lead times and bill-of-materials linkages), while the supplier master anchors commercial terms, locations and compliance status [Go22]. In current practice, these masters are optimised for economic control. Environmental and social information even when present, is often handled in separate tools or spreadsheets, which breaks traceability across transactions and impedes systematic evaluation of alternatives in day-to-day operations [WW17] [FJ24]. Isolating environmental and social indicators outside the operational data backbone leads to manual joins, media breaks and poor reuse in processes such as purchasing, scheduling and inventory analysis [Br23]. Therefore extending core masters with sustainability fields is a prerequisite for meaningful life-cycle-based assessments and automation along the materials flow [AG18] [Go22]

System analyses of ERP-systems like SAP S/4HANA show that the lack of corresponding attributes blocks the derivation of holistic life cycle assessments[Wu18]. Without fields for CO_2 factors, recycled content, durability/repair proxies or hazardous-substance flags in material masters and transaction records, life-cycle assessments, take-back logic, and reuse/remanufacturing planning remain manual, error-prone and fragmented [Ve25].

The expansion of master data by fields for sustainability data represents the lowest common denominator for covering both regulatory reporting requirements and performative analyses. This means that in future, each material object could, for example, be assigned the recycled content, a CO_2 equivalent and a hazard indicator in addition to quantity, quality and value variables. Supplier entries could be supplemented by, i.e. location emission factors and social indicators supporting supply chain due dilligence. Latter will need further research and will be discussed in a future contribution.

CIRPASS already implements the attribute logic for the Digital Product Passport by requiring the consistent assignment of emissions, circularity and quality-related data along the value chain [Jo24]. Once the data fields are integrated in the master data of ERP-systems, automated assignment procedures can pull and update values from open sources such as the Oekobaudat database, Ecoinvent database or future DPP registers, significantly reducing the cost and personnel for SMEs.

Tab 1. illustrates extended data fields based on regulatory requirements; it shows that relevant information can largely be derived from existing open data sources or future DPP registers, provided that the master data has the necessary semantic capacity.

Master data object	Example KPI	Public data source	Regulatory requirement
Material	CO_2 equivalent per unit	Ecoinvent	CBAM
	Recycled Content (%)	DPP-Register	ESPR
	Risk Index	ECHA-SVHC	REACH
Supplier	Location emission factor	EDGAR	CBAM
	Social risk score	UNDP-Atlas	CSDDD
	Certificate status	ISO 14001	CSRD

Tab. 1: Data object with sustainability attribute, public data source and regulatory requirements

3.2 A pragmatic approach to data collection

The state of research illustrates that classic ETL routines are neither viable for SMEs from considering financial and personal ressources. Virtual, ontology-based integration methods only deliver reliable results if the underlying raw data is continuously updated and quality-assured [PM24]. This leads to the need for a pragmatic data collection approach, that relies on automation, open data sources and machine learning plausibility checks.

Automated crawlers can continuously extract data from public sources such as Ecoinvent and other public data bases (see Tab. 1) or pre-stage data already available from the DPP sandbox at specified intervals. This would follow the recommendation of the CIRPASS roadmap to use open product and process data as a low-threshold entry point into the DPP logic [Jo24]. Access is legally supported by the interoperability requirements of the Data Act, reducing licensing barriers to a minimum [Eu23b].

Another important technique is semantic matching. ODBA methods are used to match the heterogeneously structured crawler results with the additional data fields of the material and supplier master data. This enables consistent and accurate assignment of data to the corresponding attributes in the master data.

An AI-supported quality module addresses the data quality risk that has been repeatedly criticised in the literature. An active learning ensemble assigns a confidence value to each data record and only forwards entries with a confidence value below i.e. 0.8 to human reviewers. This significantly reduces the manual validation effort, as only exceptional cases require personal expertise, while routine updates are fully automated. [WW17].

Such a crawl–validate–integrate workflow could close the gap between existing open data sets and the extended ERP master data. At the same time, it meets the two key requirements highlighted by the SME survey in CIRPASS: minimal staffing levels and immediate system compatibility [Al24a]. This creates a robust data basis that enables both reporting on environmental impacts, i.e. Scope 3-CO_2-emissions, required by regulations and the performative use of sustainability indicators in day-to-day business.

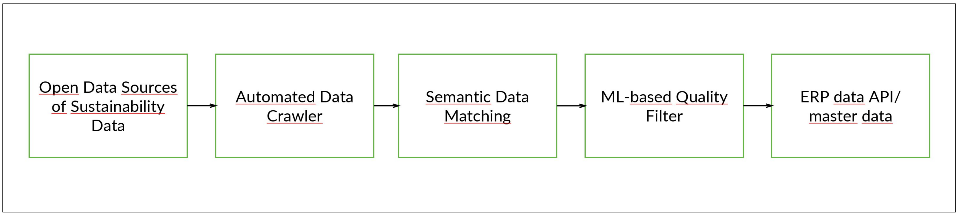


Fig.1 Crawl–validate–integrate pipeline

Figure 1 outlines the resulting crawl–validate–integrate pipeline and illustrates how the AI-supported quality filter could be inserted between data extraction and ERP persistence. By combining these approaches, SMEs could efficiently and cost-effectively collect and use their sustainability data to make informed decisions and improve their sustainability performance.

3.3 Closing the architectural gap for extended data structures in SME software

Missing sustainability fields in ERP systems [Wu18], cost-driving ETL pipelines [PM24] and DPP interfaces that will only be fragmentarily standardised by 2027 [Jo24], identified as deficites of ERP-systems (see Tab. 2), make it clear that SMEs need a new type of software architecture that extends existing business applications without requiring extensive core adjustments. Research and project experience point to a hybrid integration paradigm that combines three elements: an ontology layer for semantic interoperability, lightweight REST or GraphQL APIs for real-time data exchange, and low-code tools to automate routine ETL tasks.

Ontology-based data access allows domain-specific sustainability terms to be mapped to relational material and supplier tables without forcing physical data replication; OBDA approaches can reduce implementation effort by up to forty percent compared to classic data marts, if the ontology is encapsulated as a separate service layer [PM24]. The CIRPASS roadmap for the DPP recommends a decentralised resolver concept based on REST and GraphQL endpoints to make product and passport information accessible beyond proprietary ERP boundaries [Jo24]. Thirdly, practical reports on cloud-based environmental information systems show that low-code ETL engines can reduce the ongoing maintenance effort for period runs even in SMEs if they are deployed on container-based open source stacks [WW17].

Current capabilities of typical SME ERP systems	Missing function block	Recommended technology component
Static material and supplier master data without sustainability attributes	Semantic mapping of CO_2, risk and recycling data	Ontology layer (OBDA)
Point-to-point interfaces, proprietary formats	Open, transaction-oriented data access	REST/GraphQL-API-Gateway
Manual CSV uploads for CO_2 factors	Ongoing, automated data integration	Low-Code-ETL with Container-Deployment

Tab. 2Architectural shortcomings of today's ERP landscapes and recommended technology layers based on Wu18, WW17, PM24, CI24

Closing this architecture gap is a prerequisite for technically supporting the mandatory field logic formulated in section 3.1 and for permanently operating the crawl–validate–integrate pipeline outlined in section 3.2. The specific design of such an extension package is presented in the following chapter.

4 Proposed concept

The proposed solution is aimed at SMEs. As these enterprises lack the financial [Du25] and personal resources for collecting sustainability data as well as the required proficiency for leveraging the harvested data in there business operation this contribution proposes an integrated data structure where sustainability data is integrated into material management data and collected automatically in the background.

4.1 Integrated data structures in materials management

Foundational for an integrated data structure are three requirement classes: (i) regulatory requirements from current and upcoming EU or national legislation (e.g. CBAM, ESPR), (ii) performative requirements to enable analytics that improve enterprise performance, and (iii) functional requirements in purchasing, inventory analysis, and inventory management.

From a regulatory perspective, several EU instruments now explicitly require product- and supplier-related sustainability information in machine-readable form. CBAM mandates declarations of embedded emissions for specified goods and distinguishes verified actuals from default values with verifier-backed records [Ca23]. ESPR and the forthcoming DPP standardise information sets for carbon, resource use, substances of concern, and circularity characteristics (e.g. durability, repairability, recycled content, recyclability), with links to ISO 14044/14067, the GHG Product Standard [Gree11], and emerging circularity metrics [Is25]. CWA 18186 specifies the corresponding data structures and exchange semantics for product and component data [Gu25]. The Battery

Regulation already mandates a QR-based passport with recycled-content targets and operational traceability from 2027, coupling compliance with circular outcomes such as higher secondary-material shares and second-life design [Co22b]. Meeting these obligations without media breaks requires that recycling indicators and product-/supplier-level emissions be available as structured master-data attributes [DI25].

Anchoring triple-bottom-line (TBL) data—environmental, social, and economic—as first-class fields in material and supplier masters is therefore essential, especially for SMEs. It creates a shared "sentence structure" for analytics and interoperability [AG18][PM24][Gr14], enabling consistent capture of product- and supplier-level impacts (including Scope-3 CO_2 emissions) across life-cycle stages and their translation into operational and strategic decisions such as secondary-material sourcing, repair, reuse, and remanufacturing. Non-financial indicators (e.g. product-level carbon intensity, recycled content, due-diligence status) must be structured and linked to financial performance via cause-and-effect chains to guide resource allocation rather than merely external reporting [Fi01][Ca19]. Integrating environmental metrics into planning, budgeting, and variance analysis—emphasising data quality, targets, and threshold-based steering—requires stable, queryable attributes in operational masters instead of ad-hoc spreadsheets [Gr14][KM25].

Functionally, the integrated data structure serves three purposes. Purchasing, inventory management and inventory analysis.

In purchasing, sustainability attributes must be queryable alongside price and lead time to rank suppliers and items by combined economic-environmental-social KPIs (e.g. EUR/ton, kg CO_2e/ton, social-risk index) and to support CBAM-ready documentation at goods-receipt and invoice-line level [Ca23][Gu25].
In inventory management, master-level circularity fields—recycled content, repairability, remanufacturability, recyclability—support segregation strategies, substitution rules, and take-back allocation; ESPR/DPP guidance makes these attributes machine-resolvable for decisions across hand-over points [Gu25][Re24].
In inventory analysis and planning, life-cycle and Scope-3 factors should join the same fact structures as quantities and values to permit period close, scenario analysis, and supplier switching with sustainability deltas computed like standard variances; environmental informatics and eco-controlling literature recommends embedding such indicators in the enterprise backbone to enable alerts and thresholds in operational control loops [WW17][AG18].

To satisfy these regulatory, performative, and functional requirements—and to prepare for the DPP—the data model links the material master to a minimal, extensible set of TBL attributes, complemented by supplier parameters for location-based CO_2 emissions and due-diligence status. Applying recognised standards ensures methodological consistency and comparability: ISO 59020 and DIN SPEC 99100 for circularity-related data, ISO 14067 and the GHG Product Standard for carbon measurement [Is25][DI25][Th19]. Storing recycled content, resource-use indicators, and durability/repair proxies at master level makes them available to procurement, MRP, and recovery processes without custom joins and enables after-product-life KPIs (e.g. reuse, recyclability) in standard queries and dashboards [Gu25].

For social information, supply-chain due-diligence and certification status should be captured in supplier masters to enable risk-aware sourcing and audit planning; role-based access and long-term availability are addressed explicitly in the DPP guidance [Gu25]. Finally, forward compatibility with the DPP requires stable identifiers (e.g. GS1/QR links), versioning, and resolvable links to external registries so that machine-readable passports can be connected as delegated acts roll out and as JTC-24 APIs for searchability and life-cycle management become available [Gu25].

Data field name	Data field Description	Typical source
Dangerous-goods profile	Standard ERP profile for DG handling	ERP
Product carbon footprint	Cradle-to-gate CO_2e per declared unit (CBAM-ready)	EPD/LCI (ISO 14067; 14044/EN 15804); Supplier
Recycled content (%)	Mass-based recycled share	EPD/DPP (ISO 59020; CWA 18186); Supplier
Recycling rate at EoL (%)	Mass fraction entering recycling at end-of-life	ISO 59020; DPP/CWA 18186
Repairability index	Proxy for ease of repair/ upgrade (0–10 or scheme)	DPP/CWA 18186; ISO 59020
Durability class	Proxy for expected lifetime category	DPP/CWA 18186; ISO 59020
Primary energy demand	Total primary energy per declared unit	EPD/LCI (ISO 14044; EN 15804; CWA 18186)
GS1 GTIN	Global Trade Item Number for product identification	GS1 (forward compatibility for DPP)
DPP UID	Digital Product Passport unique identifier	DPP registry (CWA 18186; CEN/CENELEC JTC 24)
Supplier Due-diligence status	Supply-chain due diligence alignment	Supplier/Compliance (CSDDD)
Geography scope	Geographical scope of dataset	Source (CWA 18186 info sets)

Tab.3: Excerpt from the integrated data structure for materials management

Tab. 3 is showing an excerpt of the proposed data structure including the typical sources[2]. The full data structure is consisting of 40+ additional data fields that are integrated with existing, industry specific materials and supplier management master and transaction data structures.

4.2 Acquiring and processing heterogeneous sustainability data from public sources

As sustainability transformation is no exact science, publicly accessible sustainability data constitute a pragmatic foundation for building sustainability management capacities in the majority of SMEs. Therefore secondary datasets from public or private sources of life-cycle inventory (LCI) and assessment (LCA) data as well as environmental product declaration (EPD) (see Tab. 4) are considered sufficiently robust to provide a baseline and to track directional improvements over time. Current integration research recommends leveraging existing open or low-barrier sources first and increasing primary data share progressively as standards such as the Digital Product Passport (DPP) mature [PM24][CI24]. In practice, this implies acquiring and validating from two classes of sources: I) structured, machine-readable endpoints (APIs, CSV/RDF feeds) and ii) semi-structured documents (typically PDF-based EPDs or producer datasheets) that require controlled extraction and semantic anchoring.

Public data source	Content coverage	Update interval	Licence agreement
Ecoinvent v3.10	Life cycle CO_2 factors for > 20,000 materials	annually	CC-BY-NC
EDGAR v8.0	Territorial emission factors (Scope 3)	half-yearly	CC-BY-4.0
EU DPP Sandbox	Preliminary data on batteries, textiles, electronics	quarterly	EUPL-1.2
ECHA-SVHC	Hazardous substance classification	continuously	CC-BY-4.0
UNDP-Atlas	Social risk index (country level)	annually	CC-BY-NC

Tab.4: Public available Data sources with content coverage, update intervall and license

An important requirement for integrating the heterogeneous sustainability data from two classes of sources are a robust acquisition routine and integration workflow (see Fig. 2). Given the limitations of SMEs in terms of the high cost of physical ETL processes and the expertise required for robust semantics, this contribution proposes a

[2] Final structure still subject to development. Release planned for 12/25.

crawl-validate-integrate workflow that combines automation, open data sources, and an ODBA- and ML-based plausibility checks to populate the extended material and supplier data.

Low-overhead crawlers harvest public data relevant for material and supplier attributes (see Tab. 3), prioritising authoritative sources and structured endpoints where available. For structured endpoints, like Ecoinvent, background services perform scheduled retrievals that expose material- or process-level indicators (for example greenhouse-gas intensities, recycled content shares, or hazard classifications). The capture logic prioritises documented interfaces and machine-readable formats in order to minimise the effort required for downstream data processing and to guarantee the origin of the data. This strategy is advocating open product and process data, interoperable interfaces and decentralised resolvers as emphasized by CIRPASS [CI24]. The European Data Act can further reduce legal and technical barriers for programmatic access and switching between processing services from September 2025 onward, thereby strengthening SMEs' ability to collect and reuse device- and service-generated data under interoperable conditions [EU23b].

The second class of sources are semi-structured documents, which demand a different treatment. Many producers publish EPDs and sustainability datasheets as downloadable PDFs; these artefacts often encode declared units, system boundaries and impact modules in human-readable layouts rather than in clean tables. The acquisition routine therefore combines template-guided parsing with layout-aware extraction and applies explicit unit normalisation, code-list checks and completeness tests before accepting any values.

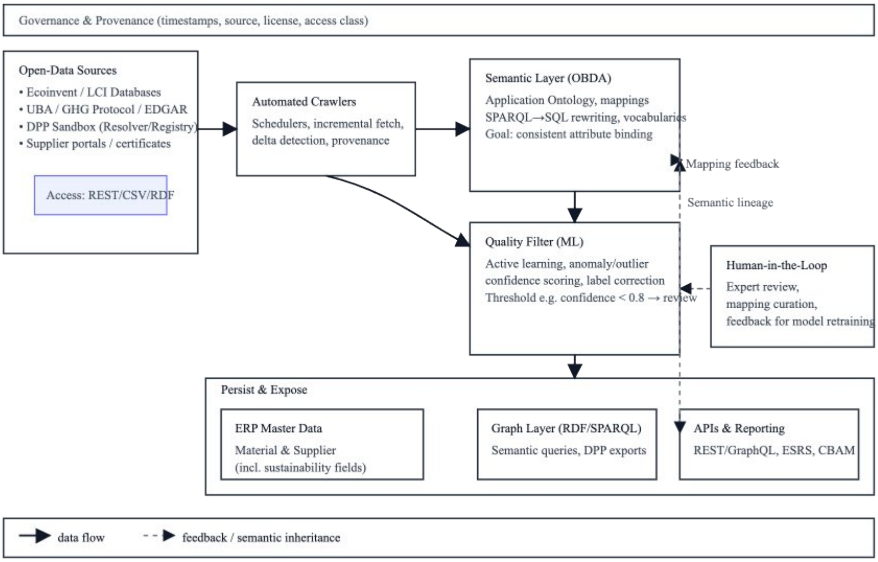


Fig.2: Automated Crawl-Validate-Integrate Workflow with Human-in-the-Loop feedback

Where templates vary substantially across issuers or product groups, a semantic anchoring step is required. The proposed solution will use application ontologies that provides a controlled vocabulary for key attributes (for example “global warming potential per declared unit,” “recycled content share,” “hazard class”), while OBDA mappings bind extracted fields to this vocabulary without forcing physical replication of the source data [Ca19]. Manufacturing case work demonstrates how such application ontologies can act as gateways into OBDA setups and thus stabilise matching logic across heterogeneous artefacts [Ba23]. This semantic step is essential to ensure that values originating from distinct layouts or issuers become comparable, searchable and auditable.

To mitigate the well-documented data-quality risks in sustainability datasets, the pipeline inserts an ML-based quality filter between ingestion and ERP persistence. Following enterprise AutoML practice, a human-in-the-loop scheme flags low-confidence assignments (e.g., confidence < 0.8) for expert review, while high-confidence cases are auto-accepted; this reduces manual effort without sacrificing accountability [FG21] [MM25]. For noisy or weakly labelled inputs—typical in scraped PDFs—modern label-quality methods can identify and correct suspect labels during training, which measurably reduces false positives in subsequent anomaly detection [SYM23].

The crawl–validate–integrate pattern closes the gap between open datasets (and future DPP registries) and extended ERP master data. It aligns with CIRPASS’ interoperability stance and the Data Act’s interoperability objectives, while matching SME realities [CI24] [EU23b]

5 Discussion

The proposed crawl–validate–integrate approach addresses key concerns in environmental informatics: material efficiency through data-based transparency and the use of open data sources. The lightweight architecture based on open-source software services (OBDA layer, API gateway, low-code ETL) can be placed in front of business software or ERP cores, minimising customisation effort. As machine learning can support the validate-phase and experiences in SMEs shows adaptable ML methods can improve efficiency and reduce waste rates [WWR22][GBB21], yet consistent schemas and domain knowledge remain limiting factors. Research needs to clarify which thresholds (e.g., confidence 0.8) are optimal, how much maintenance time active learning saves in real data flows, and how robust ontology-based matching (OBDA) remains against noisy open data feeds [SYM23].

Also how stable semantic links actually are across heterogeneous keys (material numbers, certificate IDs) and how governance/provenance is operationalised (e.g. change/audit trails) needs to be clarified.

The need for research should be addressed in lighthouse projects or reference implementations with metrics (integration effort, error rate, throughput time) and comparison to physical ETL alternatives, also in cooperation with ERP manufacturers.

Finally, the environmental impact of digital infrastructures themselves requires assessment. Always-on pipelines and cloud services consume significant amounts of energy and resources, and rebound effects can offset efficiency gains. Studies on ICT-related raw materials highlight risks across the life cycle [WOF23][WWR22]. Research should determine the net ecological balance of SME-scale crawl–validate–integrate pipelines, exploring operating modes (e.g., batch frequency, edge vs. cloud) that minimise footprints while maintaining data quality.

The proposed concept envisages a decoupled architecture: lightweight, preferably open-source services (OBDA layer, API gateway, low-code ETL) in front of the ERP core to ensure release capability and gradual introduction. This encapsulation takes into account typical SME constraints and minimises deep core customisation. It remains unclear how stable semantic links across heterogeneous keys (material numbers, certificate IDs) actually are and how governance/provenance is operationalised (e.g. change/audit trails). The need for research should be addressed in lighthouse projects or reference implementations with metrics (integration effort, error rate, throughput time) and in comparison to physical ETL alternatives, also in cooperation with ERP manufacturers.

Finally, the environmental impact of digital infrastructures themselves requires assessment. Always-on pipelines and cloud services consume energy and resources, and rebound effects can offset efficiency gains. Studies on ICT-related raw materials highlight risks across the life cycle [WOF23][WWR22]. Research should determine the net ecological balance of SME-scale crawl–validate–integrate pipelines, exploring operating modes (e.g., batch frequency, edge vs. cloud) that minimise footprints while maintaining data quality.

6 Conclusion

Empowering SMEs for the twin transformation of digitalisation and sustainability requires an integrated approach. On one hand, digital maturity must increase significantly; on the other, ecological and social data should be collected with minimal additional effort and embedded directly in operational decision-making in order to support the adaption of sustainable, circular business models in the medium to long term. Because SME management is strongly economically oriented, adoption is most likely when clear business cases show disproportionate cost savings or revenue potential. Open question remain: What economic arguments actually motivate SMEs to digitise and consider sustainability data in their decisions?

The contribution shows that integrated data structures along the triple bottom line, anchored as attributes in the material and supplier master, can form the basis for making sustainability information operationally usable in SMEs. In conjunction with the proposed crawl–validate–integrate pipeline, which combines open data sources, semantic access (OBDA) and an AI-supported quality filter, an implementation path is outlined that addresses regulatory requirements (CBAM/ESPR/DPP) and performative control needs in day-to-day business. The data set and processing architecture can be considered as a foundation for integrating circularity-relevant key figures into procurement, scheduling, inventory analysis and return processes, thereby enabling data-based cycle orientation in regular operations.

Open issues regarding scalability and robustness of semantic mappings across heterogeneous data spaces, the net effort of continuous data collection and maintenance (including human-in-the-loop components), and the net ecological impact of the digital infrastructure itself should be adressed. In a field test or a lighthouse project, e.g. with an ERP manufacturer reliable evidence of successful implementation should be created. In a representative product group the effectiveness of the approach should be measured using defined metrics such as allocation accuracy, data maintenance effort, savings and substitution potential in procurement, and material and energy consumption, including possible rebound effects.

Bibliography

[AG18] Arndt, H.-K.; Marx Gómez, J.; Wohlgemuth, V.; Lehmann, S.; Pleshkanovska, R. (Hrsg.): *Nachhaltige Betriebliche Umweltinformationssysteme: Konferenzband zu den 9. BUIS-Tagen*. Springer Gabler, Wiesbaden, 2018.

[Al24a] Alcayaga, A.; Berg, H.; Hoffmann, N.; Brüggmann, H.; Maus, F.: *The Digital Product Passport (DPP) for the Circular Economy: Recommendations for policy, business and IT*. 2024.

[Al24b] Alzahmi, W.; Al-Assaf, K.; Al Sheikh, R.; Bahroun, Z.: *Sustainable ERP Systems: Current Trends, Challenges, and Future Directions*. In: Proc. IEEE ICTMOD, IEEE, Sharjah (UAE), 1–8, 2024.

[Al25] Alzahmi, W.; Al-Assaf, K.; Alshaikh, R.; Bahroun, Z.: *Towards Sustainable ERP Systems: Emerging Trends, Challenges, and Future Pathways*. *Management Systems in Production Engineering*, 33(1), 24–38, 2025.

[An22] Statistisches Bundesamt: *Anzahl der Unternehmen in Deutschland nach Unternehmensgröße im Jahr 2022*, abgerufen am 08.08.2025.

[An24] Statistisches Bundesamt: *Anteil der Unternehmen in Deutschland mit Nutzung einer ERP-Software, nach Wirtschaftszweig im Jahr 2023*, 2024

[Br23] Brecher, C.; Müller, A.; Nauck, T.; Storms, S.: *Mapping Application Ontologies as a Gateway into OBDA Systems in the Internet of Production*. In: Arseniev, A.; Aouf, N. (eds): *Cyber-Physical Systems and Control II*, Springer, Cham, 195–204, 2023.

[Ca19] Calvanese, D. et al.: *Ontop: Answering SPARQL Queries over Relational Databases*. IOS Press, 2019.

[Ca23] European Parliament and Council: *Carbon Boarder Adjustment Mechanism (CBAM)*, 2023.

[CF21] Crisan, A.; Fiore-Gartland, B.: *Fits and Starts: Enterprise Use of AutoML and the Role of Humans in the Loop*. In: *CHI '21 Proceedings*, ACM, 2021.

[Ch14] Chofreh, A. G.; Goni, F. A.; Shaharoun, A. M.; Ismail, S.; Klemeš, J. J.: *Sustainable enterprise resource planning: Imperatives and research directions*. *Journal of Cleaner Production*, 71, 139–147, 2014.

[Ci20] European Commission: *Circular Economy Action Plan*, 2020.

[Co22] European Parliament and Council: *Corporate Sustainability Reporting Directive (CSRD)*, 2022.

[DI25] DIN/DKE: *SPEC 99100 – Anforderungen an Datenattribute des Batteriepasses*, 2025.

[Du25] Statistisches Bundesamt: *Durchschnittliche Umsatzrenditen von mittelständischen Unternehmen in Deutschland nach Branchen im Jahr 2022*, abgerufen am 08.08.2025.

[Eu23] EFRAG: *European Sustainability Reporting Standards (ESRS)*, 2023.

[Eu23b] European Parliament and Council: *European Data Act*, 2023.

[EU25] EMAS: *EU-Omnibus – Neue Regeln für CSRD-Berichtspflicht*, https://www.emas.de/…, abgerufen am 27.02.2025.

[Fi01] Figge, F. et al.: Sustainability balanced scorecard: wertorientiertes Nachhaltigkeitsmanagement mit der Balanced Scorecard, Lüneburg, 2001

[FJ24] Fröhlich, E.; Jamal, Y. (Hrsg.): *CSR und Beschaffung: Die Bedeutung des Einkaufs für eine nachhaltige Transformation*. 2. Aufl., Springer, Berlin/Heidelberg, 2024.

[GBB21] Gensch, C.-O. et al.: *Deutschland auf dem Weg zur Klimaneutralität: Welche Chancen und Risiken ergeben sich durch die Digitalisierung?*, 2021

[Go22] Goel, H.: *Handbook for SAP PP in S/4HANA*. Apress, Berkeley (CA), 2022.

[Gr14] Steinke, K.-H. et al.: *Green Controlling Leitfaden für die erfolgreiche Integration ökologischer Zielsetzungen in Unternehmensplanung und-steuerung*, 2014

[Gu25] CEN Workshop: *Guidelines to create a Digital Product Passport – the CircThread experience*, 2025.

[Is25] ISO 59020:2024-05 *Circular economy. Measuring and assessing circularity performance*, 2025

[Je00] Jeurissen, R.: *John Elkington, Cannibals With Forks… Journal of Business Ethics*, 23(2), 229–231, 2000.

[Jo24] Jousse, A.: *DPP System roadmap*, 2024

[KM25] Klindt, J.; McReynolds, M.: *Why sustainability data is as strategically important as financial data*. World Economic Forum, https://www.weforum.org/…, abgerufen am 08.08.2025.

[Ma24] Mammadov, H.: *Adopting Digitalisation: The Impact on SME Performance in Evidence from Italy and Spain*, UNIVERSIDAD DE SEVILLA, 2024

[MM25] Mumuni, A.; Mumuni, F.: *Automated data processing and feature engineering for deep learning and big data applications: A survey*. *Journal of Information and Intelligence*, 3(2), 113–153, 2025.

[PM24] Putrama, I. M.; Martinek, P.: *Heterogeneous data integration: Challenges and opportunities*. *Data in Brief*, 56, 2024.

[Pü24] Püchel, L. et al.: *On the Pivotal Role of Data in Sustainability Transformations: Challenges and Opportunities*. *Business & Information Systems Engineering*, 66(6), 831–848, 2024.

[Re24] European Parliament and Council: *Ecodesign for Sustainable Products Regulation (ESPR)*, 2024.

[Sh17] Statistisches Bundesam: *Share of large companies using enterprise resource planning software in Germany in January 2017, by industry sector*, 2017

[SYM23] Shu, J.; Yuan, X.; Meng, D.: *CMW-Net: an adaptive robust algorithm for sample selection and label correction*. *National Science Review*, 10(6), nwad084, 2023.

[Th19] European Commission: *The European Green Deal*, 2019.

[Ve22] Statistisches Bundesamt: *Verteilung der Bruttowertschöpfung zu Faktorkosten der Unternehmen in Deutschland nach Unternehmensgröße im Jahr 2022*, abgerufen am 08.08.2025

[WWR22] Willenbacher, Martina ; Wohlgemuth, Volker ; Risch, Lisa: *Neural Networks for Energy Optimization of Production Processes in Small and Medium Sized Enterprises*. In: *Advances and New Trends in Environmental Informatics*, 129–145, 2022

[Wl04] Wohlgemuth, V.; Naumann, S.; Lang, C.: *Exchanging Environmental Relevant Data between ERP Systems and BUIS Using PAS 1025*. In: *Sh@ring: 18th EnviroInfo*, Éditions du Tricorne, Genève, 2004.

[WW17] Wohlgemuth, V.; Winter, M.: *Konzeption eines cloudbasierten stoffstromorientierten BUIS*. In: *Industrie von morgen*, BWV, Berlin, 2017.

[Wu18] Wunderlich, S. et al.: *Identifizierung und Abbildung von Produktnachhaltigkeitsdaten in SAP S/4HANA*. In: [AG18], 63–71, 2018

Spatiotemporal analysis of $PM_{2.5}$ in Nanjing, China, using network theory and PCA

Georgios Georgoulis[1], Tijian Wang[2] and Kostas Karatzas[1]

Abstract: This study investigates the spatial and temporal correlation patterns of $PM_{2.5}$ concentrations recorded by over 280 air quality monitoring stations in Nanjing, China, over the course of one month (December 2019). Employing a combination of Spearman correlation analysis, network analysis, and Principal Component Analysis (PCA), we explore the dynamic connectivity between stations over the diurnal cycle. The objective is to determine whether the city behaves as a spatially cohesive unit or as a collection of localized communities in terms of $PM_{2.5}$ levels. The analysis identifies critical hours with maximum correlation coherence and reveals periods of spatial fragmentation. This approach provides insights that can inform data-driven environmental management and urban air quality assessment.

Keywords: Air quality, Spearman correlation, Network analysis, $PM_{2.5}$, PCA, Environmental informatics

1 Introduction

Air pollution, particularly due to fine particulate matter ($PM_{2.5}$) is a critical concern in urban areas worldwide, with documented impacts on human health, visibility and climate. $PM_{2.5}$ particles, which have a diameter of less than 2.5 micrometres, can penetrate deep into the respiratory tract and even enter the bloodstream, causing cardiovascular diseases. As urbanization accelerates and vehicular emissions increase, the need for detailed, high-resolution analysis of pollution patterns has become paramount. The ability to quantify and visualize how pollution behaves over time and space is essential for city-wide planning and environmental health strategies.

Traditionally, air quality monitoring has relied on statistical summaries, time-series decomposition, or interpolation techniques. While valuable, these approaches often fail to capture the dynamic and interconnected nature of pollution across dense urban networks. Emerging methodologies, inspired by data science and network theory provide a more nuanced framework. Spearman correlation, for example, allows for non-linear relationships detection between sensor outputs, and Network Analysis (NA) can reveal

[1] Aristotle University, Environmental Informatics Research Group, School of Mechanical Engineering, Thessaloniki 54124, Greece, kkara@auth.gr https://orcid.org/0000-0002-1033-5985
[2] Nanjing University, School of Atmospheric Science, Nanjing 210023, China, tjwang@nju.edu.cn https://orcid.org/0000-0001-9599-3874

doi:10.18420/env2025_04

hidden structural patterns in monitoring data, such as community structures or influential hubs [DZ17]. Principal Component analysis (PCA), on the other hand, aids in identifying latent temporal patterns and reducing the complexity of high-dimensional datasets.

In this study, we integrate these approaches to examine an extensive $PM_{2.5}$ monitoring dataset from Nanjing, China. Our aim is to assess whether the city's air quality functions as a cohesive single network or as distinct localized sub-networks, and to pinpoint key time periods that exhibit either maximum coherence or significant fragmentation in pollution patterns.

2 Materials and methods

2.1 Study area and data sources

The analysis is centred on Nanjing, a major city in eastern China. The dataset contains hourly concentrations of 7 pollutants (NO_2, SO_2, CO, O_3, TVOC, PM_{10}, $PM_{2.5}$) over a full calendar month (December of 2019), from 282 air quality monitoring stations across the city, while no additional data (like meteorological observations) were made available. The Cambri AER2 - series (Shenzhen Cambri Environmental Technology Ltd., China) is the air quality monitoring device used. The dense and widespread spatial distribution of these devices (Figure 1) makes Nanjing an ideal case for high-resolution correlation and network analysis. This means that we address these devices as a network where each unit represents a node (that has attributes, i.e. concentration values of pollutants of interest), and the lines that may be envisaged connecting all those nodes are the network edges (representing potential relationships or connections between the devices).

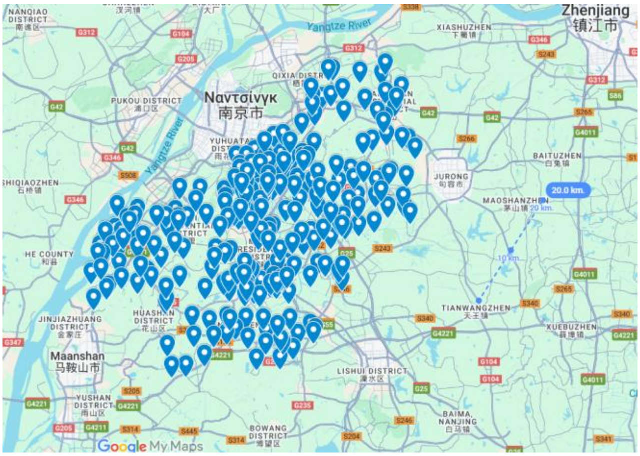


Fig. 1: Geospatial Distribution of the Air Quality Monitoring Stations Across Nanjing. An indicative scale has been included for distance estimation.

2.2 Methodology

Data preprocessing

Data cleaning involved filtering out rows with missing or non-relevant pollutant values. Timestamp fields were converted to Python datetime objects and decomposed into separate 'Date', 'Hour', and 'DateHour' columns. Concentration fields were cast to numeric types to ensure proper aggregation and correlation calculations.

Metrics and plots

We computed descriptive statistics (mean, standard deviation, minimum, maximum, skewness, and kurtosis) of concentrations across all stations for each pollutant and for each hour of each day. Boxplots were generated for each pollutant and date to visualize diurnal patterns and variability. In addition, and to evaluate the similarity in pollutant behaviour across stations, Spearman rank correlation matrices were computed for each hour of the day. Each matrix represents pairwise correlations of $PM_{2.5}$ values between all stations at that specific hour, reflecting how synchronized or divergent the pollutant dynamics are across the network. For each hour, the mean Spearman correlation of each station with all others was also calculated. This yielded a 282 × 24 matrix that characterises the temporal "centrality" or influence of each station. Stations with higher average correlations are those that consistently align with broader pollution trends. We then identified the top 10 stations with the highest average correlation at each hour therefore providing with a temporal view of the most synchronized stations in the pollution network.

Next, we examined the off-diagonal entries of the hourly correlation matrices to derive a 24-hour profile of spatial coherence. Excluding the main diagonal (self-correlations) and any perfect correlations ($|r| = 1$), we computed the average inter-station correlation for each hour [ZM21]. This temporal coherence highlights the hours of highest and lowest spatial synchronization in $PM_{2.5}$ levels across the city.

Principal Component Analysis

For each hourly correlation matrix, the upper triangular off-diagonal values were extracted and flattened into a vector representing that hour's spatial correlation profile. We applied a threshold of $r > 0.4$ to each vector to filter out weak correlations. The resulting vectors from all 24 hours were then combined into a 24 × N matrix, where N is the total number of station pairs with $r > 0.4$ in at least one hour. We performed PCA on this matrix, retaining 10 PCs to capture the dominant modes of variability [Kr23]. A scree plot was used to examine the variance explained by each component, and a 3D scatter plot of the first three principal components was generated to visualize the clustering of hours and shifts in station behaviour in the reduced-dimensional space.

Network Analysis

For each hour, we constructed a network graph where nodes represent monitoring stations and edges connect station pairs with Spearman correlation $\rho > 0.6$ (using the correlation value as the edge weight). In environmental data analysis, Spearman correlation values above 0.6 are generally considered to indicate moderate to strong monotonic associations. Such filtering has precedent in atmospheric and sensor network studies, where correlations thresholds between 0.5 and 0.7 are commonly used [LI22].

Key network metrics (node degree, density etc) are computed to assess the overall connectivity and complexity of the system. Community detection identifies densely connected groups of stations (subnetworks) that behave similarly in terms of PM2.5 dynamics. These graph-theoretic metrics quantify the evolving structural behaviour of the station network over time [Vl21].

3 Results and discussion

As part of the initial statistical exploration, descriptive metrics for each pollutant were calculated on an hourly basis. Fig. 2 shows a snapshot of the computed values for $PM_{2.5}$. Each row corresponds to a specific hour and includes the mean, standard deviation, minimum, maximum concentrations, along with skewness and kurtosis. These indicators offer a multifaceted view of the pollutant's temporal dynamics. The data reveal a clear diurnal pattern. The specific $PM_{2.5}$ levels are generally higher during early morning (e.g., 0:00–6:00) and late afternoon to evening hours (16:00–23:00), which correlates with increased human activity, traffic congestion, and possibly industrial operations. For instance, at 18:00 on Dec 1st, the mean concentration peaks at 68.7 µg/m³, indicating a potential pollution episode.

Pollutant	Date	Hour	Mean	Std Dev	Min	Max	Skewness	Kurtosis
PM2.5	2019-12-01	0	39,74206	8,846666	20,37373	66,89312	0,73731	0,255115
		1	35,47724	7,996032	18,12749	63,4289	0,62517	-0,12525
		2	35,65992	6,345172	21,96897	61,15474	0,86708	0,960845
		3	33,49925	8,058019	20,47161	61,3834	1,18887	1,293207
		4	30,85923	7,229466	18,67006	62,49007	1,379422	2,660828
		5	28,78656	8,186775	14,42691	54,60121	0,518282	-0,10337
		6	26,82668	7,531809	15,07086	47,09575	0,566218	-0,68448
		7	23,96541	6,593224	14,51925	51,80502	1,651838	3,15972
		8	27,04202	6,667591	15,00445	48,83425	1,021028	1,022579
		9	27,52143	6,332253	14,12711	67,94593	1,602616	6,285344
		10	27,94723	5,213968	14,91794	67,76965	1,707589	11,47429
		11	26,76644	4,980806	15,3346	54,53233	0,925032	3,217375
		12	26,27236	4,233085	17,06871	55,31897	1,512813	7,992318
		13	27,21015	4,166982	17,26887	50,1081	0,986952	4,321658
		14	29,45257	4,416918	17,37753	43,75729	0,273278	0,609428
		15	30,14754	4,152865	18,27735	42,79836	0,040964	0,284829
		16	55,63405	9,209799	31,05306	103,1533	0,760544	2,381652
		17	40,04287	3,745617	29,18513	59,4248	0,250166	2,391453
		18	68,73554	12,59546	36,2603	126,8273	0,257479	1,604599
		19	32,61269	3,972184	21,39698	50,09524	0,030633	2,350793
		20	38,66311	5,570438	23,02882	68,68296	0,517526	3,938416
		21	42,02836	5,611769	23,51782	70,20358	0,011411	3,928548
		22	44,92726	6,058244	25,24972	82,63899	0,413139	6,467839
		23	46,30589	6,631792	27,48297	91,38252	1,120877	8,573812

Fig. 2: Hourly Descriptive Statistics for $PM_{2.5}$ concentrations on December 1–2, 2019

Figure 3 illustrates the diurnal variation of $PM_{2.5}$ concentrations for a representative day (December 7, 2019), shown as boxplots of the hourly concentration distribution across all stations. Each boxplot summarizes the central tendency, dispersion and outliers for a given hour of the day. The results highlight a clear diurnal pattern:

- Morning to midday (7:00-12:00), concentrations begin to increase steadily, peaking around 9:00 to 10:00. This aligns with increased traffic and reduced atmospheric dispersion.
- Afternoon (13:00-16:00), a moderate decline is observed, possibly due to improved dispersion conditions.
- Evening (18:00-23:00), a second rise in concentrations is evident, particularly after 19:00, reflecting the evening rush hour and accumulation effects under cooling atmospheric conditions.

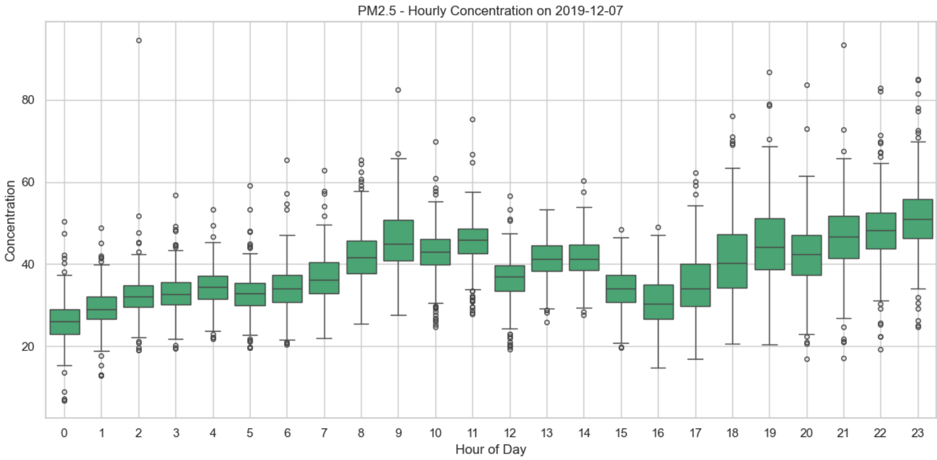


Fig. 3: Diurnal Variation of $PM_{2.5}$ concentrations across stations in Nanjing, on December 7, 2019.

The Spearman correlation matrix that was calculated provided with valuable insights. Figure 4 presents the diurnal evolution of the average Spearman correlation coefficient among all monitoring stations in Nanjing for $PM_{2.5}$ concentrations. The plot excludes perfect correlations (±1) and captures only meaningful off-diagonal relationships. This indicator reflects the overall degree of spatial coherence - i.e., how similarly pollutant levels fluctuate across the city during each hour of the day.

Key observations include:

- High coherence throughout the day: From 01:00 to 16:00, the average ρ remains consistently above 0.85, peaking between 10:00 and 13:00. This suggests that during daylight hours, stations exhibit strong synchrony in pollutant patterns—likely due to shared exposure to emissions, meteorological uniformity, or traffic flows.

- Sharp decline in the evening: After around 19:00, a marked decrease in interstation correlation is observed, reaching a minimum around 21:00. This drop may reflect the influence of localized sources (e.g., residential heating), boundary layer breakup, or other nocturnal atmospheric effects leading to heterogeneous $PM_{2.5}$ behaviour.
- Late-night recovery: The average correlation begins to rebound after midnight, indicating a partial restoration of uniform pollutant dynamics.

These findings highlight the non-stationary structure of spatial correlation over time and provide motivation for hour-by-hour network modelling. The evening period emerges as a phase of spatial fragmentation, warranting further analysis of network modularity and community formation.

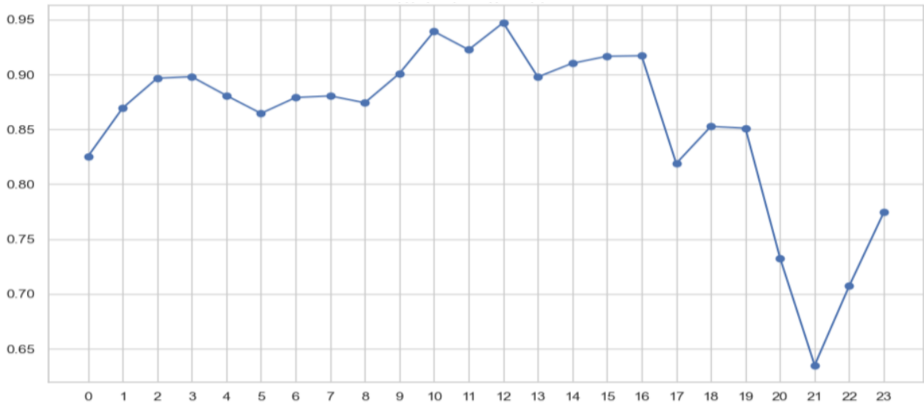


Fig. 4: Hourly profile of mean Spearman Correlation for $PM_{2.5}$ among all monitoring stations in Nanjing. Horizontal axis depicts the hour of the day (in local time) while the vertical axis presents with the correlation coefficient value.

The analysis continued with the calculation of the PCA based on the Spearman correlation matrices as presented in the prev. chapter. For each hour, the upper triangle of the correlation matrix—comprising all non-redundant, off-diagonal values—was extracted and flattened into a vector. These vectors represent the spatial configuration of $PM_{2.5}$ coherence at each hour and were used to construct a 24 × N matrix, where N is the number of unique station-to-station relationships after thresholding. The scree plot in Figure 5 shows the proportion of explained variance for the first 10 principal components. It is evident that PC1 to PC3 account for approximately 60–65% of the total variance, indicating that a significant portion of the temporal variability in spatial correlation structure can be captured using just three dimensions. For this reason, and in Figure 6, a 3D scatter plot visualizes each hour of the day as a point in the PC1–PC3 space. On this basis hours 0-19 appear to be relatively clustered, indicating that during most of the day the spatial structure of $PM_{2.5}$ correlations among all monitoring stations remains similar. In contrast, hours 20-23 diverge significantly from the main cluster, implying a distinct spatial network configuration in the late evening. These hours correspond to periods previously identified as having lower average interstation correlations and higher modularity. The separation of these late hours in PC space supports the hypothesis of a temporal shift in the network's coherence, possibly due to local emission sources, boundary layer changes, or reduced traffic homogenization effects.

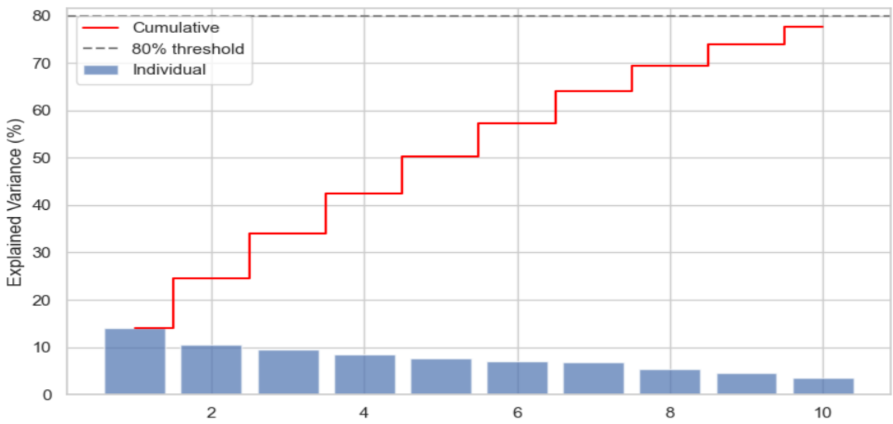


Fig. 5: Scree Plots Showing explained variance of the first 10 PCA components.

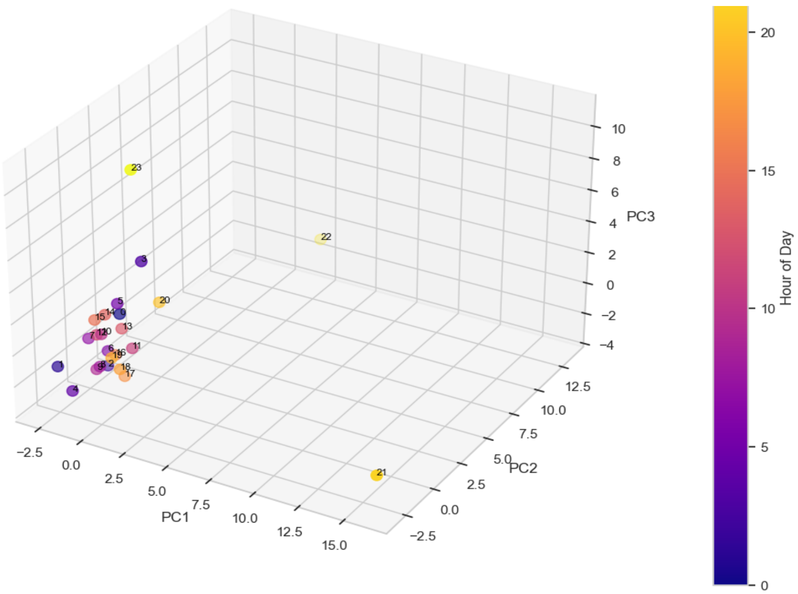


Fig. 6: 3D Visualization of PCA components (PC1, PC2, PC3)

Finally, we examined the graph-theoretic properties of the $PM_{2.5}$ correlation network at each hour (using $\rho > 0.6$ to define edges). Key network metrics are summarized as follows:

- Node count: The number of active nodes (stations) remains nearly constant at about 281 throughout the day, indicating that almost all stations reported valid data at all hours.
- Edge count and average degree: Both metrics reach their highest values during the late morning and early afternoon (approximately 10:00–15:00), reflecting a very well-connected network of stations at those times and thus high overall spatial coherence,
- Network density: Network density exceeds 0.99 during several midday hours, meaning nearly every possible station pair is connected by a strong correlation. This confirms the uniformly coherent behaviour of the pollution network in the middle of the day.
- Community count: For most hours, the correlation network splits into only 2–3 communities, indicating minimal fragmentation of the city into independent sub-regions. However, in the late evening (around 21:00–23:00) the number of communities rises (up to 4), signifying a resurgence of fragmentation as the station network divides into more distinct clusters.
- Modularity: Consistent with the above, modularity values are extremely low (< 0.02) during the high-coherence midday period, which is typical of a strongly interconnected network with little community structure. In the evening, by contrast, modularity increases sharply (for example, reaching ~0.23 at 22:00), suggesting a phase of spatial disintegration where stations form well-defined but weakly interconnected subgroups [Fe20].

Figure 7 presents a comprehensive table of graph-theoretical metrics computed from hourly Spearman correlation networks for PM2.5, using a correlation threshold of $\rho > 0.6$ to define edges. Each row represents one hour of the day, and the values are derived from the corresponding undirected weighted graph of interstation connections.

- Edge count and average degree are highest during midday (e.g., hours 10–15), suggesting a well-connected sensor network with strong overall correlations.
- Network density exceeds 0.99 in several hours, especially between 10:00 and 15:00, further confirming uniform correlation behaviour during this interval.
- Community count (NumCommunities) remains low (2–3) for most hours, indicating limited fragmentation in the sensor network. However, late evening hours (e.g., 21–23) see a resurgence of fragmentation with higher community counts (up to 4).
- Modularity values remain low (<0.02) during high-coherence periods, which is typical for strongly interconnected networks. In contrast, sharp increases in modularity appear during the evening and early night (e.g., hour 22: modularity = 0.23), suggesting a phase of spatial disintegration.

Hour	Nodes	Edges	AvgDegree	Density	NumCommunities	Modularity
0	281	37390	266,1209964	0,950432	3	0,040478585
1	281	38566	274,4911032	0,980325	3	0,021599129
2	281	38272	272,3985765	0,972852	4	0,00771696
3	281	38379	273,1601423	0,975572	4	0,011079856
4	281	37594	267,5729537	0,955618	3	0,014089596
5	280	37486	267,7571429	0,959703	3	0,014201657
6	280	38279	273,4214286	0,980005	2	0,011679466
7	280	38260	273,2857143	0,979519	2	0,012698059
8	280	37998	271,4142857	0,972811	2	0,017238184
9	280	38575	275,5357143	0,987583	2	0,008853348
10	281	39301	279,7224199	0,999009	3	0,005135129
11	281	39206	279,0462633	0,996594	2	0,012423838
12	281	39217	279,1245552	0,996873	2	0,007916528
13	281	39143	278,5978648	0,994992	2	0,015234877
14	281	39302	279,7295374	0,999034	3	0,012561559
15	281	39311	279,7935943	0,999263	2	0,008033067
16	281	38910	276,9395018	0,98907	2	0,009260324
17	281	36535	260,0355872	0,928699	2	0,039320643
18	281	38558	274,4341637	0,980122	3	0,020669806
19	281	38176	271,7153025	0,970412	2	0,022721893
20	281	31404	223,5160142	0,798271	2	0,071487052
21	281	22884	162,8754448	0,581698	3	0,232371444
22	281	28554	203,2313167	0,725826	4	0,098759449
23	281	34135	242,9537367	0,867692	3	0,055148293

Fig. 7: Hourly evolution of network metrics based on PM2.5 correlation

4 Conclusions

This study demonstrates the utility of combining correlation analysis, PCA, and network analysis to gain insights into the spatiotemporal structure of urban $PM_{2.5}$ pollution. Using hourly Spearman correlation matrices (and their upper-triangular values), we characterized the dynamic inter-station relationships of $PM_{2.5}$ throughout the day in Nanjing, China.

The results indicate a cyclic behaviour of the sensor network's coherence: midday hours revealed strong uniformity with dense graph structures and low modularity, while late evening hours marked significant fragmentation, both in terms of lower mean correlations and rising community separation. PCA provided dimensionality reduction that confirmed these structural shifts, particularly isolating hours 20–23 as outliers in the system's temporal evolution. This aligns with possible real-world causes such as shifts in emission sources, traffic decay, or boundary layer transitions.

Overall, our analytical approach offers a replicable and scalable framework for treating an urban air quality monitoring network not just as a passive collection of sensors, but as a dynamic system with distinct temporal modes of behaviour. Such insights can support targeted interventions, adaptive monitoring strategies, and early-warning systems for pollution clustering events.

Bibliography

[DZ17] Dai Y-H.; Zhou W-X.: Temporal and spatial correlation patterns of air pollutants in Chinese cities. *PLoS ONE 12(8),* e0182724, 2017. https://doi.org/10.1371/journal.pone.0182724

[Fe20] Ferreira L.N.; Vega-Oliveros D.A.; Cotacallapa M.; Cardoso M.F.; Quiles M.G.; Zhao L.; Macau E.E.N.: Spatiotemporal data analysis with chronological networks. *Nat Commun. 11(1),*4036. 2020. https://doi.org/10.1038/s41467-020-17634-2

[Kr23] Krzyśko M.; Nijkamp P.; Ratajczak W.; Wołyński W.; Wenerska B.: Spatio-temporal principal component analysis. *Spatial Economic Analysis*, *19*(1), 8–29, 2023. https://doi.org/10.1080/17421772.2023.2237532

[LI22] Liu X.; Zhao J.; Lin S.; Li J.; Wang S.; Zhang Y.; Gao Y.; Chai J.: Fine-Grained Individual Air Quality Index (IAQI) Prediction Based on Spatial-Temporal Causal Convolution Network: A Case Study of Shanghai. *Atmosphere 13(6),* 959, 2022. https://doi.org/10.3390/atmos13060959

[MA22] Maher B.A.; Gonet T.; Karloukovski V.V.; Wang H.; Bannan T.J.: Protecting playgrounds: local-scale reduction of airborne particulate matter concentrations through particulate deposition on roadside 'tredges' (green infrastructure). *Sci Rep 12*, 14236, 2022. https://doi.org/10.1038/s41598-022-18509-w

[Vl21] Vlachogiannis D.M.; Xu Y.; Jin L.; González M.C; Correlation networks of air particulate matter ($PM_{2.5}$): a comparative study. *Appl Netw Sci. 6(1):32,* 2021

[ZM21] Zafra-Mejía C.; Suárez J.; Pachón J.: Public Health Considerations for PM10 in a High-Pollution Megacity: Influences of Atmospheric Condition and Land Coverage. *Atmosphere. 12. 118*, 2021. https://doi.org/10.3390/atmos12010118.

Sustainable Software in Practice: A Use-Case-Oriented Tool for Exploring Sustainability Criteria

Anas Abukhadijah[1], Lejla Begic-Fazlic[1], Achim Guldner[1], Sebastian Weber[1], Sebastian Bast[1], Guido Dartmann[1], and Stefan Naumann[1]

Abstract: With digital transformation and the rise of artificial intelligence (AI), AI-based software systems have become essential to modern society, raising concerns about their environmental impacts. Sustainable software engineering integrates ecological considerations into software development processes. Existing sustainability frameworks and tools often lack interactivity and practical integration. This paper presents an interactive tool that introduces capabilities to support the accessibility and usability of a sustainability criteria catalog through dynamic filtering, interactive visual dashboards, and contextual markdown documentation. This approach supports decision-making, encourages practical application, and fosters feedback loops to refine sustainability standards.

Keywords: Sustainable software engineering, Software sustainability, Interactive tooling, Sustainability criteria, Green software, Decision support, Visual dashboards

1 Introduction

Software systems have increasingly become a vital part of modern infrastructure, significantly influencing economic growth, social interaction, and daily life [Ag23; KD11]. With the global acceleration of digital transformation and the rise of AI, the environmental impact associated with software systems - particularly their consumption of energy and resources - has gained significant attention in both academia and industry [La15; Na11; VSC23]. The intersection of software engineering and sustainability is addressed by the discipline of sustainable software engineering, which aims to integrate sustainability considerations into the software development life cycle to reduce environmental footprints while maintaining performance and functionality [Be15; Na11; Ve18].

Traditionally, software engineering prioritized functional correctness, performance optimization, and maintainability. However, increasing awareness of climate change and environmental degradation requires extending these priorities to include ecological sustainability [CP17]. Researchers and practitioners now recognize that software indirectly influences environmental sustainability through its operational resource demands, hardware utilization, and life-cycle management [Hi16]. Consequently, a systematic and rigorous approach to evaluating, managing, and reducing software-related environmental impacts has become essential.

[1] University of Applied Sciences Trier, Environmental Campus Birkenfeld, Institute for Software Systems, PO-Box 1380, 55761 Birkenfeld, Germany, a.abukhadijah@umwelt-campus.de; l.begic@umwelt-campus.de; a.guldner@umwelt-campus.de; Seb.Weber@umwelt-campus.de; s.bast@umwelt-campus.de; g.dartmann@umwelt-campus.de; s.naumann@umwelt-campus.de

doi:10.18420/env2025_05

Efforts such as the GREENSOFT reference model by Naumann et al. [Na11], the Green Software Measurement Model (GSMM) by Guldner et al. [Gu24], and the Sustainability Assessment Framework (SAF) Toolkit by Lago et al. [La25] offer foundational knowledge and practical frameworks for evaluating and improving software sustainability. Despite these advancements, many existing models and tools face limitations in terms of interactivity and contextual applicability, particularly regarding their integration into real-world development processes and decision-making workflows. For instance, GREENSOFT Naumann et al. [Na11] and GSMM Guldner et al. [Gu24] provide structured models and metrics but lack contextual guidance and do not support collaborative annotation. Users in our internal projects KIRA [2], EASY[3],Green Twin[4], as well as partners from industry and other sectors, have often pointed out that the large number of criteria, indicators, and metrics leads to cluttered presentations. This makes it difficult for them to quickly identify and apply the most relevant information in practice. These observations served as the motivation for developing a tool specifically designed to address this issue.

To address this gap, we introduce a tool that provides an interactive interface for navigating sustainability criteria catalogs introduced e.g. from Kern et al. [Ke18], or measurement procedures from Guldner et al. [Gu24]. Users can apply predefined or dynamic filters to explore relevant indicators based on project goals, domain contexts, or constraints. Furthermore, the tool includes a Markdown-based Wiki system for documenting use-cases and linking each to selected criteria along with their explanation. This approach supports transparency, reproducibility, and knowledge sharing in sustainability-oriented software development. The purpose of this project is to make the sustainability criteria catalog more accessible to companies, decision makers, and developers, potentially leading to valuable feedback for the scientific community and informing future research.

2 Related Work

Research in sustainable software engineering encompasses diverse tools, frameworks, and educational initiatives aimed at enhancing software sustainability through energy-efficient practices and life-cycle considerations.

Naumann et al. [Na11] introduced conceptual models and practical guidelines promoting green software development, emphasizing energy efficiency and life-cycle-oriented design. Their work significantly influenced eco-labeling standards such as the Blue Angel for Resources and Energy-Efficient Software Products (DE-UZ 215)[5], which stipulates strict criteria regarding energy and hardware usage, as well as transparency and update criteria.

[2] https:www.umwelt-campus.de/iss/projekte/laufende-projekte/kira [2025-06-18]
[3] https://www.umwelt-campus.de/iss/projekte/laufende-projekte/easy [2025-06-18]
[4] https://www.umwelt-campus.de/iss/projekte/laufende-projekte/greentwin [2025-06-18]
[5] https://www.blauer-engel.de/en/productworld/software [2025-06-11]

Other foundational concepts or frameworks such as the Sustainability Assessment Framework (SAF) Toolkit [La25] have been developed. SAF specifically enables software architects to explicitly model sustainability as a core quality attribute during the software design phase. However, it primarily addresses design-time considerations and does not account for runtime interactions or project-specific customization. Moreover, its usage is often limited to expert-driven modeling workflows and lacks interactive support for practitioners during real-world decision-making processes. In contrast, our tool bridges this gap by offering a lightweight, browser-based interface that integrates markdown-based contextualization, dynamic filtering, and visual dashboards—allowing both technical and non-technical users to explore, apply, and document sustainability criteria in concrete project settings. To complement these design-focused frameworks, various runtime measurement tools have emerged to provide quantitative insights into software sustainability during actual operation. CodeCarbon [Co24], for instance, estimates the carbon footprint of Python code execution by accounting for the execution environment and geographical context. Furthermore, it does not support contextual linking to project-specific sustainability goals or provide interactive visual analytics beyond post-hoc dashboards, which limits its utility in integrated decision-making environments. Similarly, the Green Metrics Tool [Gr24a] provides runtime energy profiling capabilities for applications, assisting developers in identifying and addressing energy-inefficient code segments. This tool emphasizes container-based runtime profiling but does not natively support linking energy measurements to contextual use-case descriptions, nor does it allow practitioners to document rationale or reference criteria directly within the tool. Additionally, GreenHub SDK [Gr24b] specifically targets energy monitoring for Android applications. Its focus is primarily on data acquisition and model training rather than on providing practitioners with interactive, project-specific tools.

Addressing sustainability in the context of AI, the GAISSA project [MFD23] emphasizes integrating sustainability into AI-based software system design. GAISSA offers architecture-centric methods and tools that assist data scientists and software engineers in creating energy-efficient AI applications, introducing the novel concept of "greenability"to evaluate and optimize AI systems' sustainability through informed architectural decisions.

Educational initiatives, such as the Sustainable Software Engineering course [De24], attempt to bridge theoretical concepts with practical applications. However, existing educational resources and research tools often lack interactive, use-case-specific exploration capabilities and systematic documentation frameworks, which are essential for effectively justifying and contextualizing sustainability decisions within practical projects.

Guldner et al. [Gu24] proposed the Green Software Measurement Model (GSMM), a comprehensive reference framework designed to structure the measurement and analysis of the energy and resource consumption of software. The GSMM synthesizes insights from various research groups into a structured model comprising five key components: measured objects and goals, metrics, procedural models, measurement setups, and data evaluation protocols. This model emphasizes detailed metadata collection, enhancing reproducibility, comparability, and extensibility of sustainability assessments. Complementing the GSMM, the structured criteria catalog introduced in Kern et al. [Ke18] (and updated in the process of the creation of the Blue Angel) offers a detailed collection of sustainability criteria derived from empirical software evaluations. Building on this foundation, the KIRA [6] project has extended the criteria catalog to include additional AI-specific sustainability aspects, reflecting recent developments and domain-specific challenges. These extended criteria are already applied in practical scenarios such as the determining the fill level of containers Cetkin et al. [Ce24]. Despite its comprehensiveness, the static nature of this catalog limits its usability due to the absence of interactive exploration features, dynamic filtering capabilities, and contextual documentation support.

To address these identified gaps, we developed the NEXUS Toolbox (Navigating Environmental Criteria (cross-criteria), Usage and Sustainability). This tool enhances existing sustainability criteria catalogs by providing an interactive environment for exploration, documentation, and visualization. It integrates embedded documentation features, enabling practitioners to effectively align their project-specific decisions with sustainability objectives through intuitive, context-aware interactions. In this way, the NEXUS Toolbox positions itself as a practice-oriented layer that contextualizes existing criteria and supports sustainability-related decision-making through concrete use-case guidance.

3 System Architecture

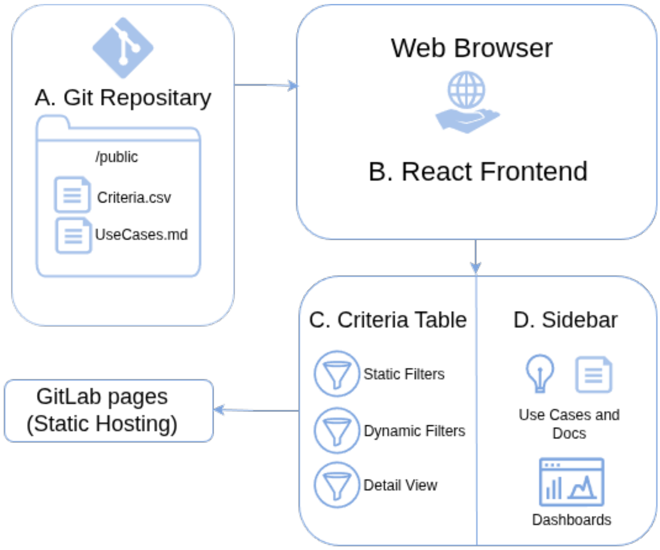


Fig. 1: System architecture

[6] https:www.umwelt-campus.de/iss/projekte/laufende-projekte/kira [2025-06-18]

The tool's structure follows a modular layout (see Fig. 1). It is implemented as a React-based frontend and deployed via GitLab Pages (Fig. 1 B.). A central component is a dynamic table (Fig. 1 C.) that displays all available criteria and provides sorting and filtering functionalities. Users can use static filters in the form of drop-down menus that allow for rapid selection of criteria by predefined categories such as *type of indicator*, *effort*, *domain* or *use case*. Additionally, dynamic filters enable more flexible querying: users can freely select any column as a filter and enter search terms to refine the dataset according to their needs.

Furthermore, the system can display details for each individual criterion. When a table entry is selected, a separate view displays the available information associated with that criterion, including its definition, context of use, and measurement details (Fig. 1 C.) . This design facilitates focused inspection of single evaluation elements without distraction from the broader dataset.

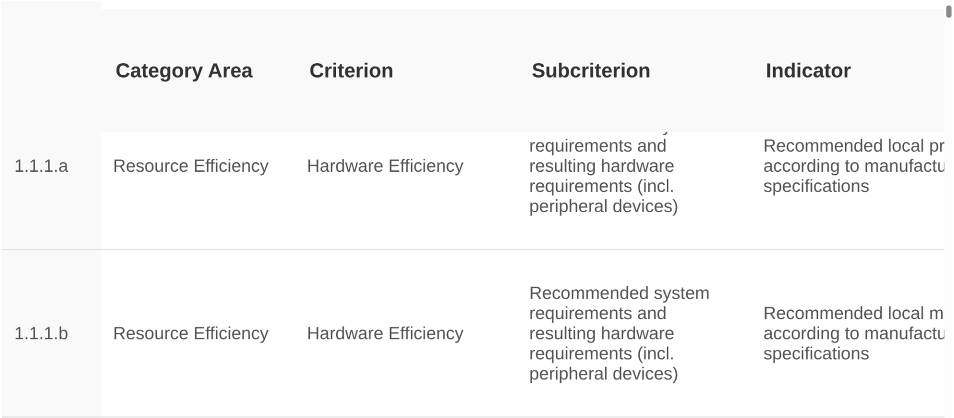
In addition to information on the criteria, the toolbox integrates a sidebar interface (Fig. 1 D.) that lists practical use-cases. These are written in Markdown and rendered directly within the application using a client-side Markdown parser. Each use-case outlines a scenario in which certain criteria were applied, providing context, rationale, and relevant background information.

The architecture also includes support for embedded visual dashboards within the Markdown pages (Fig. 1 D., bottom). These dashboards allow users to interactively generate diagrams based on selected metrics and models. This functionality enhances the tool's analytical capabilities and helps users identify patterns and relationships in the data. Since the tool is deployed via GitLab Pages, there is no backend, database, or server-side logic involved (Fig. 1 A.). Leveraging this fully client-side architecture, the NEXUS Toolbox runs entirely in the browser and requires no installation or additional infrastructure—only a modern web browser is needed, which ensures ease of deployment and broad accessibility.

4 Implementation

The front-end was implemented in JavaScript[7] using the React framework[8].The dynamic criteria table (see Fig. 2) is powered by the open-source library `react-table`[9], while Markdown rendering is handled via `react-markdown`[10]. All criteria data in CSV format and use-case documents in Markdown format are stored statically in the public directory of a GitLab repository. The application fetches this content at runtime using asynchronous requests.

[7] ECMAScript 2021 standard (ES12).
[8] React version 18.2.0, released June 2022.
[9] https://www.npmjs.com/package/react-table [2025-06-13]
[10] https://github.com/remarkjs/react-markdown [2025-06-13]



	Category Area	Criterion	Subcriterion	Indicator
1.1.1.a	Resource Efficiency	Hardware Efficiency	requirements and resulting hardware requirements (incl. peripheral devices)	Recommended local pr according to manufactu specifications
1.1.1.b	Resource Efficiency	Hardware Efficiency	Recommended system requirements and resulting hardware requirements (incl. peripheral devices)	Recommended local m according to manufactu specifications

Fig. 2: Table view of all criteria

To enable focused exploration of the criteria, the tool incorporates two filtering mechanisms. Static filters are implemented as drop-down menus that allow users to filter by predefined categories such as *indicator type*, *domain relevance*, *required effort* or *use case*. In addition, dynamic filters (see Fig. 3) allow users to interactively select any table column and enter free-text search queries to narrow down the displayed results based on their specific interests.

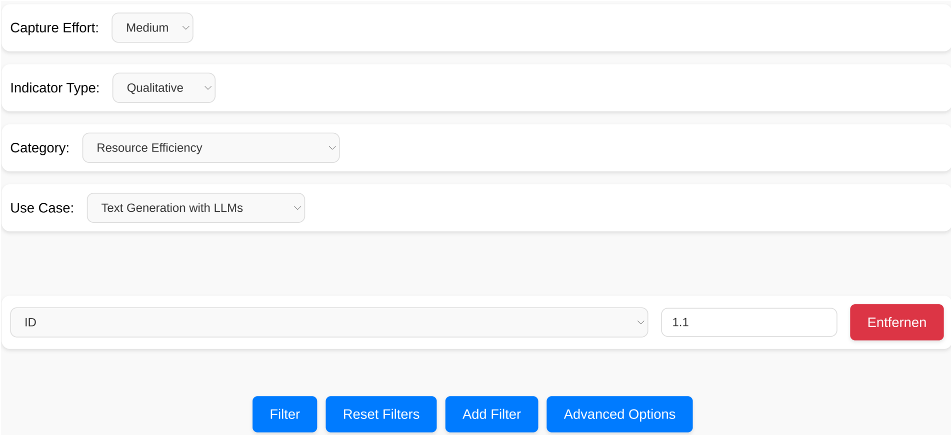


Fig. 3: Static and dynamic Filters

Moving from general criteria exploration to specific practical contexts, the tool incorporates a detailed use-case documentation. To effectively visualize the practical application of sustainability criteria, interactive dashboards are embedded directly within each use-case page, using the `recharts` [11] library. Each use-case page clearly outlines a specific scenario,

[11] https://recharts.org/en-US [2025-06-13]

detailing the context, explanation, and relevance of selected criteria. The visual dashboards within these pages generate interactive bar and line charts, enabling users to examine and interpret metrics, evaluate algorithmic performance, and compare models in an accessible and engaging manner. As an exemplar, the waste management use case estimates the fill level of municipal waste bins. By fusing spatial impulse-response analysis with machine-learning techniques, this sound-based approach remains low-cost and non-invasive, and the same pipeline can be transferred to any application that requires reliable level or occupancy detection far beyond waste-management scenarios. This direct integration of visual analytics within use-cases enhances clarity, supports informed decision-making and provides deeper insight into the practical impacts of sustainability criteria application (see Fig. 4).

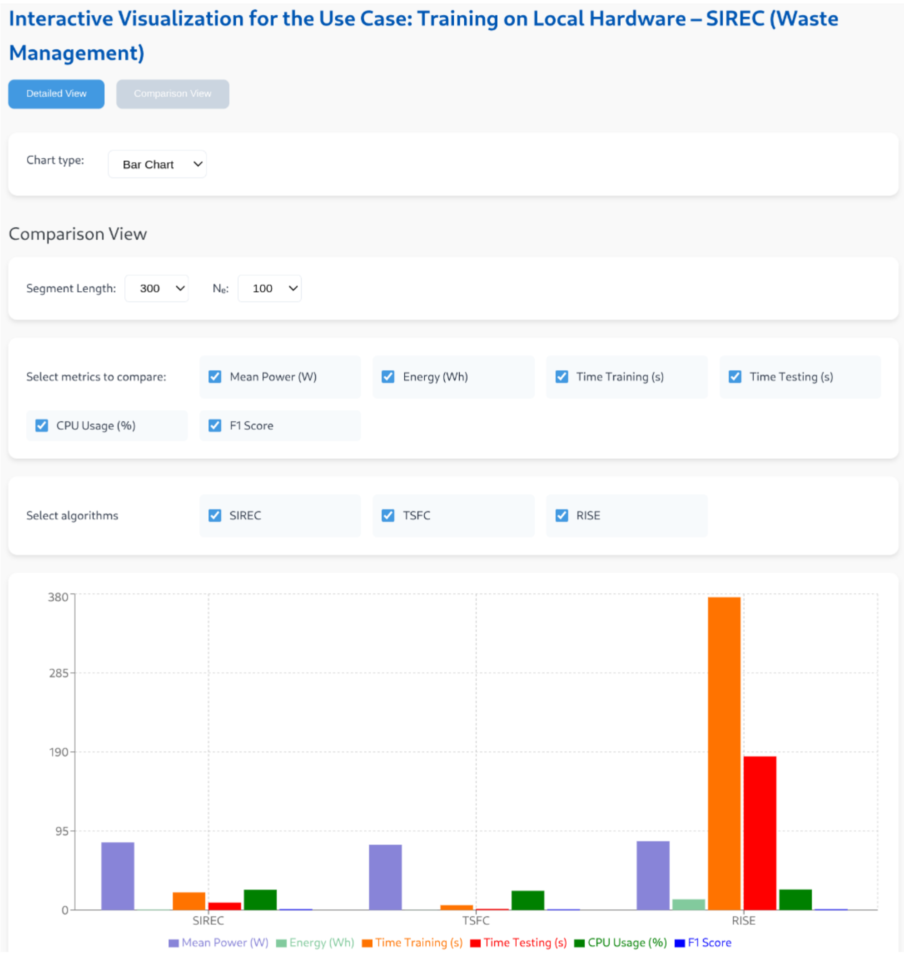


Fig. 4: Charts of a use-case

All filtering, navigation, and rendering are handled entirely within the browser. This approach ensures full transparency and version control through Git, and allows researchers and practitioners to easily extend or adapt the tool to their own needs. The tool is available as an open-source repository [12], with a live instance [13].

5 Discussion and Outlook

This tool represents an attempt to bridge the gap between comprehensive sustainability frameworks and their practical application in software projects. By offering a structured and interactive interface, it enables developers and decision-makers to explore the extensive sustainability criteria catalog proposed by Naumann, Guldner, Kern et al. [Gu24; Ke18; NGK21], which may otherwise appear too complex or abstract for practical use.

Beyond simply listing metrics, the tool provides context and explanations for selecting specific criteria in defined use-cases. This feature is essential for supporting transparent decision-making, as it allows users to understand not only what to measure but also why certain criteria are relevant in their domain or project setting. In doing so, the tool facilitates deeper reflection on sustainability goals and their operational relevance.

The tool also opens a channel for feedback between practitioners and researchers. Developers and decision-makers can engage with the criteria in realistic settings, and their experiences—ranging from applicability to limitations—can be systematically documented and used to refine the criteria catalog itself. This creates a valuable feedback loop. While the current version of the tool demonstrates the concept, it remains a work in progress. Future development will focus on extending functionality, improving usability, and integrating more diverse use-cases and practitioner feedback.

To prevent pseudo-dependencies between criteria filters, we plan to implement a filter-validation module. Before executing a query, selected filters will be checked against predefined rules (e.g., certain domains only apply to specific indicator types). If an invalid combination is detected, the user will be alerted with suggestions for valid filter options.

For the next phase, we plan to conduct a user evaluation to ensure the practical applicability and usability of the tool. We will recruit 10 experts from industry and academia, perform semi-structured interviews, and collect feedback using the System Usability Scale. The collected results will then be analyzed to refine the NEXUS Toolbox based on practitioner suggestions.

Additionally, to further enhance the practical applicability and personalization of the tool, a promising future direction involves integrating a recommender system. This system would suggest relevant sustainability criteria based on project-specific metadata. By analyzing

[12] https://gitlab.rlp.net/green-software-engineering/KIRA-Toolbox [2025-06-20]
[13] https://kira-toolbox-4afdf7.pages.gitlab.rlp.net/ [2025-06-17]

characteristics such as project domain, development phase, target platform, team size, or known sustainability goals, the recommender system could intelligently prioritize and propose the most relevant indicators and metrics from the criteria catalog. This approach would support users in rapidly identifying context-appropriate sustainability considerations, reduce cognitive load, and facilitate adoption in time-constrained development environments. Recommendation algorithms could range from simple rule-based mappings to more sophisticated machine learning models trained on annotated use-case data and historical project feedback.

Acknowledgement

This project was funded by the Federal Ministry for the Environment, Climate Action, Nature Conservation and Nuclear Safety (BMUKN) project "KIRA" under Grant 67KI32013. Parts of the text have been enhanced and linguistically revised using AI tools.

References

[Ag23] Agustian, K. et al.: The Impact of Digital Transformation on Business Models and Competitive Advantage. Technology and Society Perspectives (TACIT) 1 (2), pp. 79–93, 2023.

[Be15] Becker, C. et al.: The Karlskrona Manifesto for Sustainability Design. In: Proceedings of the 37th International Conference on Software Engineering - Volume 2. IEEE, pp. 467–476, 2015.

[Ce24] Cetkin, B. et al.: Spatial Impulse Response Analysis and Ensemble Learning for Efficient Precision Level Sensing. Discover Artificial Intelligence 4, p. 63, 2024.

[Co24] CodeCarbon Contributors: CodeCarbon: Track the Carbon Footprint of Your Computing, Accessed: 2025-05-15, 2024, https://codecarbon.io/.

[CP17] Calero, C.; Piattini, M.: Green in Software Engineering. Springer International Publishing, Cham, 2017.

[De24] Delft University of Technology: Sustainable Software Engineering Course, Accessed: 2025-05-15, 2024, https://luiscruz.github.io/course_sustainableSE/2024.

[Gr24a] Green Metrics Tool Contributors: Green Metrics Tool, Accessed: 2025-05-15, 2024, https://greenlab.di.uminho.pt/green-metrics-tool.

[Gr24b] GreenHub Contributors: GreenHub: Energy-Aware Android App Analysis Platform, Accessed: 2025-05-15, 2024, https://greenhub.ai/.

[Gu24] Guldner, A. et al.: Development and Evaluation of a Reference Measurement Model for Assessing the Resource and Energy Efficiency of Software Products and Components—Green Software Measurement Model (GSMM). Future Generation Computer Systems 155, p. 86, 2024.

[Hi16] Hindle, A.: Green Software Engineering: The Curse of Methodology. In: 2016 IEEE 23rd International Conference on Software Analysis, Evolution, and Reengineering (SANER). IEEE, pp. 46–55, 2016.

[KD11] Kitchin, R.; Dodge, M.: Software and Everyday Life. MIT Press, Cambridge, MA, 2011.

[Ke18] Kern, E. et al.: Sustainable software products—Towards assessment criteria for resource and energy efficiency. Future Gener. Comput. Syst. 86, pp. 199–210, 2018.

[La15] Lago, P. et al.: Exploring Initial Challenges for Green Software Engineering: Summary of the First GREENS Workshop. ACM SIGSOFT Software Engineering Notes 40 (6), pp. 36–39, 2015.

[La25] Lago, P. et al.: The Sustainability Assessment Framework Toolkit: A Decade of Modeling Experience. Software and Systems Modeling 24, Theme Section Paper, pp. 361–383, 2025.

[MFD23] Martínez-Fernández, S.; Franch, X.; Durán, F.: Towards Green AI-Based Software Systems: An Architecture-Centric Approach (GAISSA). In: 2023 49th Euromicro Conference on Software Engineering and Advanced Applications (SEAA). IEEE, Durres, Albania, pp. 411–418, 2023.

[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.

[NGK21] Naumann, S.; Guldner, A.; Kern, E.: The Eco-label Blue Angel for Software—Development and Components. In (Page, B. et al., eds.): Advances and New Trends in Environmental Informatics. Springer, Cham, pp. 55–67, 2021.

[Ve18] Venters, C. C. et al.: Software Sustainability: Research and Practice from a Software Architecture Viewpoint. Journal of Systems and Software 138, pp. 174–188, 2018.

[VSC23] Verdecchia, R.; Sallou, J.; Cruz, L.: A Systematic Review of Green AI. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 13 (4), e1507, 2023.

Hybrid Forecast of Phytoplankton in the California Current

Ida Hausmann [1] and Christine Preisach [1]

Abstract: Most of the oxygen breathed by humanity is produced by phytoplankton. Like many things on Earth, these organisms are severely affected by climate change. We apply Machine Learning (ML) to the long-term data set provided by CalCOFI to analyze the influence of different environmental parameters on phytoplankton. For forecasting of phytoplankton - up to the year 2100 for various climate change scenarios - we propose a hybrid approach that combines ML models and physical climate models (CMIP6 models). In our work, we show that the best ML model is able to estimate the phytoplankton population in the California Current with high accuracy. This model is then used in conjunction with the CMIP6 climate models, to predict the development of the organisms for the far future and different climate scenarios. Our hybrid approach is capable of integrating the results of complex climate models into a forecast without increasing the computational effort. It offers the possibility of predicting phytoplankton based on different scenarios without running simulations.

Keywords: Forecasting, Phytoplankton, Climate Change

1 Introduction

This research examines the influence of climate change on phytoplankton in the California Current. We analyze which environmental parameters have a large influence on the phytoplankton population. Afterwards, these parameters are used to train a regression model that estimates the phytoplankton population with high accuracy. Finally, this ML model is combined with physical climate models to predict phytoplankton up to 2100.

Industry, agriculture, transportation, living standards, land use, all these are anthropogenic activities that have a significant influence on climate change [De23]. It is essential to understand the consequences of climate change and therefore to be able to take appropriate measures to mitigate its effects [Sá21]. Scientists agree that the key to this is to research the oceans [Sá21; Wo21]. These large and complex ecosystems cover 71 percent of the Earth's surface and are home to the majority of all living organisms on the planet [Um09].

The regulation of climate, temperature, carbon dioxide and oxygen levels, as well as its use as a food source, are only a fraction of the functions of this ecosystem. However, the oceans are increasingly suffering from intensive use by humans and from the effects of climate change. In addition to global warming and rising sea levels, the oceans are also acidifying. This has serious consequences for humans as well as for all other species. [Um09]

[1] University of Applied Sciences Karlsruhe, ISRG, Moltkestr. 30, 76133 Karlsruhe, Germany,
ida.hausmann@h-ka.de, https://orcid.org/0009-0006-4331-9777;
christine.preisach@h-ka.de, https://orcid.org/0009-0009-1385-0585

doi:10.18420/env2025_06

Micro- and macroorganisms play a fundamental role in the oceans [Sá20]. This includes, for example, diatoms and flagellates, which are divided into zooplankton and phytoplankton [Um09]. Zooplankton refers to animal organisms that spend their entire life as plankton or are born only in the larval stage. Phytoplankton, on the other hand, refers to unicellular plant organisms. Through photosynthesis, these organisms produce most of the oxygen breathed by humanity [Na23]. For this reason, these species form the basis for a healthy and diverse life on Earth.

One of the four most important eastern boundary upwelling systems (EBUS) in the world is the California Current. EBUS are characterized by equatorward currents transporting cool, fresh and oxygen- and nutrient-rich water along the west coast of a continent. These wind-driven upwelling systems ensure that nutrients reach the surface, making these regions highly productive and very important for fisheries. In summary, the California Current constitutes a highly relevant ecosystem for the planet. Combined with the abundance of freely available data, research in this region contributes significantly to advancing our understanding of complex marine ecosystems and their responses to climate change. [Th22]

2 Background

Related Work Studying the oceanic systems is an essential part of climate research. Observing the marine ecosystem as well as its micro- and macroorganisms plays an important role in analyzing climate change [Wo21]. According to Sánchez-Pi et al. [Sá20], these organisms are key to understanding the complex functions of ecosystems and the risks posed by climate change. Meanwhile, scientists are using ML methods to explore the oceans and predict and mitigate the consequences of climate and biodiversity changes [Sá20].

Righetti et al. [Ri19] use an ML approach to make predictions about biographic patterns of phytoplankton species. The data used are based on 180 000 locations distributed globally. Additionally, different phytoplankton species are distinguished to investigate diversity. They tested three ML methods to train a classification model, which informs about the presence or absence of a species. Due to its high degree of interpretability and flexibility, they selected the generalized additive model (GAM). [Ri19]

In addition to machine learning methods, which primarily learn relationships within the training data to predict a target variable, environmental parameters can also be forecasted using physical climate models. These models simulate physical and chemical processes, as well as the dynamic interactions between different atmospheric and oceanic components, in order to produce projections for the 21st century and beyond or to reconstruct past states [KB24]. However, such simulations are highly computationally intensive [No18]. For instance, the computations for the CMIP6 models are largely carried out on high-performance computing systems, such as the Mistral supercomputer operated by the German Climate Computing Center [Gm]. Moreover, several years are typically required for model development and the completion of simulations [Ey16]. Therefore, it can be assumed that individual climate model simulations may take several weeks, months, or even years to complete.

Particularly for research questions at the local scale or for target variables that are not explicitly contained by climate models, a hybrid modeling approach can be advantageous. This concept is implemented, for example, in Nowack et al. [No18], where existing outputs from a climate model (e.g., temperature) are used as input features in a regression framework to estimate another parameter (e.g., ozone concentration).

A comparable approach is presented by Rivera et al. [Ri23], who not only investigate how plankton abundance can be modeled using machine learning methods, but also assess how this abundance may change under future climate conditions. First, a classification model is trained to determine the geographic distribution of zooplankton in the Humboldt current system by calculating the presence or absence of phytoplankton per pixel. Then, two Representative Concentration Pathways (RCP) scenarios are used to determine the values of the oceanographic variables for the periods 2040-2050 and 2090-2100. Lastly, the trained model can be projected onto the future with the help of these values. In contrast to [Ri23], our approach learns a regression model that predicts the exact chlorophyll-A concentration for the CalCOFI measuring stations.

In summary, previous work has mainly focused on classification approaches, which indicate only the presence or absence of phytoplankton, without providing information on its actual abundance. Many machine learning models simply extend past trends, often failing to capture unexpected shifts in climate dynamics, such as those driven by political decisions. Scenario-based analyses offer the advantage of examining whether phytoplankton is influenced by climate change or develops independently of it. While physical climate models simulate various environmental parameters and their interactions, they require significant computational resources. Therefore, there is a need for more efficient approaches that utilize existing data without relying on costly simulations. Future research should address the following questions: Which parameters influence phytoplankton abundance? How is it linked to climate change? And what forecasts can be made under different SSP scenarios up to 2100?

Data The data for this research is provided by CalCOFI[2], this marine ecosystem research program is one of the longest-running sampling programs in the world. In general, it studies the physics, biochemistry and biology of the marine environment. Data collection started in the late 1940s and was initiated to investigate relationships between oceanographic conditions and sardine abundance. These goals and the volume of data associated with them quickly expanded. The data set contains 113 defined stations distributed along the California coast from the northern part of San Francisco Bay to San Diego on a uniform grid, with stations placed near the coast to approximately 670 km offshore (Fig. 1). The research ships collect samples from each station for 2 to 4 weeks every quarter of the year

[2] California Cooperative Oceanic Fisheries Investigation: https://calcofi.org/

using Conductivity-Temperature-Depth (CTD) rosettes at different depths. Thus, profiles are created from the sea surface to the bottom or to a maximum depth of 515 m. [Sa21]

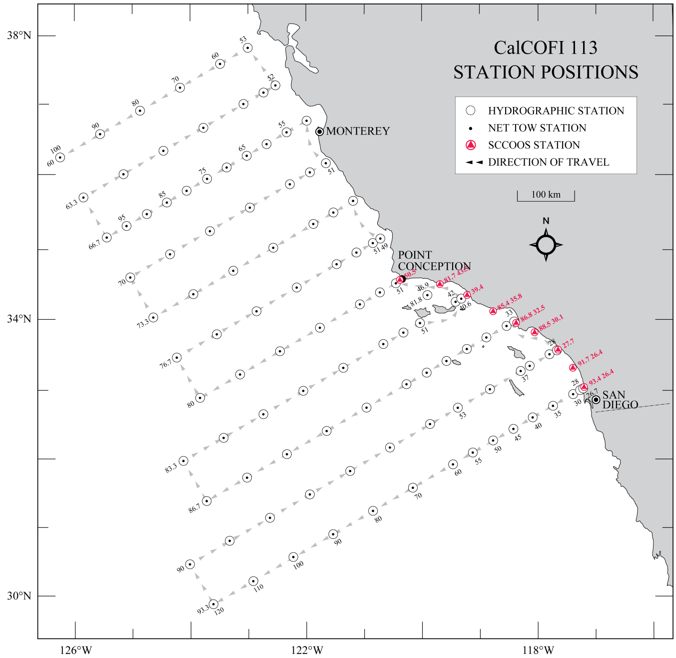


Fig. 1: Map of the CalCOFI stations.[3] [Ca23]

In addition to the oceanographic data used to train the ML model, we used estimations from climate models, including various Shared Socioeconomic Pathways (SSP) scenarios to be able to predict into the future. For this task, we chose a high-quality data set provided by Bio-ORACLE v3.0[4]. It contains comprehensive raster data sets on the sea surface and benthos [Bi24]. These data include geophysical, biotic and ecological information [Bi24] for the present as well as for the future, based on CMIP6 models [As24]. The data products of the sixth phase of CMIP are considered an important source of robust and reliable climate information [Gm].

The Bio-ORACLE layers are global data sets that can be used for both large- and small-scale research due to its high spatial resolution. The spatial resolution is displayed in 5 arcmin, which corresponds to approximately 9.2 km at the equator [As18]. The calculation of the Bio-ORACLE layer is based on averaging the results of the different CMIP6 models [As18]. All the data used in our research therefore use the same combination of climate models.

[3] SCCOOS = Southern California Coastal Ocean Observation System

[4] Bio-Oracle: https://www.bio-oracle.org/index.php

3 Methodology

Our hybrid approach consists of two main parts (Fig. 2): First, an explainable regression model is trained using the CalCOFI data, then in the second part, this model is used to create a forecast for phytoplankton occurrence for the years 2030 to 2100.

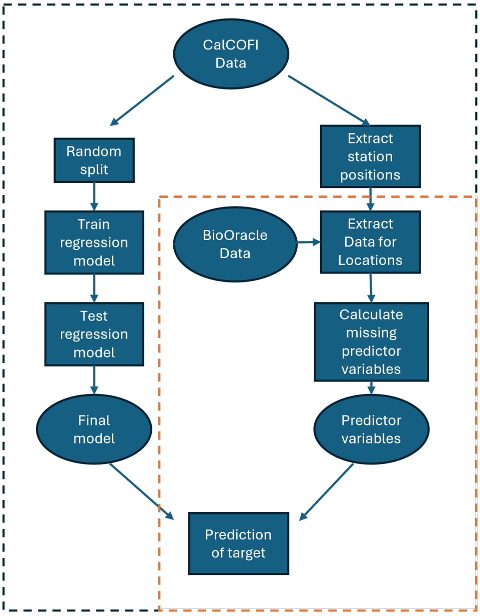


Fig. 2: Flowchart of the Workflow.

Chlorophyll-A was selected as the target variable for the regression model. Chlorophyll-A is a pigment found in all phytoplankton species [Da18] and is widely used as a standard proxy to estimate phytoplankton biomass [Je80]. The unit of chlorophyll-A is µg/l. An initial examination of the dataset revealed that chlorophyll-A concentrations exhibit a highly skewed distribution, which complicates both the interpretation of temporal trends and the analysis of relationships between chlorophyll-A and environmental variables. In cases where the data approximate a log-normal distribution, as observed here, a commonly applied approach is to perform a logarithmic transformation [Fe14; Ze21]. This transformation adjusts the relative differences between values, thus modifying the distributional characteristics of the variable [Ze21]. The transformation is applied using the formula: $x' = \log(x + 0.001)$

To ensure a good interpretation, only algorithms that guarantee a high level of interpretability were tested. This means that the flexibility of these algorithms is limited and, consequently, no complex models are generated [Ja13]. For this reason, linear regression (LR), generalized additive models (GAM) and random forest (RF) are tested. LR is considered the simplest and fastest approach, giving the user a good overview of the parameters that have the strongest influence on the target variable [Ja13]. Furthermore, the LR can be used as a baseline. GAM allows non-linear relationships to be represented, which often increases the accuracy of the predictions and the results are still easy to interpret [Ja13]. Furthermore, this method is used in some comparable works, such as that of Rivera et al. [Ri23]. RF combines a high degree of flexibility and still provides some interpretability due to feature importance (FI) scores [Ja13]. All future values of environmental parameters fall within the range of the training data, validating the applicability of the RF model for projection.

The models were trained and evaluated using an 80/20 train-test split. To compare the performance of the different algorithms, evaluation metrics such as R^2, root mean squared error (RMSE) and mean absolute error (MAE) were calculated (table 1). Among the tested models, RF achieved the best performance across all evaluation metrics. Consequently, the RF model was selected for the prediction of the occurrence of phytoplankton.

Tab. 1: Comparison R^2, RMSE and MAE of the LR-, GAM- and RF-Model based on the same training data.

Error Metrics	LR	GAM	RF
R^2	0.7696	0.8390	**0.9242**
RMSE	0.9037	0.7651	**0.5151**
MAE	0.6766	0.5684	**0.3264**

A correlation analysis was carried out before selecting the input parameters. This showed that there are few parameters that exhibit collinearity. Examples of these are the parameters density, water temperature and sampling depth. In general, RF is robust against correlations between features, and forecasts are not affected by this. For this reason, it was experimentally tested which input characteristics lead to the lowest error rates (the same evaluation metrics as in table 1). The input parameters selected are listed in Table 2.

The aim of the second part of our approach is to predict phytoplankton for different climate change scenarios as well as for different time periods in the future (2030-2100). The workflow for this part is visualized in the orange section in fig. 2. The basis for this are the Bio-Oracle data, which contain global raster data sets for various environmental parameters. Based on the geographical position of the measuring stations, the corresponding values for different future points in time can be extracted from the raster data sets and converted into tabular format. The format of this must correspond to that of the regression model's training data. Unfortunately, not all environmental parameters used in the regression model are covered by Bio-Oracle, requiring the missing parameters to be estimated for future points in time. Depending on the parameter, the mean or median (from the CalCOFI data) is used. As only a small subset of parameters was affected (nitrite, density & wind conditions), they were not excluded from the analysis to assess their relevance for chlorophyll-A and thus improve the interpretability of the prediction quality. Once the data set is generated, the final regression model can be applied to create the long-term forecast, i.e., predict the chlorophyll-A value until 2100. In order to interpret the model predictions, they must first be transformed back, as they are still subject to the logarithmic transformation.

4 Results and Discussion

The use of the explainable ML method RF enables the interpretation of the influence of individual environmental parameters on the target variable by calculating the feature importance (FI) scores (Tab. 2). These indicate that sampling depth, water density and water temperature have a significant influence on the abundance of phytoplankton (FI 29%-13%). Nutrients such as nitrate, silicate and phosphate as well as parameters related to position (distance to the coast and bottom depth) have a medium influence on phytoplankton with a FI <10%. Atmospheric parameters such as air temperature, air pressure and wind conditions have a FI of <2% and therefore do not have a major influence.

Tab. 2: Feature Importance score for all environmental parameters of the final regression model.

Feature Importance of environmental parameters			
Parameter	FI (%)	Parameter	FI (%)
Sampling depth	29.36	Phosphate	3.48
Water temperature	18.13	Salinity	2.14
Density	13.96	Air temperature	1.32
Nitrate	7.49	Quarter	1.11
Silicate	6.49	Pressure	0.99
Nitrite	5.69	Wind speed	0.78
Distance	4.66	Wind direction	0.70
Bottom depth	3.70		

The focus of this research was to compare different climate change scenarios to analyze whether the development of phytoplankton is affected by climate change. When inspecting the chlorophyll-A predictions over time and stations (Fig. 3) , we notice that each scenario leads to a reduction in phytoplankton occurrence . This trend is already evident in past data measured by CalCOFI. However, the extent of the loss varies greatly depending on the scenario considered. The sustainable SSP-scenarios SSP1-1.9 and SSP1-2.6 show a loss up to 2040 and change only slightly thereafter. Predictions of Chlorophyll-A concentrations decrease more steeply for scenarios that expect a high increase in CO_2 emissions until 2100. It can be seen that the higher the expected CO_2 increase, the lower the predicted chlorophyll-A concentration in 2100.

The main reason for the lower concentration of chlorophyll-A is the increase in water temperature. As already mentioned, the water temperature correlates significantly with the chlorophyll-A concentration. Bio-Oracle data show that the increase in CO_2 will also have a significant impact on the water temperature. Fig. 4 shows the predicted water temperature for different SSP scenarios as modeled by Bio-Oracle, which also shows an upward trend for nonsustainable scenarios. An increase in the water temperature leads to a higher grazing rate of zooplankton, exceeding the production of phytoplankton [Ol13]. Furthermore, warming leads to a reduction in the supply of nutrients. There is no exchange of water layers, which means that nutrients from the depths no longer reach the surface [Ho11].

These developments have a significant impact on this ecosystem and humanity. A reduced phytoplankton population leads to a loss of ecological resilience [He21], as well as to a negative development in the productivity of the marine ecosystem [He21; Ho11; Ol13]. This reduction in productivity affects fisheries, meaning that an important food source disappears [Fr21]. Phytoplankton also plays a crucial role in the global carbon cycle. If phytoplankton declines, the ocean loses its ability to absorb CO_2, which in turn further exacerbates climate change [Ho11].

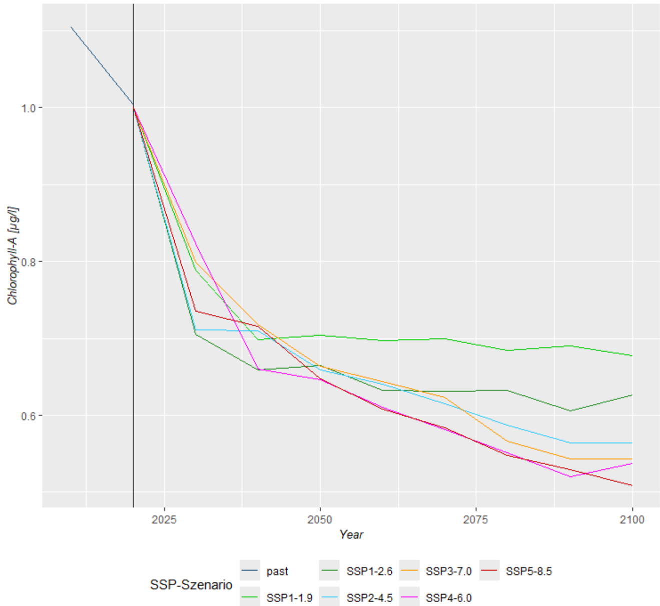


Fig. 3: Phytoplankton (chlorophyl-A) forecast

When reviewing the quality and plausibility of these results, two aspects must be considered. First, the quality of the regression model's ability to depict the relationship between various environmental parameters and the target variable is crucial. In our work, the RF model achieved an R^2 of 0.92, indicating good quality. Secondly, the Bio-Oracle data only contain predicted values, meaning an unknown level of uncertainty is included in the calculation. However, as Bio-Oracle combines several complex climate models from CMIP6 [As18], it can be assumed to be of suitable quality for further processing.

This hybrid forecasting approach offers several advantages over using either ML or physical climate models alone. Using only ML could lead to very inaccurate long-term predictions, as these would only be based on the distant past. However, the climate is constantly changing, which can lead to changes in the trend or unexpected fluctuations. Furthermore, various political decisions, as well as other events, can significantly influence climate development. Therefore, it makes sense to integrate different scenarios, like the SSP scenarios, into the analysis when creating such forecasts.

The simulation of such scenarios is usually carried out with atmosphere-ocean general circulation models (AOGCM). Atmospheric and oceanic models are combined to simulate the dynamics and interactions between the systems [KB24]. However, the calculations of these models are extremely complex and require a great amount of computing power and

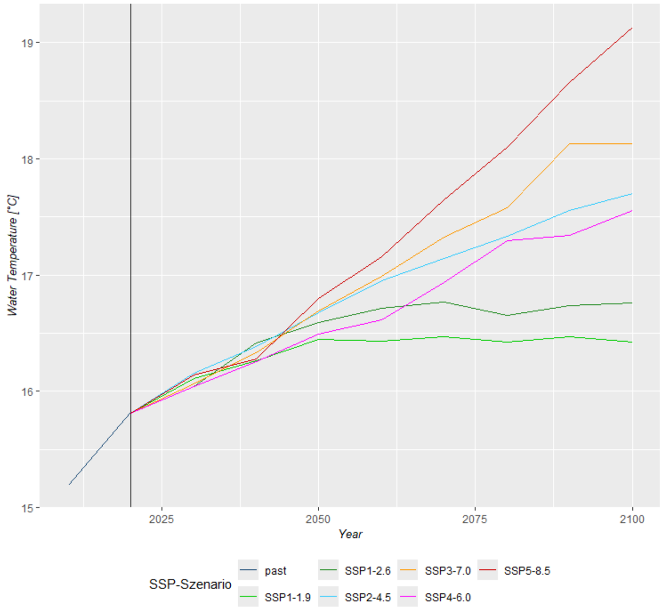


Fig. 4: Water temperature forecast

time [KB24]. In contrast, our approach uses fast ML methods to analyze relationships between environmental parameters and phytoplankton. The values already calculated by various AOGCMs can then be used as input for the forecasts. This allows the heterogeneous correlations of the simulations to be integrated into the prediction without having to execute complex calculations. This approach is therefore ideal to create what-if scenarios.

5 Conclusion

This research shows that explainable ML algorithms perform well in representing the correlation between different environmental parameters and phytoplankton. In particular, the RF approach achieves more accurate prediction results than GAM, which is typically used in this area.

To summarize the results, all SSP-scenarios lead to a reduction in phytoplankton occurrence. However the extent of the loss varies greatly depending on the scenario. It can be noted, that the higher the expected CO2 increase, the lower the predicted chlorophyll-A concentration in 2100. This variation in predictions depending on the scenarios supports the hypothesis that phytoplankton could be affected by climate change. According to the feature importance of the regression model, the main factors responsible for phytoplankton development are water temperature, sampling depth, and density. Furthermore, the influence of nutrients and geographical conditions can also be observed.

Overall, our hybrid approach is able to integrate results of physical climate models into a forecast without increasing the computational effort or calculation time. Furthermore, this approach offers the possibility to calculate the development of phytoplankton depending on different scenarios without running simulations. In this paper, the analysis is focused on the average development over time in the different scenarios, future research could explore spatial differences in phytoplankton occurrence and development.

References

[As18] Assis, J. et al.: Bio-Oracle v2.0: Extending marine data layers for bioclimatic modelling. Global Ecology and Biogeography 27 (3), pp. 277–284, 2018.

[As24] Assis, J. et al.: Bio-Oracle v3.0. Pushing marine data layers to the CMIP6 Earth System Models of climate change research. Global Ecology and Biogeography 33 (4), 2024.

[Bi24] Bio-Oracle: Bio-Oracle : Marine data layers for ecological modelling, bio-oracle, 2024, https://bio-oracle.org/index.php.

[Ca23] CalCOFI: Station Positions, California Cooperative Oceanic Fisheries Investigations, 2023, https://calcofi.org/sampling-info/station-positions/.

[Da18] Davies, C. H. et al.: A database of chlorophyll a in Australian waters. Scientific Data 5 (1), 2018.

[De23] Deutscher Wetterdienst: Wetter und Klima - Deutscher Wetterdienst - Basisfakten zum Klimawandel, Deutscher Wetterdienst, 2023, https://www.dwd.de/DE/klimaumwelt/%20klimawandel/klimawandel%5C_node.html.

[Ey16] Eyring, V. et al.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geoscientific Model Development 9 (5), 2016.

[Fe14] Feng, C. et al.: Log-transformation and its implications for data analysis. Shanghai Archives of Psychiatry 26 (2), 2014.

[Fr21] Friedland, K. D. et al.: Trends in phytoplankton communities within large marine ecosystems diverge from the global ocean. Canadian Journal of Fisheries and Aquatic Sciences 78 (11), 2021.

[Gm] GmbH, D. K.: CMIP6, DKRZ, https://www.dkrz.de/de/kommunikation/klimasimulationen/cmip6-de, accessed: 12/13/2024.

[He21] Henson, S. A. et al.: Future phytoplankton diversity in a changing climate. Nature Communications 12 (1), 2021.

[Ho11] Hofmann, M. et al.: Declining ocean chlorophyll under unabated anthropogenic CO2 emissions. Environmental Research Letters 6 (3), 2011.

[Ja13] James, G. et al.: An Introduction to Statistical Learning. Springer NY, New York, NY, 2013.

[Je80] Jeffrey, S. W.: Algal Pigment Systems. In (Falkowski, P. G., ed.): Primary Productivity in the Sea. Springer US, 1980.

[KB24] Kasang, D.; Böttinger, M.: Komplexe Klimamodelle, Klimanavigator, 2024, https://www.klimanavigator.eu/dossier/artikel/012808/index.php.

[Na23] National Geographic Society: Marine Food Chain, National Geographic, 2023, https://education.nationalgeographic.org/resource/marine-food-chain.

[No18] Nowack, P. et al.: Using machine learning to build temperature-based ozone parameterizations for climate sensitivity simulations. Environmental Research Letters 13 (10), 2018.

[Ol13] Olonscheck, D. et al.: Decomposing the effects of ocean warming on chlorophyll *a* concentrations into physically and biologically driven contributions. Environmental Research Letters 8 (1), 2013.

[Ri19] Righetti, D. et al.: Global pattern of phytoplankton diversity driven by temperature and environmental variability. Science Advances 5 (5), 2019.

[Ri23] Rivera, R. et al.: Modeling present and future distribution of plankton populations in a coastal upwelling zone: the copepod Calanus chilensis as a study case. Scientific Reports 13 (1), 2023.

[Sá20] Sánchez-Pi, N. et al.: Artificial Intelligence, Machine Learning and Modeling for Understanding the Oceans and Climate Change. In: Tackling Climate Change with Machine Learning workshop at NeurIPS 2020. 2020.

[Sa21] Satterthwaite, E. V. et al.: Long-Term Marine Ecosystem Observations to Support Sustainability & Resilience in the California Current System,: California Cooperative Oceanic Fisheries Investigations (CalCOFI) 2017 - 2021, 2021.

[Sá21] Sánchez-Pi, N. et al.: OcéanIA: AI, Data, and Models for Understanding the Ocean and Climate Change. Inria – Institut national de recherche en sciences et technologies du numérique, Lille, Paris, Saclay, Santiago, Sophia-Antipolis, 2021.

[Th22] Thompson, A. R. et al.: State of the California Current Ecosystem in 2021: Winter is coming? Frontiers in Marine Science 9, 2022.

[Um09] Umweltbundesamt: Klimawandel und marine Ökosysteme - Meeresschutz ist Klimaschutz, 2009, https://www.umweltbundesamt.de/sites/default/files/%20medien/publikation/%20long/3805.pdf.

[Wo21] Wolff, T. d. et al.: Assessing Physics Informed Neural Networks in Ocean Modelling and Climate Change Applications. In: AI: Modeling Oceans and Climate Change Workshop at ICLR 2021. 2021.

[Ze21] Zelený, D.: Data Preparation, Analysis of community ecology data in R, 2021, https://www.davidzeleny.net/anadat-r/doku.php/en:data_preparation.

A Self-Supervised Vision Foundation Model for All-Sky Imagery

Preliminary Validation of a Multi-View Self-Supervised Learning Framework

Yassine Ribouh [1], Grit Behrens [2], Andreas Boschert [3] , Mike Zehner [4]

Abstract: Accurate short-term forecasting of solar irradiance is essential for maintaining the stability of renewable energy grids, yet it remains a difficult task due to the complex and dynamic nature of cloud formations. In this work, we present a vision-based foundation model built using self-supervised learning (SSL) on a large dataset of unlabeled all-sky imagery. Our hybrid SSL framework employs a Vision Transformer (ViT) and integrates masked autoencoding (MAE) to capture fine-grained atmospheric textures, alongside self-distillation to enhance semantic feature learning. To enforce robustness across multiple perspectives, we incorporate a cross-view consistency loss using data from four synchronized camera viewpoints. Early training results show strong convergence, underscoring the effectiveness of our approach as a high-fidelity feature extractor for downstream tasks in solar forecasting.

Keywords: Self-Supervised Learning, Vision Transformer, All-Sky Imagery, Foundation Models, Multi-View Learning, Masked Autoencoding.

1 Introduction

The large-scale integration of solar power is fundamentally challenged by grid instability, a direct result of its inherent intermittency. This variability, driven by dynamic cloud cover, demands accurate short-term irradiance nowcasting to keep the grid reliable. Ground-based all-sky imagers (ASIs) are well suited to the task because they offer finer temporal and spatial resolution than satellites and can capture the local cloud features that

[1] Solar Computing Lab, Bielefeld University of Applied Sciences and Arts, Artilleriestraße 9, 32427 Minden, yassine.ribouh@hsbi.de
[2] Solar Computing Lab, Bielefeld University of Applied Sciences and Arts, Artilleriestraße 9, 32427 Minden, grit.behrens@hsbi.de, https://orcid.org/0009-0009-0247-8204
[3] Laboratory for Solar Energy Systems, Rosenheim Technical University of Applied Sciences, Hochschulstraße 1, 83024 Rosenheim, Germany, https://orcid.org/0000-0002-1540-8098
[4] Laboratory for Solar Energy Systems, Rosenheim Technical University of Applied Sciences, Hochschulstraße 1, 83024 Rosenheim, Germany, https://orcid.org/0000-0001-9642-0496

doi:10.18420/env2025_07

matter for plant-level forecasts.

The key problem is turning these high-frequency image streams into physically meaningful feature vectors. Generic computer-vision backbones pre-trained on ImageNet learn edges and object parts that have little to do with atmospheric optics, leaving a semantic gap between the model and the task.

Only a few studies have tried to close this gap with self-supervised learning (SSL) on ASI data. *Johnson et al.* adapted SimCLR to auroral all-sky images and showed that contrastive pre-training boosts aurora-type classification, but the method relies on single-view crops and ignores generative signals [Jo21]. *Fabel et al.* used DeepCluster plus inpainting/super-resolution pre-text tasks on single-camera sky frames to improve four-class cloud segmentation, but they did not exploit multi-camera geometry or joint objectives [Fa22]. Both papers confirm that SSL helps, yet neither tackles view-invariance or combines contrastive and reconstruction losses in one pass.

In this work-in-progress we develop a Vision Transformer (ViT) based foundation model for all-sky imagery. The encoder is pre-trained on a large set of unlabeled sky frames with a hybrid SSL recipe: masked-image reconstruction, DINO-style self-distillation, and a lightweight cross-view consistency loss that ties together synchronized fisheye cameras in our dataset. We hypothesize and early results support that this blend yields a feature extractor that is more attuned to cloud physics than models trained from scratch or with generic pre-training. The remainder of the paper describes the method and reports preliminary gains in downstream nowcasting accuracy.

2 Methodology

2.1 Dataset and Pre-processing

Figure 1: ASI-16/55 with radiation shield in the middle, left the pyranometer and right the temperature and humidity sensor

The dataset for this study was collected at a dedicated site in Buttenwiesen, Germany. The setup includes four ASI 16 all-sky imagers from CMS Ing. Dr. Schreder GmbH [Sc25], each paired with a pyranometer and temperature and humidity sensors, enabling synchronized multi-view captures of the sky along with environmental metadata like solar irradiance, air temperature, and relative humidity. Data collection occurred over several discontinuous periods August through October 2024, briefly in November 2024, and again in April and May 2025 interrupted intermittently for system recalibration. Images were originally captured at 15-second intervals and then thinned to one-minute spacing to reduce redundancy while retaining meaningful temporal dynamics. For the self-supervised pre-training, the entire dataset was pooled and randomly split, with 98% (146,932 images) used for training and 2% (7,733 images) for validation. While sensor metadata was collected, it was intentionally excluded during SSL training to ensure that representations are learned purely from the image content.

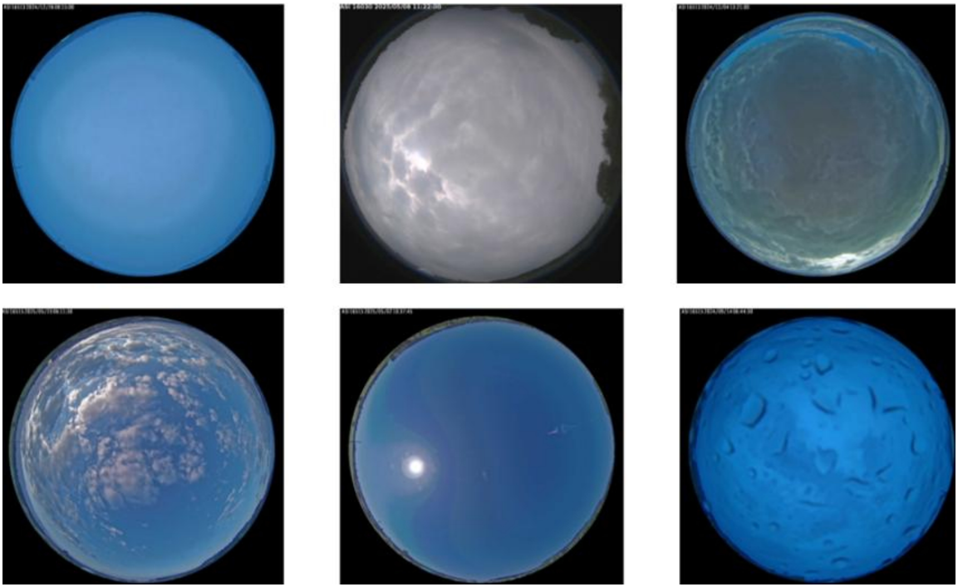


Figure 2: Examples of All-Sky Images with different cloud conditions.

2.2 A Multi-View Hybrid Self-Supervised Architecture

To create a feature extractor that is deeply specialized for the atmospheric domain, we designed a hybrid self-supervised learning framework. This framework is built upon a Vision Transformer (ViT) backbone and combines three distinct but complementary learning objectives: pixel-level reconstruction, high-level feature distillation, and multi-view consistency. The total loss for the model is a weighted sum of the individual loss components:

$$\mathcal{L}_{total} = \lambda_{pixel}\mathcal{L}_{Mae} + \lambda_{distill}\mathcal{L}_{Dino} + \lambda_{xview}\mathcal{L}_{xview}$$

where λ_{pixel}, $\lambda_{distill}$, and λ_{xview} are scalar weights that balance the contribution of each objective.

Backbone: Vision Transformer (ViT)

We selected the Vision Transformer (ViT) architecture as the model's encoder backbone [Do20]. Unlike convolutional neural networks that process images with localized filters, ViT partition an image into a sequence of patches, embed them, and process them using a global self-attention mechanism. This allows the model to capture long-range dependencies across the entire image from the earliest layers. For all-sky imagery, this global context is paramount, as the significance of a cloud in one part of the sky is often dependent on the conditions of the entire sky dome, such as the position of the sun.

Masked Autoencoding (MAE)

The first objective is based on Masked Autoencoding (MAE), a self-supervised method that excels at learning rich, low-level feature representations [He22]. In this approach, a significant portion of the input image patches (e.g., 75%) are randomly masked and removed before being fed to the ViT encoder. The model is then tasked with reconstructing the original pixel values of these missing patches from the latent representation of the visible patches. This task forces the model to develop an internal understanding of the "visual grammar" of the sky learning about typical cloud textures, colour gradients, and the shape of solar glare. The MAE loss is the Mean Squared Error (MSE) computed only on the masked patches:

$$\mathcal{L}_{mae} = \frac{1}{|M|} \sum_{p \in M} ||\mathrm{Decoder}\big(\mathrm{Encoder}(I_{\mathrm{vis}})\big)_p - \big(I_{\mathrm{orig}}\big)_p ||^2$$

where I_{vis} is the original image, I_{orig} are the visible patches, and M is the set of masked patches.

Self-Distillation (DINO-style)

While MAE is powerful for learning local texture, it can be less effective at capturing high-level semantic concepts. To address this, we integrated a self-distillation objective inspired by DINO and DINOv2 (self **DI**stillation, **NO** labels) [Ca21, Oq23]. This method uses a student-teacher framework, where the student model (the one being actively trained) learns to match the feature-level output of a teacher model. The teacher model's weights are not updated by backpropagation but are instead an exponential moving average (EMA) of the student's weights from previous training steps.

The student is fed different augmented views of an image and is trained to match the probability distribution produced by the more stable teacher from its own [CLS] (classification) token. This token is prepended to the image patch sequence to act as a summary of the entire image. The objective is formulated as a cross-entropy loss:

$$\mathcal{L}_{\mathrm{Dino}} = -\sum_{x \in V} P_t(x) \log P_s\,(x)$$

V is the set of augmented views of an input image. $P_s(x)$ is the probability distribution for view x produced by the student network, obtained by applying a temperature-scaled SoftMax function to its [CLS] token embedding. $P_t(x)$ is the target distribution from the teacher network, which is similarly computed but with a different temperature and is sharpened and centered to provide a stable, non-collapsing learning target. This process discourages the model from focusing only on trivial details and pushes it to learn robust, abstract features that are invariant to changes in color, brightness, or orientation.

Multi-View Consistency Loss

The key aspect of our architecture is a cross-view consistency loss designed to explicitly leverage our multi-camera dataset. Because we have four cameras capturing the same scene at the same time from different viewpoints, we can enforce a powerful constraint: the model's high-level understanding of the scene should be independent of the camera from which it was viewed.

To implement this, we augment the input to the ViT. For each image patch sequence, we prepend not only a standard [CLS] token but also a learnable "view embedding" that corresponds to the camera ID (1, 2, 3, or 4). The model is then trained on batches containing simultaneous images from different cameras. The cross-view loss function penalizes the model if the final [CLS] token outputs for these simultaneous views are dissimilar. We use a simple MSE loss for this objective:

$$\mathcal{L}_{xview} = ||z_{\text{cls}}^{(i)} - z_{\text{cls}}^{(j)}||^2$$

where $z_{\text{cls}}^{(i)}$ and $z_{\text{cls}}^{(j)}$ are the [CLS] token embeddings for two images of the same scene at the same time, captured from different camera views i and j. This loss forces the model to learn a truly view-invariant representation of the atmosphere, focusing on the inherent physical state rather than camera-specific artifacts or perspectives. This is a critical step towards building a generalized and robust feature extractor.

2.3 Training Protocol

The model is being trained for a total of 400 epochs. As of this writing, training is in progress and has completed 250 epochs on an NVIDIA A100 80 GB GPU, with each epoch taking approximately 20 minutes. We employ the AdamW optimizer [Lo17], configured with a base learning rate of 1.5×10^{-4} and a weight decay of 0.05. The learning rate follows a cosine decay schedule with a linear warmup over the first 10 epochs, allowing the model to stabilize its initial learning before gradually decreasing the rate to converge on a robust solution. Training uses a global batch size of 512 images.

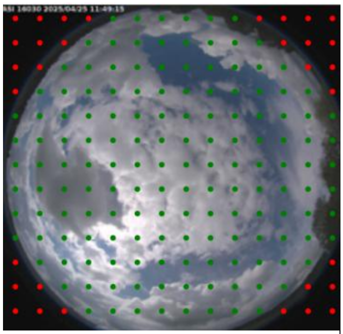


Figure 3: Fisheye mask: valid (green) vs. masked (red) patches

A key implementation detail is the use of a fisheye-aware mask. Since all-sky images have a circular data region surrounded by black, uninformative borders, we apply a static mask during pre-processing. This mask prevents the model from wasting computational resources attempting to reconstruct or learn features from these black pixels, focusing its capacity entirely on the relevant atmospheric data. Furthermore, to stabilize the hybrid learning process, we use a scheduling mechanism for two key hyperparameters. The momentum for the DINO teacher's weight update (an exponential moving average) gradually increased from 0.996 to 0.9995 over the first 30 epochs.

3 Preliminary Results and Analysis

To assess the effectiveness of our self-supervised pre-training strategy at this intermediate stage, we conducted an analysis of the feature space being learned by the ViT encoder.

3.1 Emerging Structure in the Feature Space

We extracted [CLS] embeddings for each validation image and applied t-SNE to map this high-dimensional space into two dimensions. In Figure 4, points colored by withheld solar irradiance form a clear, structured manifold rather than a random scatter. Low-irradiance (purple) and high-irradiance (yellow/green) regions separate smoothly, showing that our SSL training inherently clusters images by atmospheric conditions and captures irradiance-related features.

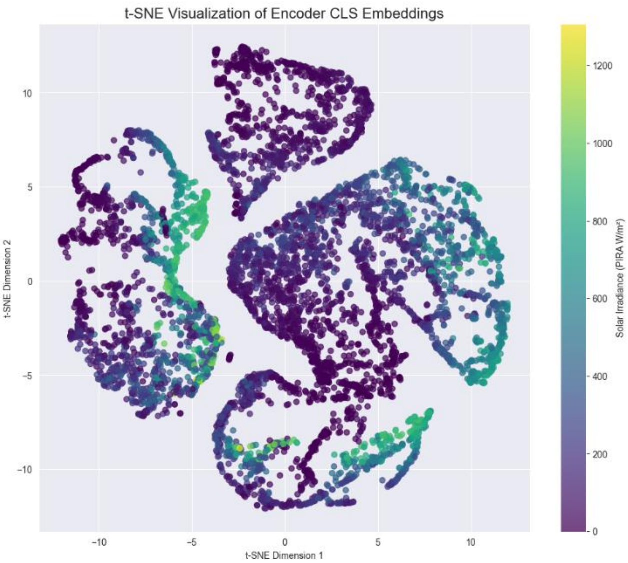


Figure 4: t-SNE Visualization of Encoder Embeddings

We ran K-Means (k = 3) on the 2D t-SNE embeddings, yielding three clusters. Sampling images from each cluster then reveals the semantic distinctions among the regions of the learned manifold.

Cluster 1: Overcast Conditions

Figure 5: Sample Images from Cluster 1

As shown in **Figure 5**, the first cluster is dominated by images depicting thick, uniform, overcast skies or dense cloud cover.

Cluster 2: Clear Sky Conditions

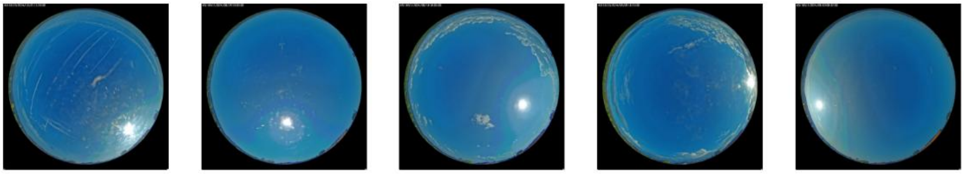


Figure 6: Sample Images from Cluster 2

The second cluster, illustrated in **Figure 6**, contains images of predominantly clear blue skies, sometimes with wispy clouds. The sun is often clearly visible and unobscured.

Cluster 3: Partly Cloudy Conditions

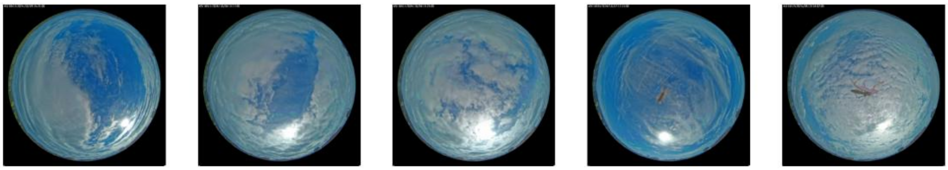


Figure 7: Sample Images from Cluster 3

The final cluster, shown in **Figure 7**, captures the intermediate and most complex conditions. It consists of images with broken clouds, scattered clouds, or hazy skies.

The t-SNE visualization and the subsequent clustering analysis provide initial evidence that our model is successfully learning to disentangle complex atmospheric scenes into a semantically meaningful feature space, confirming our methodology is on the right track.

3.2 Cloud Classification

We assembled a balanced dataset of 1 000 all-sky images, evenly divided among four classes: **clear sky**, **partial clouds**, **heavy clouds**, and **full overcast**. We randomly split the data into Training (800 images), Validation (200 images), Test (200 images).

We trained the SSL encoder with a lightweight classifier head. Input images were resized to 224×224 and fed in batches of 16, with random horizontal flips and color-jitter augmentations. Each model was trained for up to 60 epochs (early stopping on validation loss) using the cross-entropy loss. As baselines, we evaluated ViT-Base pretrained on ImageNet and a randomly initialized (scratch).

Table 1 reports the per-class F1-scores for cloud classification across the three ViT variants, and our SSL pre-trained model. Our SSL backbone achieves the best overall F1.

Cloud Class	Random F1 (%)	ImageNet F1 (%)	SSL F1 (%)
Clear Sky	78.57	**100.00**	98.99
Full Overcast	70.69	90.72	**93.62**
Heavy Clouds	40.91	80.77	**82.46**
Partial Clouds	38.24	82.83	**83.87**

Table 1: Per-class F1-scores for cloud classification

4 Discussion and Future Work

4.1 Interpretation of Preliminary Findings

The preliminary results presented in this paper offer strong validation for our hybrid self-supervised learning methodology. The clear, emergent structure in the t-SNE feature space where images cluster by cloud condition demonstrates that even at the halfway point of pre-training, the encoder has learned physically meaningful representations. When we attached a lightweight classifier head and fine-tuned this halfway-trained encoder for cloud classification, it achieved per-class F1-scores of 98.99% (“clear sky”), 93.62% (“full overcast”), 82.46% (“heavy clouds”), and 83.87% (“partial clouds”), outperforming ImageNet and random baselines.

As this is a work in progress, the final performance remains to be determined. Full-schedule pre-training and downstream evaluations will reveal the ultimate efficacy of our approach and guide any necessary tuning of loss weights (λ_{pixel}, $\lambda_{distill}$, λ_{xview}) or other hyperparameters. Nonetheless, our halfway-training results already establish a promising foundation for domain-specific sky-image models.

4.2 Roadmap to Completion and Evaluation

Following this assessment, we will employ a larger, more challenging dataset for the **cloud-classification** task building on our initial small-scale experiment to fine-tune the pre-trained encoder. We will also extend fine-tuning to **cloud segmentation** and **solar-irradiance forecasting**, using task-appropriate datasets. Performance on these downstream tasks, benchmarked against models trained from scratch, will deliver a definitive evaluation of our foundation model's effectiveness.

4.3 Broader Vision and Future Applications

Beyond the immediate goals of cloud classification and irradiance forecasting, the long-term vision for this project is to establish the utility of domain-specific foundation models in the atmospheric sciences. The powerful feature representations learned by our encoder have the potential to serve as a versatile starting point for a range of other important downstream tasks relevant to renewable energy and environmental monitoring. Furthermore, the current model is trained on a dataset capturing several seasons; a logical future step would be to expand the dataset to encompass a full, continuous year of observations. Retraining the model on such a comprehensive dataset would allow it to learn a complete cycle of seasonal variations, further enhancing its robustness and creating an even more powerful tool for atmospheric analysis.

5 Conclusion

In this work-in-progress paper, we have presented the methodology and preliminary validation of a foundation model for all-sky imagery, designed to address the challenge of solar forecasting. Our approach utilizes a hybrid self-supervised learning framework that combines the fine-grained texture learning of masked autoencoding with the high-level semantic feature learning of self-distillation. The key contribution of our architecture is a multi-view consistency loss that leverages simultaneous data from four distinct cameras to enforce a robust, view-invariant scene representation.

Our initial findings are highly encouraging. The analysis of the learned feature space, at an intermediate training stage, demonstrates that the model is successfully organizing complex atmospheric imagery into a coherent manifold that correlates strongly with ground-truth physical properties like cloud conditions. This was achieved without any exposure to explicit labels, providing evidence that our methodology is effectively guiding the model to learn the underlying physics of the atmosphere. This work validates our approach and represents a significant step toward creating specialized, powerful tools for the renewable energy sector. The completed feature extractor is poised to enhance the performance of downstream tasks such as cloud classification and solar nowcasting, ultimately contributing to improved grid stability and the wider adoption of solar power.

Bibliography

[Ca21] Caron, M., Touvron, H., Misra, I., Jégou, H., Mairal, J., & Bojanowski, P. (2021). Emerging Properties in Self-Supervised Vision Transformers. *arXiv preprint arXiv:2104.14294*.

[Do20] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., & Houlsby, N. (2020). An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. *arXiv preprint arXiv:2010.11929*.

[Fa22] Fabel, Y., Nouri, B., Wilbert, S., et al. Applying Self-Supervised Learning for Semantic Cloud Segmentation of All-Sky Images. *Atmospheric Measurement Techniques*, 15, 797–809, 2022.

[He22] He, K., Chen, X., Xie, S., Li, Y., Dollar, P., & Girshick, R. (2022). Masked Autoencoders Are Scalable Vision Learners. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 15979–15988).

[Jo21] Johnson, J. W., Hari, S., Hampton, D., Connor, H. K., & Keesee, A. A Contrastive Learning Approach to Auroral Identification and Classification. arXiv:2109.13899, 2021.

[Lo17] Loshchilov, I., & Hutter, F. (2017). Decoupled Weight Decay Regularization. arXiv preprint arXiv:1711.05101.

[Oq23] Oquab, M., Darcet, T., Moutakanni,T., Vo, H., Szafraniec, M., & Bojanowski, P. (2023). DINOv2: Learning Robust Visual Features without Supervision. *arXiv preprint arXiv:2304.07193*.

[Sc25] Schreder CMS, http://schreder-cms.com/de/index.htm, accessed: 18/06/2025.

Predicting the Annual Dirty Ten Based on Open Data about Pollution Releases

Heiko Thimm [1], Vanessa Schmidt[2]

Abstract: The German Pollutant Release and Transfer Register (PRTR) contains over 80,000 reports from more than 5,000 companies, offering substantial transparency regarding environmental pollution. This research uses PRTR data to predict the annual "Top Ten Dirty," which are industrial entities expected to report the ten largest releases of a specific substance. Following the CRISP-DM process model, an analytical dataset was constructed and used to train two machine learning models: one to predict the annual, branch-specific top ten polluters for greenhouse gas emissions into the air and another for heavy metal discharges into water. The results demonstrate that both models produce predictions closely aligned with actual pollution reports. These predictions can support various stakeholders, including corporate environmental managers, regulatory agencies, business organizations, and authorities combating environmental crimes.

Keywords: Corporate Pollution Reporting, Data Science, Machine Learning, Pollution Prediction, German PRTR register

1 Introduction

Historical pollution data from official registers, such as the European Pollutant Release and Transfer Register (E-PRTR), is crucial for many reasons, including environmental monitoring, regulatory purposes, and transparency. Equally important is the ability to predict future environmental impacts for stakeholders such as corporate sustainability managers, investors, regulatory bodies, communities, and individuals living close to industrial facilities. For example, accurate forecasts of greenhouse gas (GHG) emissions are essential for developing effective pollution reduction strategies [JC22]. This study explores using Machine Learning (ML) techniques to predict future polluters, focusing on identifying the top ten annual polluters by industry branch, also referred to as the `annual dirty ten´ companies [EE23]. These companies report the ten largest annual pollution amounts above the legal threshold for a particular type of pollution.

This analysis is based on the German PRTR dataset. Two Random Forest models were trained to predict the top ten polluters: one for GHG emissions into the air and another for heavy metal discharges into water.

[1] Pforzheim University, School of Engineering, Tiefenbronner Str. 65, Pforzheim, D-75175, heiko.thimm@hs-pforzheim.de, https://orcid.org/0000-0002-6200-9655
[2] Pforzheim University, School of Engineering, Tiefenbronner Str. 65, Pforzheim, D-75175, vanessa.schmidt@hs-pforzheim.de

doi:10.18420/env2025_08

The study details the data preparation techniques and ML methods applied, following the CRISP-DM process model, which is widely used in Data Science and ML research and practice [Ma21]. Performance evaluation metrics demonstrate that the models' predictions closely align with actual pollution reports.

Existing prediction methods often rely on large datasets, such as Earth observation data and Internet of Things (IoT) sensor data. These methods typically focus on short-term predictions of air and water quality. These approaches require complex data management systems and significant computational resources. In contrast, this study uses structured PRTR data to train ML models that can make annual predictions without requiring extensive infrastructure or high computational costs.

The following sections present a review of related research, an overview of the research methodology and design, and a description of the German PRTR dataset. The subsequent section details the main steps of the approach to using ML to predict the top 10 polluters. The article ends with the researchers' conclusions.

2 Related Work

Numerous studies across different disciplines have employed data from national PRTRs or combined registers to explore industrial pollution and its impacts. For instance, an empirical financial analysis examined how firms' environmental disclosures influence market valuation and investments in pollution prevention using toxic release data from the Japanese PRTR [HM10]. Ultimately, the study found that the Japanese financial market does not adequately price the risks associated with toxic chemical releases.

In Sweden, researchers developed a methodology to estimate the national environmental impact of chemical emissions using PRTR data [SPF16]. They aggregated and transformed emission data for 54 substances released into the air and water in 2008 using a life-cycle impact assessment approach. This enabled them to calculate potential human toxicity and ecotoxicity at a national level.

Erhart et al. [Er25] investigated the exposure of freshwater biodiversity in Natura 2000 sites to industrial pollution by spatially linking protected areas with emission data from the E-PRTR. They found that many protected sites are subject to elevated ecological risks, highlighting the urgent need for integrated conservation and pollution control measures.

Other researchers have explored the potential of PRTR data to support cleaner production initiatives. One study examined an Australian fertilizer manufacturing facility to determine if PRTR reporting could help identify and implement cleaner production strategies [KS04]. While the study emphasized the utility of PRTR data for program development and monitoring, it also noted a key limitation: cleaner production programs typically require process-level pollution data, whereas PRTRs report emissions at the facility level.

A New Zealand study [HM10] used Australian PRTR data as a proxy for environmental performance and introduced a toxicity-weighting scheme for over 90 chemicals. The authors argued that toxicity-weighted indicators offer a more meaningful alternative to mass-based metrics for stakeholders, including corporate managers, policymakers, and analysts.

Erhart and Erhart [EE23] contributed a large-scale environmental ranking of over 10,000 European industrial facilities by integrating toxicity and global warming potentials using a harmonized life cycle assessment framework. Their results demonstrate substantial disparities in environmental performance across industries and regions, offering guidance for targeted policy interventions and corporate environmental improvements.

3 Research Methodology and Design

The study employed an experimental, quantitative research methodology and used historical PRTR data as the primary data source. This methodological approach resulted in a research design that ensures rigorous, reproducible, multi-step work, establishing a foundation for future research and improvements to predictive models. The steps are described in the remaining paragraphs.

Data collection. The dataset was obtained from the German PRTR register as of May 2023. It includes notifications from companies reporting their annual pollutant emissions and amounts of dangerous and non-dangerous waste across different pollution categories.

Data preparation. The collected data underwent preprocessing to handle missing values identified as `NaN´ (not a number), encode categorical variables, and normalize numerical features. Reports of non-targeted pollution types were retained, as they are relevant for identifying hidden patterns for the prediction task.

Exploratory Data Analysis (EDA) and Feature Selection. EDA was conducted to understand the distribution, patterns, and relationships within the dataset to help identify key variables influencing pollutant emissions. Descriptive statistics, visualizations, and correlation analyses were applied to assess data characteristics, detect outliers, and identify missing values.

Model development. Two binary classification models were developed using supervised ML techniques: one that predicts the top 10 polluters regarding GHG emissions into the air (AirT10P) and another that predicts the top 10 polluters regarding the release of heavy metals into the water (WaterT10P). These models predict whether a company falls within the top ten polluters based on reported emissions. Random forest classifiers [Gu22] were chosen for their interpretability, accuracy, and ability to cope with imbalanced datasets.

Data resampling. Since the dataset exhibited class imbalance, especially with fewer notifications belonging to the top 10 polluters, data resampling techniques were applied.

Downsampling was performed to reduce the number of non-top-10 cases to match the number of top-10 notifications. Upsampling was applied to increase the number of top-10 cases by duplicating existing data points, thus balancing the dataset.

Model training and validation. Two binary classification models were trained to predict whether a company would rank among the top 10 polluters (T10P): a model to predict top 10 polluters regarding GHG emissions into the air (AirT10P) and a prediction model regarding heavy metal emissions into water (WaterT10P). K-fold cross-validation [Ya23] with k=5 was used to evaluate model performance.

Secondary validation. To confirm the robustness of the models further, an additional validation step was performed. During resampling, notifications with NaN values for AirT10P and WaterT10P were removed from the training dataset. These notifications were then reintroduced as a separate validation dataset to test the models' ability to correctly classify instances outside the training set. The performance results were consistent with those of the initial validation step, with only minor differences at the second decimal place.

Model deployment and validation based on new data. To obtain insights into model deployment, the models were applied to predict the top ten polluters in terms of GHG emissions and heavy metal emissions for the year 2023. The results were then compared to the actual top ten polluters. The confusion matrix obtained from these comparisons revealed that the ML models provide highly accurate predictions. However, both models predicted the top 10 polluters with relatively low, yet still acceptable, precision.

4 The German PRTR Dataset

The German Federal Environmental Agency publishes the German PRTR dataset on its official website and updates it annually within the first two months of the calendar year. This study used a dataset downloaded in October 2024 that contained approximately 80,000 pollution reports from 2007 to 2022. These reports were submitted by approximately 5,000 industrial facilities, including small businesses, SMEs, and large multinational corporations. Over the 16-year period, 105,163 business activities were registered, 83,484 of which were designated as primary business activities.

Companies are classified annually according to the NACE Code standard [Eu22], which identifies their primary business activities. Each activity is also assigned a PRTR-specific code and categorized into one of nine industry branches: waste and wastewater management, the chemical industry, the energy sector, intensive live-stock farming and aquaculture, the food industry, the metal industry, the mineral processing industry, the paper and wood industry, and other industrial sectors.

The relevant key characteristics of the German PRTR dataset are as follows. Each company is assigned a unique identifier for each reporting year, along with updated information such as the company's name, address, and NACE code. An additional identifier remains consistent across all reporting years. Each company's activities are described annually using PRTR-specific activity codes and categorized into industry branches. Reports classify pollution as hazardous or non-hazardous waste, wastewater shipments, or controlled or inadvertent releases into the air or water. For wastewater, air, and water releases, annual amounts of specific substances are reported and categorized by substance group, such as heavy metals and GHGs.

5 Machine Learning-Based Prediction of Top 10 Polluters

This experiment implements the aforementioned research design. The prediction task is modeled as a binary classification problem in which companies are classified as either a top 10 polluter (1) or a non-top 10 polluter (0) for a given year, branch, and pollution type.

5.1 Data Preparation

The first step in the data preparation process involved selecting PRTR reports and variables essential for the ML model from the PRTR register. The reports were selected based on specific constraints to ensure accurate, up-to-date predictions. Companies were included if they had reported data beyond 2012 and had notifications spanning at least 15 years, including one after 2020. This filtering process resulted in a dataset in which each row represents a company's main annual activity and reported pollution amount.

The dataset was then supplemented with population density data from the German Federal Statistical Office (DStatis) and the official municipality list (June 2024 version). If necessary population numbers were also retrieved from official municipal websites.

Annual branch-specific rankings and top 10 polluter labels were created by comparing reported pollution amounts. Each year, companies ranked from one to ten received a label of one, while others received a label of zero. This process added three columns for GHG emissions: `AirAmnt´ (annual GHG amounts), `AirRank´ (branch-specific rank), and `AirT10P´ (top 10 polluter indicator). Similarly, three columns were added for heavy metal releases into water: `WaterAmnt´, `WaterRank´, and `WaterT10P´.

5.2 Analysis of Potential Predictor Variables

To improve prediction accuracy, the independent (feature) variables should be strongly correlated with the dependent variables. Using AirAmnt and WaterAmnt as proxies for AirT10P and WaterT10P allowed for more precise correlation analysis. This approach was chosen because numerical variables provide more detailed insights into relationships than binary categories do, particularly given the class imbalance in the dataset.

The following features were analyzed for potential correlations with pollution amounts:

- **Business Sector (NACE Code):** Different sectors have distinct production methods, equipment, and environmental impacts, making this a relevant predictor.
- **Type of Activity (PRTR Key):** Variations in facility activities affect pollution levels, supporting its inclusion as a predictor.
- **Location (Postal Code):** Geographic factors, such as infrastructure, proximity to communities, and environmental regulations, may impact pollution levels.
- **Reporting Year:** Economic conditions and technological progress change annually and may influence pollution amounts.
- **Population Density:** Higher population densities may lead to increased regulatory scrutiny and compliance pressure, which could result in lower pollution levels.
- **Secondary Activities:** A high number of non-main activities with pollution reports indicates a production site with high pollution levels, making this a relevant predictor.
- **Release above limit absolute:** Companies can vary greatly in the amount by which they exceed the regulatory limit. Companies with a large excess amount are more likely to be ranked among the top 10 polluters than companies with a small excess amount.
- **Release above limit relative:** The relative excess amount can also be considered a dependent variable.

The results of ANOVA tests [Co21] indicate that the categorical variables `Business Sector´, `Type of Activity´, and `Location´ significantly influence annual pollution amounts. Pearson correlation tests [Co21] revealed statistically significant but very weak positive correlations for continuous variables (year, population, secondary activities, and release above limit absolute/relative) with the following exceptions. The test between `Release above limit absolute´ and `AirAmnt´ revealed a moderate correlation. Weak positive correlations were obtained for `Secondary Activities´ and the two dependent variables, `AirAmnt´and `WaterAmnt´.

5.3 Resampling and Modeling

The analytical dataset contained several columns with missing values, especially for pollution amounts, pollution rank, and top 10 polluter labels. This was because the dataset included reports for all PRTR pollution types, not just the types targeted by the prediction models. Additionally, even within subsets of pollution reports focusing on specific pollution types, years, and industry branches, the data exhibited significant class imbalance. This imbalance occurred because only 10 reports could be labeled `1´ to indicate a top polluter, while the rest were labeled `0´ to indicate non-top polluters. Consequently, in branches with numerous companies, the proportion of `1´ labels was substantially outnumbered by `0´ labels.

To ensure a robust foundation for model development, it was necessary to address this imbalance using appropriate data resampling techniques. Therefore, five downsampling steps were applied to the majority class. First, all records from firms that did not submit any reports after 2017 were removed from the dataset. Second, firms that submitted five or fewer activity reports up to and including 2020 were identified, and all of their corresponding records were excluded from the dataset. Third, eighty percent of reports from 2007 to 2012 containing missing values in any independent variable were randomly selected and removed from the dataset. Fourth, for reports from 2013 to 2018 that had missing values in an independent variable, 40% were randomly selected and removed from the dataset. Fifth, across the entire period from 2007 to 2022, 70% of reports containing missing values in an independent variable were randomly removed from the dataset. In addition to downsampling oversampling was applied using the SMOTE method [Ch02] in order to increase the minority class.

The training and validation of Random Forrest Models for the two classification tasks was performed based on K-fold cross-validation with 5 folds.

5.4 Evaluation – First Experimental Results

The validation results for the two models that predict the binary dependent variables AirT10P and WaterT10P are presented in Table 1. The performance measures reported use common ML indicators: accuracy, precision, recall, F1 score, Cohen’s kappa (CK), and Area Under the Curve (AUC) [RTK24][21].

Variable	**Accuracy**	**Precision**	**Recall**	**F1 Score**	**CK**	**AUC**
AirT10P	0.980	0.981	0.980	0.981	0.829	0.986
WaterT10P	0.978	0.980	0.978	0.978	0.797	0.982

Tab. 1: Performance validation results obtained through k-fold cross-validation

All indicators are close to the maximum possible value of 1.0, except for Cohen’s kappa, which is slightly lower. These performance metrics suggest that the predictions of both models are highly consistent with the true labels. Additionally, the Cohen's kappa values show that the models' agreement with the true labels is significantly better than would be expected by chance.

5.5 Model Deployment and Validation based on New Data for the Year 2023

The models were trained using pollution reports from 2007 to 2022. In March 2025, a new version of the PRTR dataset was released that included data for the 2023 reporting year. This update permitted a meaningful out-of-sample evaluation of model performance by comparing the predicted top ten polluters for 2023 with the actual reported outcomes.

To support model deployment, an input dataset was constructed to mirror the structure and features of the training data. Since the models were designed to perform binary classification the input dataset had to include firms with a reasonable likelihood of having reported pollution in 2023. To this end, firms that reported any pollution related to their primary activity in 2022 were selected. This decision was informed by prior research [TS25], which indicated that firms active in the most recent reporting year are more likely to continue reporting in subsequent years.

The base input dataset was derived from the 2022 subset of the training data and included 5,371 company records. Standard preprocessing steps, were applied to all numerical features of the dataset, consistent with the original training pipeline. Input vectors were constructed using the same set of predictor variables as during training. Static variables, such as `Business sector´, `Type of activity´, and `Location´, retained their 2022 values because they are unlikely to change from year to year. New values were computed for time-sensitive variables, such as `Secondary Activities´, `Release of GHG into the air above limit abs.´, and `Release of GHG into the air above limit rel.´ as the average over the years 2007 to 2022. This strategy was adopted due to the lack of 2023 values and is supported by findings in [TS25], which observed variability in these indicators from year to year. The variable `Year´ was set to `2023´ to reflect the prediction target.

The prediction results were compared to the actual top 10 lists of the year 2023. Model performance was evaluated using standard classification metrics: True positives (TP), true negatives (TN), false positives (FP), false negatives (FN), accuracy, and precision. Tables 2 and 3 present the detailed performance indicators for each industrial sector.

Branch	**TP**	**TN**	**FP**	**FN**	**Accuracy**	**Precision**
Waste and wastewater management	7	1998	3	3	0.997	0.7
Chemical industry	7	567	3	3	0.989	0.7
Energy sector	6	226	4	4	0.966	0.6
Intensive livestock farming and aquaculture	0	503	10	0	0.980	0.0
Food industry	0	344	10	0	0.971	0.0
Metal industry	8	985	2	2	0.995	0.8
Mineral processing industry	8	218	2	2	0.982	0.8
Paper and wood industry	1	135	9	6	0.900	0.1
Other industrial sectors	2	285	8	0	0.972	0.2

Tab. 2: Evaluation of predicted Top 10 Polluters for Release of Heavy Metals into the water

Across most sectors, the models achieved high accuracy, often exceeding 95%. However, precision varied significantly between sectors. For instance, the metal and chemical industries yielded high accuracy and precision, indicating robust predictive capability. In contrast, sectors such as intensive livestock farming, the food industry, and other industrial sectors demonstrated low precision. Some cases, such as heavy metal discharge in the food industry, showed a precision of 0%. These results are largely attributable to the class imbalance inherent in the data, where the number of negative cases (non-Top 10 polluters) significantly exceeds the number of positive cases, especially since each branch can contain a maximum of ten Top polluters by definition.

Interestingly, model performance was consistent across pollution types in certain branches, such as waste and wastewater management and the chemical industry. This suggests similar patterns of pollution behavior and model interpretability. For other branches, substantial variation was observed between the two pollution types.

These outcomes highlight the strengths and limitations of the current model design. While the model is accurate overall, the lower precision in some branches indicates the need for further refinement, potentially through the use of additional sector-specific features, improved class balancing techniques, or ensemble methods.

Branch	**TP**	**TN**	**FP**	**FN**	**Accuracy**	**Precision**
Waste and wastewater management	6	1997	4	4	0.996	0.6
Chemical industry	7	567	3	3	0.989	0.7
Energy sector	8	228	2	2	0.983	0.8
Intensive livestock farming and aquaculture	0	503	10	0	0.980	0.0
Food industry	5	341	5	3	0.977	0.5
Metal industry	9	986	1	1	0.997	0.9
Mineral processing industry	5	215	5	5	0.956	0.5
Paper and wood industry	7	138	3	3	0.960	0.7
Other industrial sectors	3	283	7	2	0.969	0.3

Tab. 3: Evaluation of predicted Top 10 Polluters for Release of GHG into the air

6 Conclusions

This article describes an experiment that uses PRTR data to develop ML models that predict the top 10 polluters based on their releases of GHG into the air and heavy metals into the water. The validation results indicate that both models produce predictions that are consistent with the true labels. The dataset can support the development of additional ML models, including regression models that predict future pollution levels and detect cases warranting site inspections. These developments can advance environmental protection applications. However, it is important to note that ML predictions are probabilistic and depend on high-quality real-world data. Misleading predictions due to data limitations could lead to unjustified accusations, which emphasizes the importance of verifying prediction results, especially when they are shared publicly.

Future work will involve developing prediction models for additional pollution types beyond those addressed in this work. By augmenting the dataset, global rankings across all pollution types can be generated. This expanded dataset may lead to new features, improving ML models or enabling the development of regression models that predict companies' future pollution levels. Furthermore, the analysis will investigate patterns of notifications related to individual companies, business sectors, cities, regions, and federal states. Identifying patterns across different pollution types could reveal interdependencies caused by technical processes, enabling further model refinements.

Bibliography

[Ch02] Chawla, N. V. et al.: SMOTE: Synthetic Minority Oversampling Technique. Journal of Artificial Intelligence Research 16, pp. 321–357, 2002.

[Co21] Coolidge, F.L. Ed.: Statistics: A Gentle Introduction. SAGE Publications, Inc, 2455 Teller Road, Thousand Oaks California 91320, 2021.

[EE23] Erhart, S.; Erhart, K.: Environmental ranking of European industrial facilities by toxicity and global warming potentials. Scientific reports 1/13, p. 1772, 2023.

[Er25] Erhart, S. et al.: Freshwater biodiversity risk exposure of Natura 2000 sites to industrial pollution. Environmental Impact Assessment Review 113, p. 107797, 2025.

[Eu22] European Commission: Regulation (EC) No 1893/2006, Statistical classification of economic activities NACE Revision 2. https://eur-lex.europa.eu/eli/reg_del/2023/137/oj, accessed 4 Aug 2024.

[Gu22] Gu, Q. et al.: A novel Random Forest integrated model for imbalanced data classification problem. Knowledge-Based Systems 250, p. 109050, 2022.

[HM10] Hibiki, A.; Managi, S.: Environmental Information Provision, Market Valuation, and Firm Incentives: An Empirical Study of the Japanese PRTR System. Land Economics 2/86, pp. 382–393, 2010.

[JC22] Javed, S. A.; Cudjoe, D.: A novel grey forecasting of greenhouse gas emissions from four industries of China and India. Sustainable Production and Consumption 29, pp. 777–790, 2022.

[KS04] Kolominskas, C.; Sullivan, R.: Improving cleaner production through pollutant release and transfer register reporting processes. Journal of Cleaner Production 7/12, pp. 713–724, 2004.

[Ma21] Martinez-Plumed, F. et al.: CRISP-DM Twenty Years Later: From Data Mining Processes to Data Science Trajectories. IEEE Transactions on Knowledge and Data Engineering 8/33, pp. 3048–3061, 2021.

[RTK24] Rainio, O.; Teuho, J.; Klén, R.: Evaluation metrics and statistical tests for machine learning. Scientific reports 1/14, p. 6086, 2024.

[SPF16] Sörme, L.; Palm, V.; Finnveden, G.: Using E-PRTR data on point source emissions to air and water—First steps towards a national chemical footprint. Environmental Impact Assessment Review 56, pp. 102–112, 2016.

[TS25] Thimm, H.; Schmidt, V.: A Straight Forward Method to Analyze the Variation Dynamics of Corporate Top Polluter Rankings Exemplified Through the German PRTR Register. J. Earth & Environment Science 1/5, pp. 1–9, 2025.

[Ya23] Yates, L. A. et al.: Cross validation for model selection: A review with examples from ecology. Ecological Monographs 1/93, 2023.

Green Coding Techniques and Methodologies

EnerAgent: Energy-Aware Code Refinement with LLM-Based Agents

Jonas Scheffler [1] and Max Westing [1]

Abstract: Large Language Models (LLMs) have shown strong performance in generating functionally correct source code, but their outputs often lack consideration for runtime efficiency and sustainability. In this work, we introduce EnerAgent, an agentic framework that uses iterative feedback to refine LLM-generated code toward reduced energy consumption. EnerAgent first ensures functional correctness, then performs iterative optimization guided by energy usage measurements. We evaluate the agent on a set of complex algorithmic problems in C++, comparing against both single-attempt LLM outputs and expert-written solutions. Our results show that EnerAgent consistently improves the energy efficiency of its initial code, with energy savings in over half of the tasks, often matching or surpassing human-written code. These findings highlight the promise of feedback-driven LLM agents for greener software development.

Keywords: Energy-Efficient Software, Large Language Models, Code Optimization, Autonomous Agents, Green Software Engineering

1 Introduction

Large Language Models (LLMs) have demonstrated impressive capabilities in code generation, achieving near-human performance across a range of programming tasks [Ji24]. This progress has sparked interest in using LLMs not just for functional correctness, but also for optimizing non-functional properties such as runtime, memory usage, and energy consumption.

Energy-efficient programming remains challenging as it requires deep knowledge of algorithmic trade-offs, hardware behavior, and the specific problem domain. For LLMs, the challenge is greater: efficiency is not reflected in their training objectives, and energy-aware behavior must emerge implicitly [SWC25]. To address this, recent works have explored agentic approaches where LLMs iteratively refine their outputs based on runtime feedback [Hu24; Ya23].

In this work, we introduce *EnerAgent*, an energy-aware agent that combines LLM-based code generation with feedback-driven refinement. The agent generates candidate solutions,

[1] Environmental Campus Birkenfeld, Trier University of Applied Sciences, Germany,
jonas.scheffler1@gmail.de, https://orcid.org/0009-0001-8606-8916;
m.westing@umwelt-campus.de, https://orcid.org/0009-0001-5044-7578
This paper was written as part of the course “Nachhaltige Softwaretechnik” taught by Prof. Dr. Stefan Naumann.

doi:10.18420/env2025_09

executes them to obtain runtime feedback, and iteratively improves the code by addressing inefficiencies. Unlike static, single-attempt code generation, this loop enables targeted exploration of the solution space to improve energy performance.

We evaluate EnerAgent on a suite of C++ programming tasks, comparing its performance to both human-written solutions and unrefined LLM outputs. EnerAgent often matches or surpasses human baselines in energy use, highlighting the potential of agentic LLMs for sustainable code generation.[2]

2 Related Works

As LLMs gain traction in software development, interest is growing in using them not just for functional correctness but also for efficient, sustainable code generation. Future software engineering will likely integrate LLMs with energy-aware capabilities, both in code output and in model operation [SYL24].

Initial work has focused on prompt design to steer models toward energy-efficient solutions. Minor changes in phrasing or explicit instructions can impact energy characteristics, though such gains are often inconsistent and task-specific [CRO24; Ni24].

More robust are agentic approaches, where LLMs refine outputs based on feedback from real execution [Ya23]. Huang et al. [Hu24] proposed *EffiLearner*, an iterative system for optimizing Python code using runtime and memory feedback, showing the promise of performance-aware refinement.

Our work builds on this paradigm with a fully automated design focused on energy as the primary metric, and aims to generalize across languages via Dockerized execution and energy tracking.

3 Methodology

3.1 EnerAgent Framework

In this work, we define **energy** as the amount of computational energy consumed by a program during its execution, measured using the CodeCarbon framework [CSL+24]. A solution is considered **energy-efficient** if it minimizes energy consumption while preserving functional correctness.

EnerAgent is a multi-stage, feedback-driven agent built to generate energy-efficient solutions to algorithmic programming tasks. Implemented using LangGraph [CC25], it structures the

[2] All source code and data for reproducing our results are available at: https://gitlab.rlp.net/s20a8d902543/eneragent-public

generation and refinement process as a directed state machine, where each node represents a distinct phase in the workflow.

The agent proceeds through the following structured phases:

- **Initial Code Generation:** An initial solution is generated via a zero-shot prompt [Ma20] that instructs the LLM to solve the task in C++ using the provided problem description and insert the solution into the given function template.
- **Compilation and Testing:** The generated code is compiled and executed inside a Docker-based execution environment. Code correctness is verified against predefined functional tests. If compilation or execution fails, the agent triggers a reflection step.
- **Reflection and Repair:** If the code fails to compile or pass the functional tests, the agent prompts the LLM with the problem description, current code, and error messages, explicitly instructing it to fix the issue based on the feedback. This process iterates until a working solution is found or a maximum number of repair attempts is reached.
- **Energy Optimization:** Once functional correctness is achieved, the agent transitions into the optimization phase, where it iteratively prompts the LLM to improve the code's energy efficiency. The prompt follows a chain-of-thought structure [We22], asking the model to analyze the code step by step, identify inefficiencies related to control flow, data structures, or memory use, and then propose an improved version. If no inefficiencies are found, the model is instructed to respond with `NO_CHANGE` to terminate the optimization phase, preventing unnecessary computation and overfitting to noise in energy measurements. Energy consumption is measured during each execution using the CodeCarbon framework [CSL+24]. The agent keeps track of the best-performing solution and continues the loop as long as improvements are detected. If a newly proposed solution is functionally incorrect, it is discarded, and the agent resumes optimization from the last verified correct version. This prevents the agent from becoming stuck attempting to fix broken variants that may not yield energy gains. This energy feedback mechanism is designed to be fast and lightweight, allowing the agent to iteratively optimize solutions without introducing significant overhead. More complex or hardware-based measurement techniques (e.g., external power meters) would make real-time agent optimization impractically slow. Our approach prioritizes a low-latency, transferable energy signal that can run fully autonomously, making the agent applicable in real-world settings where runtime constraints matter.

The exact prompt templates used during code generation, error repair, and energy optimization are listed in the appendix.

To encourage exploration, EnerAgent increases the LLM's temperature after each failed optimization attempt—starting with low-temperature outputs to address obvious inefficiencies, then gradually generating more diverse variants. While this strategy is intuitively motivated and supported by prior work on temperature effects in LLMs [Pe24; Re24], we did not isolate its contribution in our benchmark.

Fig. 1 illustrates EnerAgent's workflow as a directed graph of states, supporting modular handling of code generation, verification, and energy-aware refinement.

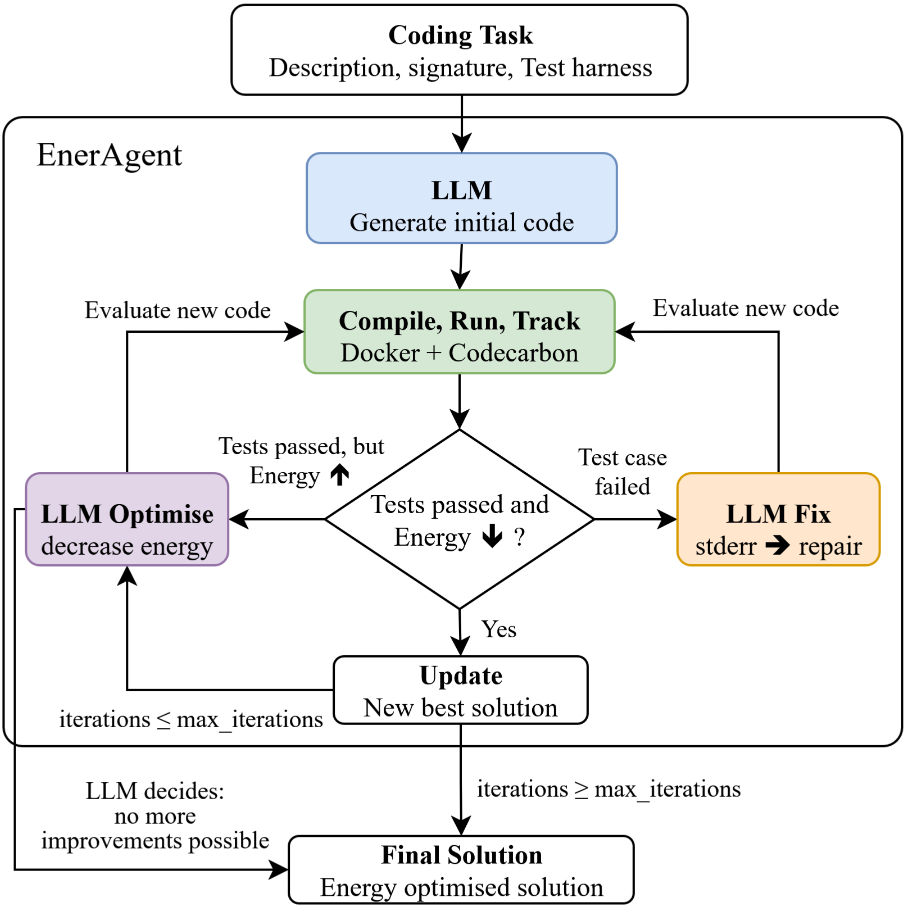


Fig. 1: **Eneragent Workflow.**The agent generates an initial solution using an LLM and a code template, then compiles, tests, and measures energy consumption. If incorrect, it enters a repair loop. Once correct, it performs iterative optimization to reduce energy use. The process ends after a fixed number of non-improving steps or when refinement is halted.

3.2 Experiment Design

To evaluate EnerAgent, we conducted experiments on a benchmark of algorithmic programming tasks drawn from LeetCode [3], focusing on problems labeled as "hard" to ensure sufficient complexity and challenge. These tasks typically require non-trivial algorithmic reasoning, such as dynamic programming, graph traversal, backtracking, or advanced data structure manipulation. Each task includes a natural language problem description, a C++ function signature template that defines the expected interface, and a set of functional tests used to verify correctness of candidate solutions. We chose C++ because its low-level nature allows for fine-grained optimizations, making it a more suitable language for evaluating energy-aware code refinement.

We compare the following methods:

- **Human Baseline:** Solutions manually written by expert programmers, measured for energy consumption.
- **LLM Single-attempt:** Code generated by prompting the LLM once, without refinement or feedback.
- **EnerAgent:** Our agentic framework that iteratively refines its output using feedback. It may attempt up to three repairs for incorrect code and continues energy optimization until either three consecutive non-improving attempts occur, or the LLM explicitly halts the loop.

Energy measurements were collected using CodeCarbon within Dockerized execution environments.

4 Results

We evaluate the methods along two primary axes: **effectiveness**, measured by the success rate of generating functionally correct code and the number of repair attempts required; and **efficiency**, assessed by the energy consumption of the generated code compared to human-written solutions and the degree of improvement achieved through iterative optimization.

To standardize our comparison of energy efficiency, we define the **Relative Energy Speedup (RES)** as the proportion of energy saved compared to a given baseline:

$$\mathrm{RES} = \frac{E_{\mathrm{baseline}} - E_{\mathrm{agent}}}{E_{\mathrm{baseline}}} \tag{1}$$

This metric reflects the relative reduction in energy consumption achieved by the agent. A higher RES indicates better energy efficiency. We apply this metric both in relation to the agent's initial solution (Sect. 4.2) and to human-written code (Sect. 4.3).

[3] https://leetcode.com

4.1 Functional Correctness (Effectiveness)

Tab. 1 summarizes EnerAgent's effectiveness in producing functionally correct solutions. Overall, the agent achieves a high success rate of over 80% for two different LLMs. A substantial portion of problems (over 60% for both LLMs) are solved in a single generation without any need for repair.

This high single-attempt success rate is unsurprising given that modern LLMs are extensively trained on code datasets, particularly on well-structured domains like competitive programming problems. However, initial generations can still suffer from minor issues, such as syntax errors, or failure to handle edge cases, that prevent compilation or correct execution.

A simple retry in which the error message is fed back into the LLM addresses many of these minor failures effectively. Almost all solved errors were corrected after just one additional refinement. Notably, very few tasks could be repaired with more than a single retry, which suggests that while a simple retry effectively corrects minor implementation errors, persistent failures are likely due to fundamental misunderstandings that cannot be resolved through additional attempts. Consequently, for tasks where functional correctness is achieved, the focus naturally shifts to optimizing other quality dimensions, such as energy efficiency, which we investigate in the following section.

Model	**Tasks**	**Solved (%)**	**1-attempt (%)**	**Fix attempts**		
				0	1	2
GPT-4.1	44	88.6	70.5	31	7	1
GPT-4.1-nano	44	81.8	61.4	27	8	1

Tab. 1: Summary of task outcomes: total attempted, percentage solved, single-attempt rate, and repair attempts required (0–2).

4.2 Energy Consumption (Efficiency)

One of the core capabilities of EnerAgent lies in its ability to iteratively improve code with respect to energy efficiency. As shown in Tab. 2, both model variants successfully reduced the energy consumption of their initial functional solutions in a significant portion of tasks. This demonstrates that even when an LLM can already produce functionally correct code, there is often considerable room for optimization, which may initially be discarded by prioritizing a simpler, but less efficient solution. The fact that over half of the tasks benefited from optimization reinforces the importance of a refinement stage for improving the quality of generated code.

The histogram in Sect. 4.2 visualizes the distribution of energy improvements across optimized tasks. Most improvements fall within the 10–20% range. While some degree of

Model	**Successful opt.**	**Mean opt. attempts**	**Mean RES vs. Initial (%)**
GPT-4.1	28	4.86	23.25
GPT-4.1-nano	18	3.17	18.54

Tab. 2: Optimization outcomes: success count, attempts, and RES vs. Initial code.

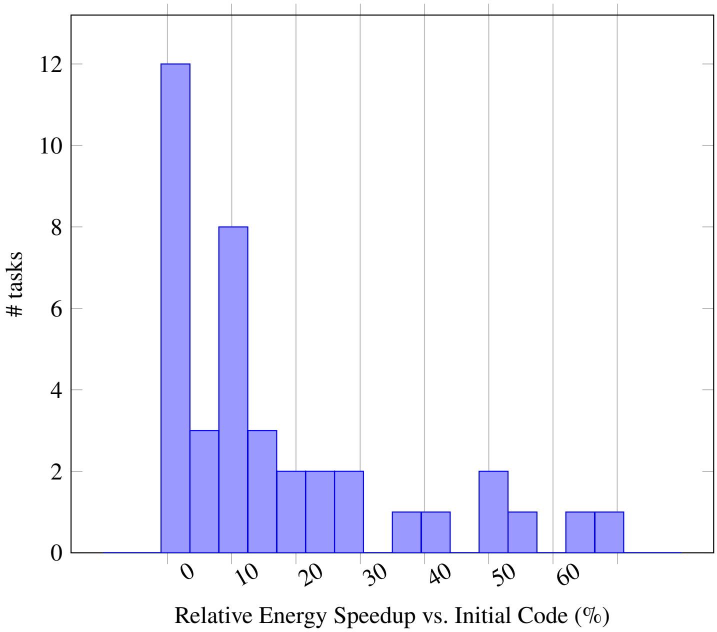


Fig. 2: Distribution of Relative Energy Speedup (RES) between initial and final agent solutions. Only functionally correct solutions are included.

the measured improvements may be explained by observational error, manual inspection of selected solutions indicated that they often stemmed from low-level optimizations tailored to the specific test conditions. These changes did not affect algorithmic complexity but instead reduced runtime overhead, such as avoiding unnecessary memory allocations or unrolling short loops. This can yield measurable speedups on small input sizes. However, these optimizations may not generalize and could even degrade performance on larger inputs due to reduced scalability.

More substantial improvements exceeding 30% were also observed, with some tasks showing energy reductions greater than 60%. These larger gains typically stem from the correction of more severe inefficiencies in the initial solution. In such cases, the model often replaced suboptimal algorithms or inappropriate data structures with more efficient alternatives, leading to the best possible runtime and memory complexity to solve the problem with.

These examples underscore the agent's ability to not just polish code, but to meaningfully rework it for greater efficiency when needed.

We emphasize that reported energy savings reflect relative differences under controlled conditions and are intended primarily as a feedback signal for guiding agent behavior, as explained in Sect. 3.1. These measurements may not be accurate estimates of absolute energy use. A more precise evaluation of true energy savings would require re-running optimized code independently, over extended durations, using external measurement tools—an avenue we leave to future work.[4]

4.3 Energy Efficiency Relative to Human Solutions

To assess the energy efficiency of the agent-generated code in a broader context, we compare its performance against human-written reference solutions. Tab. 3 shows the percentage of energy consumed by the final agent-generated code relative to the human implementation, considering only those tasks where the agent produced a functionally correct solution.

Model	Tasks	Mean RES vs. Human (%)	Median RES vs. Human (%)
GPT-4.1	39	16.53	10.00
GPT-4.1-nano	36	12.88	8.32

Tab. 3: Relative Energy Speedup (RES) of final agent solutions compared to human-written code.

These results suggest that the code generated through multiple rounds of refinement is competitive with human implementations in terms of energy consumption. However, they must be interpreted with nuance. In most cases, the final agent solutions converged to implementations of similar algorithmic complexity as their human-written counterparts. The observed energy savings primarily stemmed from small optimizations that exploited the limited input sizes used during testing (as discussed in Sect. 4.2).

In a few tasks, the agent achieved substantially greater efficiency by intentionally favoring brute-force implementations that performed well on the tested input sizes. Although these approaches are theoretically suboptimal, they incurred lower overhead and, in some cases, led to energy reductions exceeding 50%. These gains reflect input-aware simplifications rather than algorithmic superiority.

These findings highlight both the potential and the contextual sensitivity of LLM-generated code. While the agent is capable of producing highly efficient code, its optimizations are often tuned to the specific input characteristics present during measurement. This underscores the importance of careful benchmarking design when evaluating LLM-driven code generation under real-world conditions.

[4] All programs typically execute in under one second. Resulting energy values fall in the millijoule to low-joule range. Task-level absolute measurements and LLM-generated programs are available in the project repository.

5 Discussion

Our results show that EnerAgent can autonomously produce functionally correct and energy-efficient C++ code. Even within a limited benchmark, several insights emerge.

First, modern LLMs like GPT-4.1 achieve high success rates on competitive programming tasks, with over 70% correct on the first try. Many remaining failures are minor and often fixed with one or two repair attempts, highlighting the advantage of iterative, feedback-driven approaches over single-attempt generation.

Second, the optimization loop yields tangible energy gains. Over half of functionally correct solutions were improved, with average savings over 20%. While some reductions fall within measurement noise, manual inspection confirmed they often reflect subtle, but meaningful code changes.

Third, agent-generated solutions frequently match or surpass human-written code in energy usage. In some cases, the agent exploited specific test conditions (e.g. small inputs) to apply simple, low-overhead solutions, such as brute-force approaches that outperformed more general human solutions. While such behavior raises concerns about generalization beyond test conditions, it illustrates the agent's ability to adapt to specific runtime contexts.

EnerAgent demonstrates how LLM agents can serve as practical tools for sustainability-focused software development.

References

[CC25] Chase, H.; Contributors, L.: LangGraph: A low-level orchestration framework for controllable LLM agents, version 0.4.8, Python package, retrieved June 8 2025, 2025, https://github.com/langchain-ai/langgraph.

[CRO24] Cappendijk, T.; de Reus, P.; Oprescu, A.: Generating Energy-efficient code with LLMs. arXiv preprint arXiv:2411.10599, 2024.

[CSL+24] Courty, B.; Schmidt, V.; Luccioni, S., et al.: mlco2/codecarbon: v2.4.1, version v2.4.1, 2024, https://doi.org/10.5281/zenodo.11171501.

[Hu24] Huang, D. et al.: Effilearner: Enhancing efficiency of generated code via self-optimization. Advances in Neural Information Processing Systems 37, pp. 84482–84522, 2024.

[Ji24] Jiang, J. et al.: A survey on large language models for code generation. arXiv preprint arXiv:2406.00515, 2024.

[Ma20] Mann, B. et al.: Language models are few-shot learners. arXiv preprint arXiv:2005.14165 1, p. 3, 2020.

[Ni24] Niu, C. et al.: On evaluating the efficiency of source code generated by llms. In: Proceedings of the 2024 IEEE/ACM First International Conference on AI Foundation Models and Software Engineering. Pp. 103–107, 2024.

[Pe24] Peeperkorn, M. et al.: Is temperature the creativity parameter of large language models? arXiv preprint arXiv:2405.00492, 2024.

[Re24] Renze, M.: The effect of sampling temperature on problem solving in large language models. In: Findings of the Association for Computational Linguistics: EMNLP 2024. Pp. 7346–7356, 2024.

[SWC25] Solovyeva, L.; Weidmann, S.; Castor, F.: AI-Powered, But Power-Hungry? Energy Efficiency of LLM-Generated Code. arXiv preprint arXiv:2502.02412, 2025.

[SYL24] Shi, J.; Yang, Z.; Lo, D.: Efficient and green large language models for software engineering: Vision and the road ahead. ACM Transactions on Software Engineering and Methodology, 2024.

[We22] Wei, J. et al.: Chain-of-thought prompting elicits reasoning in large language models. Advances in neural information processing systems 35, pp. 24824–24837, 2022.

[Ya23] Yao, S. et al.: React: Synergizing reasoning and acting in language models, 2023. URL https://arxiv. org/abs/2210.03629, 2023.

Appendix: LLM Prompts

LLM Single-Attempt

You are an expert competitive-programming C++ developer. Write clean, efficient, code that follows the template exactly. Only produce C++ code — no commentary.

Reflection Fix Prompt

A programmer tried to solve a problem using C++ code, but the code failed to compile or failed functional tests. In the following you will be presented with a description of the problem, the incorrect code, and the compiler or runtime error logs. Your task: Find the errors in the code and fix them so the code compiles and passes all tests. Keep the same public interface and class structure. Do not add a main() function.

Reflection Optimize Prompt

A programmer solved a problem using C++ code. The code is functionally correct, but may be inefficient. In the following you will be presented with a description of the problem and the code. Your task: Optimize the code to make it as runtime and memory efficient as possible. Make sure the code remains functionally correct and keeps the same public interface and class structure. Do not add a main() function. Other than that you are free to change the code in any way you think increases efficiency. Think step by step how to make the code more efficient and then write the new improved code. If no improvements are possible, you may output `NO_CHANGE`.

Requirements Study for a Professional Green Coding Course Format

A Focus Group Study to Explore Strategies for Implementing Green Coding Training in Software Development

Dennis Junger [1] and Volker Wohlgemuth[1]

Abstract:

Although Green Coding (GC) is increasingly discussed in academia and, to some extent, within the industrial sector, it has not yet become a standard part of software development across the IT industry. GC focuses on reducing the environmental impact of software throughout its lifecycle, from production to use and disposal. To achieve this, it is essential to integrate GC knowledge and practices into the software development process. Previous work has explored integration into academic curricula and how junior developers can introduce this knowledge into the industry. In contrast, this study focuses on bringing GC knowledge directly into the industrial environment, addressing key challenges and discussing suitable training formats.

The methodology applied was a semi-structured expert interview conducted in the form of a focus group during a workshop at EnviroInfo 2024 at the GUC in Cairo.

The expert workshop revealed that the successful adoption of Green Coding depends on its integration into existing software development processes, tools, and training structures. Among various strategies discussed, training emerged as a central success factor. In-house training was considered the most effective format, as it enabled interactive learning, peer exchange, and a stronger commitment to behavioral change. Professional gatherings and workshops also play a crucial role in fostering expert exchange and continuous knowledge transfer. While online courses and webinars offer scalability, they were perceived as less impactful unless combined with certification or interactive components. Certificates, particularly those recognized by industry, were viewed as valuable incentives that could increase the perceived relevance of GC skills in the job market. However, their success depends on the recognition and integration of the broader market into recruitment practices. The integration of GC topics into academic curricula and technical training programs, either as standalone courses or embedded content, was highlighted as an essential long-term measure. Participants emphasized the need for regular updates to the content of GC training, ideally driven by expert communities, conferences, and changes in development frameworks. Overall, the workshop participants advocated for a flexible, multi-format training strategy. Combining in-house workshops, formal certification, academic courses, and integration into daily tools and processes ensures that GC knowledge reaches both current professionals and future developers. Training must be practical, context-sensitive, and regularly updated to remain effective and support a sustainable transformation in the software industry.

Keywords: Green Coding, Software Engineering, Requirements, Workshop Review, Focus Group, Green IT, Green Computing, Sustainable Software Engineering, Course, Training

[1] HTW Berlin, Environmental Informatics Unit, Wilhelminenhofstr. 75A, 12459 Berlin, Germany, dennis.junger@htw-berlin.de, https://orcid.org/0000-0002-9271-2261; volker.wohlgemuth@htw-berlin.de

doi:10.18420/env2025_10

1 Introduction and Motivation

The ’Potentials of Green Coding’ [In24] research project explored environmentally sustainable programming practices and their integration into software development processes and educational curricula. This study included a comprehensive literature review, analysis of industry practices, and validation through workshops and expert interviews. In the following project, ‘Unlocking Potentials of Green Coding into Practice’, the recommendation to put these mechanics into practice was used to develop them. The third research part of the project addresses the lack of GC knowledge among practitioners, identified in previous studies, as [He23] or also in part of the Literature Report as [Ju23]. However, the research identified practical challenges, emphasizing the need to train not only junior developers but also those already working in the industry. The previous research was focused on integrating the knowledge of GC in possible university curricula as in [Ju25a], due to the most (71, 5%) of software development professionals worldwide who had a university degree (41% bachelor’s, 25.6% master’s and 4.9% professional degrees like Ph.D., while 12.6% had some university education without earning a degree. In contrast, 11.6% of developers entered the field without a university background [St24]. This resulted in the creation of ’Academic Guidelines and Recommendations for Integrating Green Coding Aspects into Curricula of Higher Education’ [Ju25a], which outlines materials and methods for embedding GC into the education of future software developers. The first step was to find out what knowledge is most lacking for young professionals with a GC background (either educational or self-taught [JW24] to A) develop a training concept that can be applied broadly across Information and Communications Technology (ICT) and software development fields and B) to bring this together with the methodologies and content of GC in the professional field through training.

This paper is a step in the process where we aim to evaluate possible materials and a training concept to address this knowledge gap. We strive to evaluate and identify the most effective training formats. To find and create a training form, training methodologies, and content, and in the third step by developing a training course that is implemented and evaluated on a test basis. The aim is to make this training available to the public.

The article is structured first to present the motivation and the current state of research. This is followed by an explanation of the methodology and tools used for the focus group. The evaluation section discusses the findings. Based on these results, clear recommendations are provided in the conclusion section for integrating GC into training programs.

2 Methodology

To investigate the opportunities and challenges of integrating GC methods into software engineering, a focus group [Ki95] approach was selected as the central qualitative methodology. This choice was based on the need for in-depth, interactive discussions that could reveal nuanced perspectives on training formats, content relevance, and practical implementation issues across different sectors and roles. The authors followed the approach ideas of [Ro20] and [MM].

The Chair-Guided Structured Expert Interviews, particularly the Focus Groups method, were chosen for several reasons:

- It allows direct exchange between participants from diverse backgrounds, enabling the exploration of varied expectations and experiences.
- It is well-suited to early-stage idea development, especially for a complex and evolving topic like GC.
- It supports a collaborative setting, in which group dynamics can surface shared challenges and innovative solutions that individual interviews might miss.

The primary aim was to collect input for the development of an internationally applicable GC training course. Therefore, the group was formed during an international conference, ensuring geographic and professional diversity. There were 3 online and 10 in-house participants, including:

- Doctoral students
- Academics with links to industry
- Full-time researchers
- Industry practitioners
- Policymakers with ties to sustainable digitalization

Selection criteria focused on assembling a mix of individuals already engaged or interested in sustainable software development, ensuring both depth and relevance in the discussion.

2.1 Acquisition Methodology

The Focus Group was conducted on Wednesday, 13.11.2024, on the EnviroInfo Conference [Ge24] in Cairo as the last session of the ’Green Coding Techniques and Methodologies’ track. The track contained 6 paper presentations before, and the workshop was announced throughout the day. The workshop was titled: ’Workshop Unlocking the Potentials of GC into Practice - Focus Groups on: A) Labeling Schemes as Tools for Software Sustainability B) Opportunities and Challenges of Integrating GC Methods in Software Engineering Processes’ and contained the first three questions for Topic A, corresponding to the research package of the Umweltcampus Birkenfeld, presented and moderated by Mr. Westing followed by the three questions presented by the author for the working package of the HTW Berlin. The focus groups were time-boxed to 5 minutes each question.

2.2 Evaluation Methodology

The focus group was overseen by two individuals, one of whom was responsible for facilitating the session in person, while the other oversaw the online segment of the event. To ensure clarity and inclusivity for all participants, microphones were the sole form of communication, distributed to the speakers for their use. The video and audio excerpts for this presentation were recorded by the event organizer. The present recording has undergone conversion from the video format MP4 to the audio format MP3. The current iteration of the MP3 has undergone a reduction in length, with the result that only questions 4-6 of [Ju25c] are addressed. These questions (3.1 to 3.3) are the ones pertinent to the present document. The transcription of the file was subsequently executed using the Turboscribe [Tu25] page, with the speakers' details and time stamps recorded for each speaker. Subsequently, the data was imported into Google NotebookLM [Go25] as the only source so that the conversation tendency and the main topics could be analyzed and identified. In addition, a text file was added with the information on the PowerPoint slides used during the focus group as a reference.

3 Evaluation and Discussion

The complete anonymized transcript of the audio and the transcriber's notes can be found in the 'supplementary material A' at **https://doi.org/10.5281/zenodo.15729451** [Ju25c]. We also provide a mindmap addressing all the following topics under **https://doi.org/10.5281/zenodo.15728863** [Ju25b]. Even though the evaluation of the focus groups deals exclusively with the opinions and ideas of the participants, additional references have been inserted by the author at unclear points in order to provide the reader with a clearer understanding of the context discussed.

3.1 Challenges and Strategies in Integrating Green Coding Practices

The integration of GC practices into existing software development processes presents several significant challenges for companies, alongside potential strategies to overcome these obstacles. This discussion synthesizes expert perspectives on these issues. The Question asked to start this time-box was: 'What are the main challenges that companies face when integrating Green Coding practices into existing software development processes, and what strategies can be used to overcome these obstacles?'.

3.1.1 Main Challenges Identified:

A primary challenge is the lack of awareness among software developers regarding the energy consumption implications of their code and underlying infrastructure. Specifically, there is a lack of knowledge about inefficiencies in data centers, such as the significant idling of servers. Furthermore, companies often prioritize value creation and rapid revenue generation over sustainability initiatives. Sustainability may not be embedded within the core corporate strategy, remaining merely a recommendation rather than a directive. Another obstacle is the difficulty for companies to train their sophisticated models (e.g., for code classification or assessment) due to limited capacity. There is a clear need for accessible pre-trained models. While knowledge about GC exists, it is not easily accessible or consolidated in one central location. This fragmentation makes it hard for developers to find and apply relevant information. Finally, sustainability is often not yet treated as a formal functional or non-functional requirement within the software development process.

3.1.2 Strategies for Overcoming These Obstacles:

Several strategies were proposed to address these challenges:

Enhance Awareness: Actively work to create awareness among developers about the importance and impact of GC. Providing catchy examples that demonstrate the significant scaling effect of software on energy consumption can highlight why developers should invest effort in this area.

Leverage Gamification: Introduce gamification elements, such as challenges, high score boards, or transparent 'blackboards', to motivate developers and foster healthy competition in optimizing code for energy efficiency. Even when the part of the focus group in this context was quite brief, the author refers to the positive outcomes of gamification in other learning environments, as described in [BRW19].

Develop and Integrate Tooling: Tooling is considered a crucial strategy, particularly focusing on measurement. Tools can provide developers with real-time feedback on code 'greenness' directly within their Integrated Development Environment, offering recommendations for improvement. Tools can also assess the energy consumption of software architectures and suggest optimizations. Solutions such as co-pilot [Mi25] were mentioned as potentially contributing to energy reduction. Providing visualization of Key Performance Indicators related to energy consumption can make progress transparent for both developers and management.

Provide Accessible Pre-trained Models: Publishing pre-trained models related to GC (e.g., for code analysis or assessment) on accessible platforms like Machine Learning as a Service or AI as a Service hubs (similar to Hugging Face [Hu25]) would allow companies to reuse them without needing to train their own. Such hubs could include KPIs on the model's effectiveness or environmental sustainability to aid in selection. This promotes code reusability and scaling effects.

Establish a Reputable Body of Knowledge: Consolidate existing knowledge on GC practices into a reputable 'Body of Knowledge'. This knowledge should be easily accessible, possibly through comprehensive guidelines similar to web standards for green web development. Concepts like 'Energy-Safe by Design' or 'Low Energy Consumption by Design' and sustainable design patterns that offer solutions to recurring energy consumption problems could be developed and disseminated.

Integrate Sustainability into Company Strategy: Elevate sustainability from a recommendation to a mandatory component of corporate strategy.

Define Sustainability as a (Non-)Functional Requirement: Formally incorporate sustainability as a (non-)functional requirement within the software development process (as suggested for example, by [Ra14]). This includes integrating it into the requirements engineering phase. Utilizing artifacts from models like a Green Software Engineering Model (e.g., the GREENSOFT Model [NKD14]), such as sustainability journals or introspection techniques, can support this integration within methodologies like Scrum or traditional processes.

Integrate Green Coding into Existing Frameworks: Embed GC patterns, philosophy, and practices directly into widely used software development frameworks. Updating these frameworks would force developers to learn and apply GC principles, ensuring the content remains current.

3.1.3 Results of the Expert Voting:

The participants voted on what topic would be most important to focus on in the integration process. The options were soft skills (change management), measurement (tools), and software engineering process information. In the vote, five participants were in favor of Measurement (Tools) and seven in favor of Software Engineering Process Information, while Soft Skills (Change Management) received one vote.

This indicates that the experts consider the areas of tooling (measurement) and integration into the software development process (software engineering process information) to be particularly important for overcoming the challenges and implementing strategies in GC.

3.2 Possible Training Formats

Training programs for GC can be tailored to address the specific needs and challenges faced by different roles in software engineering, including developers, project managers, and quality assurance engineers. The primary objective of these programs is to support the practical integration of GC practices within companies. Starting with the focus group question, 'How can training programs be designed to address the specific needs and challenges faced by different roles in software engineering, such as developers, project managers, and quality assurance engineers? What format is preferred?', multiple formats and approaches were proposed to ensure effective delivery of the training:

3.2.1 Suitable Training Formats as Described in the Focus Group

Face-to-face and In-house Training: This approach is considered highly effective for fostering a shared understanding and cultivating a cultural shift towards GC within an organization. Participants are more likely to internalize the importance of the topic when engaging in face-to-face interactions, workshops, and discussions, which can also help address anxieties and encourage a collective commitment. In-house training specifically provides an environment where individuals can learn alongside like-minded colleagues, enhancing the learning experience. It is also crucial for sensitizing employees and convincing leadership about the benefits of GC.

Certificates and Online Tests with Certifications: The provision of certificates is seen as a strong motivator for individuals to actively engage with and study the content of the GC, rather than simply passively consuming information. Certificates can also enhance a professional's curriculum vitae and potentially encourage recruiters and companies to prioritize candidates possessing certified GC skills, although motivating recruiters to focus on such certifications remains a challenge. A formal certification program, such as becoming a 'certified green coder', could serve as a foundational step.

Conferences and Gatherings: These events offer a valuable platform for disseminating knowledge, sharing methodologies, and presenting up-to-date research findings related to GC. They facilitate networking and information exchange among specialists from different companies and research institutions, promoting a wider adoption of GC practices.

Integration into Academic Curricula: From an academic perspective, it is recommended that GC topics be systematically integrated into existing university curricula for software engineering and DevOps. Offering GC courses as electives can provide students with optional pathways to acquire these essential skills.

Webinars and Online Courses: While mentioned as a potential format, some participants expressed concerns that online-only methods, such as clicking through slides, might not effectively convey the deeper significance or foster a tangible 'feeling' for the topic.

3.2.2 Results of the Voting Regarding Training Formats

During the discussion on preferred training formats, a vote was conducted offering the following options: Webinar / Online Course, In-house Training, and Online Test with Certificate. The results indicated a clear preference for in-house training, which received the highest number of votes (7). This was followed by spontaneously added conferences and gatherings (6 votes), certificate-based online tests (3 votes), and webinars or online courses (2 votes).

These results reflect a general inclination towards interactive and presence-based learning formats.

3.3 Integrating and Updating Green Coding in Technical Training

This focus group section aimed to identify strategies for meaningfully integrating GC training into existing technical training programs and ensuring its long-term relevance. The following sections summarize the discussed approaches.

Key guiding questions included: 'How can Green Coding training be integrated into existing technical training programs without overwhelming developers with additional requirements?' and 'How can we ensure that GC content remains up to date? What are suitable forms of integration?'

3.3.1 Ideas of the Focus Group of possible Integrations

Integration into Existing Technical Training Programs: Training programs for GC can be integrated into existing technical training programs and maintained for up-to-date content through several proposed strategies, to prevent developers from being overwhelmed by additional requirements.

Leveraging Development Frameworks for Green Coding Principles: A particularly effective strategy involves embedding GC patterns, foundational philosophies, and related principles directly into existing development frameworks that are widely utilized by software professionals. This method inherently encourages developers to adopt coding and thinking practices aligned with GC, as updates to the framework would necessitate continuous learning and adherence to current GC standards.

Motivating Developers through Practical Examples and Impact Awareness: It is crucial to present compelling and impactful examples to developers, illustrating the practical benefits of investing in GC, especially considering the substantial scaling effect of software. By fostering sufficient awareness of these advantages, developers may be intrinsically motivated to proactively seek out and engage with relevant GC material independently.

Ensuring Green Coding Content Stays Up-to-Date: To ensure the continuous relevance and currency of GC content, conferences and specialized gatherings are identified as vital platforms for disseminating the latest research findings and methodological advancements. These events could feature dedicated 'GC for specialists' days or tracks, fostering valuable exchange and collaboration among experts from various companies and research institutions.

Role of Conferences and Expert Gatherings: Conferences and professional gatherings offer essential opportunities to share current findings, methodologies, and best practices. The inclusion of specialized sessions focused on GC ensures ongoing engagement with the latest developments.

Framework Updates as a Mechanism for Continuous Learning: The integration of GC principles into actively maintained and updated development frameworks inherently contributes to the currency of the content, as developers would continuously learn about new GC developments as part of routine framework updates.

3.3.2 Fitting Integration Forms (Vote Results)

In the discussion on suitable integration forms for GC training, a vote was conducted with the following options: Stand-alone Training, Full Integration into Courses, Other Options, and Both/All Options. The most favored approach was Both/All Options, receiving six votes. Full Integration into Courses followed with four votes, while Stand-alone Trainings received three votes. Other Options did not receive any votes.

These results suggest a preference for flexible and combined training approaches.

4 Conclusion

The findings of this expert workshop highlight a clear need for the structured integration of GC practices into the broader software engineering domain. Experts emphasized that successful implementation depends not only on raising awareness but also on embedding GC principles into existing software development processes, tools, and training programs. The voting results underline the importance of practical integration—particularly through updates in development frameworks and the use of measurement tools—as well as the relevance of software engineering process information. Tooling and process alignment are seen as enablers that make sustainable development tangible and manageable in daily practice.

Training formats emerged as a key success factor. Interactive and presence-based formats, especially in-house training and professional gatherings, were strongly preferred over passive online formats. These approaches are more effective in creating shared understanding, encouraging behavior change, and fostering a sense of ownership. Certification programs were also mentioned as valuable motivators, though their effectiveness depends on market acceptance. Academic integration through elective courses and embedding GC content into existing technical curricula were seen as strategic entry points for long-term cultural change in the industry.

To remain relevant, GC content requires regular updates. This can be supported by expert communities, conferences, and knowledge sharing. Integrating GC principles into widely used software development frameworks helps ensure that developers encounter sustainable practices as part of their routine work, making sustainability a standard consideration in software engineering.

Ultimately, the discussions and results indicate that there is no one-size-fits-all approach. A flexible combination of training formats and integration strategies—tailored to different organizational and individual needs—was the most supported direction. GC must become a cross-cutting concern that is interwoven with existing roles, responsibilities, and workflows in software development. Only then can its potential be fully realized to reduce the environmental impact of digital technologies in a scalable and lasting manner.

Acknowledgments

This work was funded by the Internet Society Foundation[2]. The authors extend their sincere thanks to the EnviroInfo Conference at the German University in Cairo for providing the opportunity to conduct this expert interview as part of the workshop.

References

[BRW19] Bernik, A.; Radošević, D.; Wattanasoontorn, V.: Influence, principles and good practice of computer game elements: Mechanics and Dynamics. In: 2019 23rd International Computer Science and Engineering Conference (ICSEC). Pp. 241–245, 2019.

[Ge24] German University in Cairo: EnviroInfo 2025- Transboundary Environmental Challenges: Digital Inclusion for Sustainable Development, 2024, https://enviroinfo.org/, accessed: 06/07/2025.

[Go25] Google: NotebookLM, 2025, https://notebooklm.google.com/, accessed: 06/05/2025.

[He23] Henze, C. et al.: Potentials of Green Coding: Perstepctives on Implementation, Berlin, 2023, https://gi.de/meldung/green-coding-publikation, accessed: 06/05/2025.

[Hu25] Hugging Face, I.: Hugging Face Hub, https://huggingface.co/, Accessed: 2025-07-29, 2025.

[In24] Informatik, G. f.: Green Coding - Gesellschaft für Informatik e.V, 2024, https://gi.de/aktuelles/projekte/green-coding, accessed: 08/16/2024.

[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.

[Ju25a] Junger, D. et al.: Academic Guidelines and Recommendations for Integrating Green Coding Aspects into Computer Science Curricula of Higher Education. In: Recent Advances in Environmental Science from the Euro-Mediterranean and Surrounding Regions (6th Edition). Accepted for publication, Marrakesh, Morocco, 2025.

2 https://www.isocfoundation.org/ [2025-06-06]

[Ju25b] Junger, D. M.: Mind Map Requirements Study for a Professional Green Coding Course Format, 2025, https://doi.org/10.5281/zenodo.15728863.

[Ju25c] Junger, D. M.: Requirements Study for a Professional Green Coding Course Format - Supplementary Material A, 2025, https://doi.org/10.5281/zenodo.15729451.

[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.

[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.

[Mi25] Microsoft: Copilot (GPT-4) [Large Language Model], https://copilot.microsoft.com/, Accessed: 2025-07-29, 2025.

[MM] MMG Connect: Focus Group Brief, http://www.mmgconnect.com/projects/userfiles/file/focusgroupbrief.pdf, Accessed: 2025-07-29.

[NKD14] Naumann, S.; Kern, E.; Dick, M.: Classifying Green Software Engineering - The GREENSOFT Model. Softwaretechnik-Trends 33, pp. 18–19, 2014.

[Ra14] Developing a sustainability non-functional requirements framework, GREENS 2014, Hyderabad, India: Association for Computing Machinery, 2014, pp. 1–8, https://doi.org/10.1145/2593743.2593744.

[Ro20] Roberts, R.: Qualitative Interview Questions: Guidance for Novice Researchers. The Qualitative Report, 2020.

[St24] Stack Overflow and Amazon: Levels of Formal Education for Software Developers, Worldwide, as of 2024 [Graph], 2024, https://www.statista.com/statistics/793568/worldwide-developer-survey-level-formal-education/, accessed: 08/20/2024.

[Tu25] TurboScribe: TurboScribe - Unlimited audio & video transcription, 2025, https://turboscribe.ai/, accessed: 06/05/2025.

Resource Efficiency in Software: A Comparative Study of Two Text Editors

Lena Fischer [1], Marian Messer [2], Dennis Junger [3], and Volker Wohlgemuth [4]

Abstract: The following report documents a practical measurement to compare the two text editors Visual Studio Code (VS Code) and Pulsar. Due to the ongoing increase in available software products for a developer's toolkit, this study aims to compare the energy and hardware consumption to give developers a guideline which text editor to choose when considering energy and resource efficiency. In the used measurement setup, based on the requirements for the Blue Angel for Software, VS Code proves to be the more resource-efficient text editor, particularly in terms of CPU, RAM and hard disk usage. A key finding of this analysis is that VS Code, a software developed by a large corporation, is not more resource-intensive than Pulsar, a community-led open source project. The results indicate that VS Code is significantly more efficient in terms of CPU usage, memory load, and hard disk activity, both in absolute terms and after adjusting for system baseline consumption. Pulsar consumes more energy and system resources in the proposed usage scenario. A critical finding—observed during this and a second comparable study—concerns the use of Power Automate: energy measurements show that the automation tool itself may consume more power during idle phases than the editors during active usage. Hence, this study recommends alternative scripting tools for future experiments and careful validation of the baseline setup to avoid a distorted set of results due to the used measurement methodology. The findings point to the need for improved measurement setups in future studies, including longer measurement periods, more complex usage scenarios, and more efficient automation solutions.

Ultimately, the study aims to contribute to the field of sustainable software by delivering reproducible data and encouraging critical discussion about resource efficiency in everyday programming tools.

Keywords: Ressource Efficiency; Green IT; Open Source Software; Energy Consumption; Text Editor; Sustainability of Software; Green Coding

1 Thematic Introduction and Theoretical Background

The tension between economy and ecology is a well-known concept. In the software industry in particular, experience shows that ecology often loses out to economics. New software is entering the market, computers are becoming more and more power-intensive and data centers are already causing the same amount of CO_2 equivalents as global air traffic [As21]

[1] HTW Berlin, Environmental Informatics Unit, Wilhelminenhofstr. 75A, 12459 Berlin, Germany, lena.fischer@student.htw-berlin.de, https://orcid.org/0009-0001-8030-624X

[2] HTW Berlin, Fachbereich 2, Wilhelminenhofstraße 75A, 12459 Berlin, Deutschland, marian.messer@student.htw-berlin.de, https://orcid.org/0009-0009-9861-5401

[3] HTW Berlin, Fachbereich 2, Wilhelminenhofstraße 75A, 12459 Berlin, Deutschland, dennis.junger@htw-berlin.de, https://orcid.org/0000-0002-9271-2261

[4] HTW Berlin, Fachbereich 2, Wilhelminenhofstraße 75A, 12459 Berlin, Deutschland, volker.wohlgemuth@htw-berlin.de

doi:10.18420/env2025_11

in 2021. This topic is addressed by the subject area of green IT, which was explored in greater depth as part of a newly developed module at HTW Berlin [JW23]. The presented measurement setup, based on the methodology of [JW22], is intended to build on this idea and provide information about sustainable, green IT practices through the detailed analysis of the energy consumption and hardware utilization[Ju23], which has been growing steadily for years, with data and its interpretation. Since text editors are part of the daily toolbox this study aims to offer transparency on a software's energy and resource efficiency, giving developers the opportunity to choose sustainable software in their day-to-day work. This report also aims to tackle the topic of open source software and the associated sustainability potential.

1.1 Importance of Text Editors for Software Development

Technologies in software development are subject to constant change. In order to be able to program efficiently and clearly, Integrated Development Environments (IDEs) with a variety of integrated functionalities gained in importance as early as the 1990s and 2000s [Ka24]. While IDEs can still be found in the software landscape of most developers today, another type of coding software has gained importance: lightweight text editors. Text editors such as VS Code, Sublime Text or Notepad++ offer a lean basic framework of functionalities compared to heavyweight IDEs. Required functionalities can then be added as required using extensions and plug-ins.

1.2 Open Source Approach

Open source approaches in software development have been a common practice for decades. While companies often keep software projects internal in order to protect business secrets or gain a competitive advantage, the open source concept illustrates a different approach. The code can be viewed and used by the public. In theory, this approach is intended to increase quality through its practically unrestricted peer review - i.e. the evaluation and improvement of the code by other developers [Ab07]. Since the comparison in this paper deals with two products with opposing approaches, the energy consumption could provide an indication of possible advantages or disadvantages of OSS in the case of lightweight software such as a text editor.

1.3 Frame of Reference

VS Code is one of the most widely used text editors on the market. It is a free software product for both private and commercial use that is developed and maintained by Microsoft. The basic version of VS Code is an intuitive text editor that allows you to open, edit and save source code files. Furthermore, the ‘out of the box’ version of VS Code includes some standard formatting templates such as the color coding of keywords in the code or automatic indentation of lines. However, the majority of the functionalities offered are not pre-installed during download, but are available via the extension marketplace. Both verified software publishers such as Microsoft and individual users can publish extensions there. These then enable a range of different functionalities, starting with the opening of file types such as PDF files, individualized colour themes and code formatting through to the integration of ChatGPT to support the programming process. For our comparison, the basic version of VS Code without extensions was used due to the comparability with the second software product.

In addition to VS Code, Pulsar was chosen as the comparison product for this practical measurement. This tex editor is an evolution of the text editor Atom, which has been retired since 2022 due to security risks [Gi25]. Atom’s popularity in the software development community gave way the software Pulsar, ‘A Community-led Hyper-Hackable Text Editor’[Pu25]. It was chosen for this study’s comparison because, unlike VS Code, it is not a software product developed by a corporation like Microsoft that takes over maintenance and further development. Section 1.2 has already raised important arguments in favor of using open source software, which is why Pulsar was chosen as the software to contrast with VS Code. All further developments are open source, the software itself and all updates are free of charge and the Pulsar code is publicly available on GitHub. Similar to VS Code, the download version of Pulsar contains a small number of different standard functionalities such as code highlighting and formatting. Like in VS Code, plugins can be integrated via a marketplace. For a valuable comparison, the ‘atom-terminal’ extension was selected for the integration of a terminal, as this was necessary for the proposed usage scenario. Apart from this, the out-of-the-box version of the software was used here in order to achieve a comparable result.

2 Approach and Methodology

The following sections describe the proposed usage scenario, structure of the measurement setup, and the final evaluation of the collected data using the O.S.C.A.R. software (https://oscar.umwelt-campus.de/).

2.1 Use Scenario

In order to keep the measurements of the two text editors as close to reality as possible, a standard usage scenario, consisting of basic interactions with the software, was selected. This was then executed in both VS Code and Pulsar and used for the later measurements. The scenario was then recorded by the automation software *Power Automate* (https://make.powerautomate.com/) and repeated 30 times in succession to obtain a representative result. Before the iterations are started, the variables *process time* and *start* *time* are created and saved. The following behavior diagram describes the sequence the individual interactions with the two text editors.

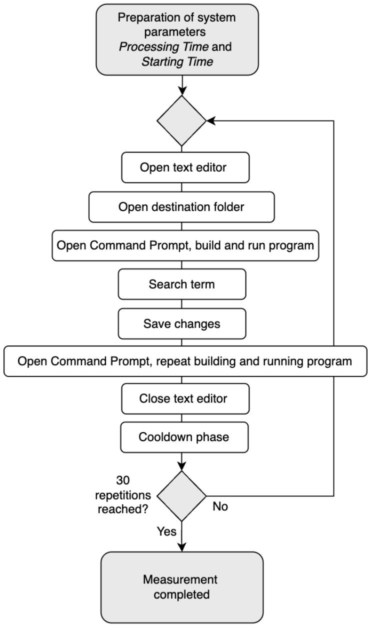


Fig. 1: Behavior diagram for usage scenario (own figure)

2.2 Measuring Setup

In order to obtain representative and comparable measured values, the measurement laboratory provided by HTW Berlin was used. The setup is based on Gröger et. al. [Gr18] and Junger et. al. [JW22]. For the evaluation, the hardware utilization of the System Under Test (SUT) was measured by the built-in Performance Monitor and saved for later processing. The action logs of the usage scenario were also saved. Finally, the electrical power consumption of the SUT ('DJ Darwin') was recorded and saved on a separate PC ('DJ Einstein') using a GUDE Power Control 1202 measuring device while the usage scenario was being played. In addition to the measurements of the usage scenario in both text editors, baseline measurements were recorded that do not contain any user interaction with the software.

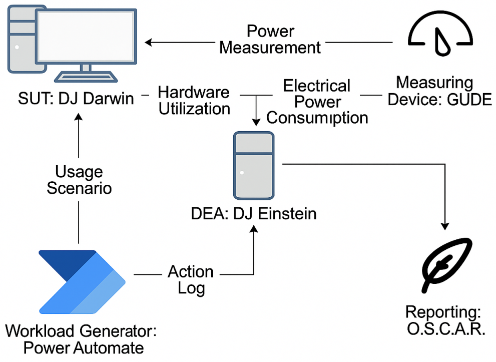


Fig. 2: Measurement setup (own figure, based on Gröger et. al.)

2.3 Evaluation using O.S.C.A.R.

The web version of the O.S.C.A.R. software, developed by the Umweltcampus Birkenfeld, was used to evaluate the recorded measurement data. It uses a defined methodology to generate a report on various factors related to the environmental impact of software. The hardware utilization, electrical power consumption and the action log described in the previous chapter are uploaded as input. This is done for the measurements of the usage scenario and for the baseline measurement. O.S.C.A.R. generates a report, which links the baseline and usage scenario measurements and provides a variety of different results.
Due to technical problems with O.S.C.A.R., not all measurement data could be used fully, meaning that the usage scenario for the final evaluation could only be evaluated with ten repetitions instead of the initially planned 30 repetitions.

3 Results and Evaluation

This chapter presents the results of the measurements for the VS Code and Pulsar text editors. Both the absolute values and the baseline-adjusted values are analyzed and finally evaluated.

3.1 Results VS Code

This section describes and analyzes the results of the VS Code measurements in particular.

3.1.1 Measurement Values

The measurements show that the resource utilization of VS Code varies depending on the usage scenario. During the first ten seconds in which the editor is opened, the average processor utilization is 3.78 %, while the memory consumption is 54.77 %. Network usage shows 2645.75 bytes of data sent and 3453.31 bytes of data received. The hard disk activity is 10,259.68 bytes for read accesses and 130,904.8 bytes for write accesses (see O.S.C.A.R. - report VS Code [FM25]).

Between the 20th and 30th second, in which the code is compiled and executed, an increase in CPU and hard disk utilization can be observed. During the thirtieth to thirty-fifth second, when a text search and replace is performed, CPU and RAM remain almost constant. The greatest hard disk activity occurs between the thirty-fifth and fortieth second when the file is saved.The energy consumption measurement for VS Code resulted in an average electrical work of 0.48 Wh in the usage scenario.

3.1.2 Corrected Measurement Values

After subtracting the baseline values, Visual Studio Code shows an additional CPU usage of 0.61%, and an increase in RAM usage of 7.82%. The corrected network traffic amounts to 2402.88 bytes sent and 2209.17 bytes received. Regarding hard disk activity, baseline-adjusted values result in 10,259.68 bytes for read operations and 130,904.80 bytes for write operations.

The corrected energy consumption is calculated at –0.12 Wh. This negative value is not physically plausible and indicates a possible distortion caused by the measurement setup – for example, background load introduced by the automation tool. Nevertheless, this value still reflects the relatively low energy consumption of VS Code within the measured usage scenario. (see O.S.C.A.R. - report VSCode [FM25]).

3.2 Pulsar Results

This chapter describes and analyzes the results of Pulsar's measurements in particular.

3.2.1 Measurement Values

The CPU utilization of Pulsar is 5.11 % on average, while the memory consumption is 61.50 %. The network usage is 1161.33 bytes sent and 2394.31 bytes received. The hard disk activity of 40,028.18 bytes for read accesses and 360,231.8 bytes for write accesses is significantly higher than with VS Code. An electrical work of 0.56 Wh was measured for Pulsar in the usage scenario. (see O.S.C.A.R. - Report [FM25]).

The higher CPU and RAM utilization is particularly noticeable when the editor is started. During the compilation process between the twentieth and thirtieth second, the hard disk activity remains significantly higher. The network load remains relatively constant in almost all scenario steps.

3.2.2 Corrected Measurement Values

For Pulsar, the baseline-adjusted CPU usage amounts to 2.01%, and RAM usage increases by 10.74%. The corrected network traffic is 1089.64 bytes sent and 1393.71 bytes received. The disk activity is 40,028.18 bytes for reads and 360,231.80 bytes for writes.

The energy consumption corrected for baseline is –0.03 Wh. As with VS Code, this negative value suggests that the measurement setup introduced inconsistencies, possibly due to overhead from the automation tool. However, when comparing the two editors, Pulsar still consumed approximately 0.08 Wh more energy than VS Code in the same scenario, which is consistent with the observed differences in CPU, RAM, and disk utilization. (see O.S.C.A.R. - report [FM25]).

Metric	**VS Code**	**Pulsar**	**Difference (Pulsar – VS Code)**
CPU usage (%)	0.61	2.01	+1.40
RAM usage (%)	7.82	10.74	+2.92
Network – sent (B)	2402.88	1089.64	–1313.24
Network – received (B)	2209.17	1393.71	–815.46
Disk – read (B)	10,259.68	40,028.18	+29,768.50
Disk – written (B)	130,904.80	360,231.80	+229,327.00

Tab. 1: Comparison of resource usage between VS Code and Pulsar

3.3 Discussion

The comparison shows that VS Code is more resource-efficient overall than Pulsar. The baseline-adjusted values confirm this trend: Pulsar requires more CPU and RAM resources and has higher hard disk activity. VS Code, on the other hand, shows increased network usage, which could indicate online services or synchronization processes.

Another possible technical influencing factor is the underlying architecture of the programs: While VS Code is based on Electron and implemented with TypeScript/JavaScript, Pulsar also uses the Electron platform, but originally stems from the Atom project (also developed in JavaScript), whose code structure is considered less efficient. Differences in garbage collection, thread handling, or rendering can therefore have a direct impact on the measured CPU and RAM utilization.

External background processes during the measurement as well as the selected usage scenario limit the general validity of the results. Further analysis with alternative scenarios and longer test durations would be necessary to derive broader conclusions.

The energy consumption in watt hours (Wh) calculated by O.S.C.A.R. was also used to evaluate the energy efficiency. The corrected values adjusted for the system baseline show slightly negative energy consumption for both text editors (VS Code: –0.12 Wh, Pulsar: –0.03 Wh). While not physically plausible, these values indicate measurement distortions, potentially caused by the automation tool Power Automate or inaccuracies in the baseline reference.

Nevertheless, it is noticeable that Pulsar consumed approximately 0.08 Wh more energy than VS Code, which corresponds with the observed differences in CPU, RAM, and disk usage. This supports the overall finding that VS Code demonstrates higher runtime efficiency within the chosen scenario.

However, it is important to emphasize that energy consumption alone does not determine software sustainability. While runtime efficiency is a relevant metric, sustainability also encompasses long-term maintainability, update frequency, hardware compatibility, licensing, and community engagement. Open-source software like Pulsar can offer significant advantages in these areas for instance, by allowing transparent code inspection, avoiding vendor lock-in, and supporting older systems longer.

Therefore, this study's findings should not be interpreted as a general assessment of proprietary software being more sustainable than open source alternatives. Rather, they highlight that in this specific test case, VS Code exhibited greater energy and resource efficiency. Broader sustainability assessments should include diverse tools, scenarios, and lifecycle phases to ensure a more holistic understanding.

4 Critical Reflection and Additional Findings

The evaluation results revealed that the use of Power Automate as an automation tool significantly influenced the energy consumption measurements, especially during idle phases. In some cases, the baseline consumption while the system was idle but Power Automate was active exceeded the consumption during actual test execution. This behavior introduces distortions in the calculated energy values, as baseline consumption is subtracted from the measured usage scenario.

These findings align with those from a related study [MR25a], [MR25b], which observed similar issues when Power Automate was used to automate energy measurements. One likely reason is that Power Automate introduces system load even when no active input is being performed, possibly due to inefficient polling or active waiting processes.

To improve future experiments and reduce external influences on the measurement data, we recommend using lightweight and script-based automation tools. Tools such as *Shell scripts*, *Actiona*, or *AutoHotkey* could be more resource-efficient and easier to control in terms of background load and processor utilization. In addition, they allow more precise timing and scenario control, which can improve reproducibility.

Further technical improvements could involve pinning processes to specific CPU cores or using real-time process monitoring to exclude side effects caused by system background processes.

5 Conclusion

The measurements carried out show that Visual Studio Code works more resource-efficiently overall than the open source editor Pulsar in the usage scenario examined. This applies both to the utilization of central system components (CPU, RAM, hard disk) and, with restrictions, to the measured energy consumption. Despite possible distortions due to Power Automate, the adjusted consumption values indicate a consistent efficiency advantage of VS Code.

One key takeaway of this study is that software developed by a large corporation is not necessarily more resource intensive than a community-led open source project. However, this observation is limited to runtime efficiency in a controlled usage scenario. Broader sustainability implications such as long-term maintainability, support for older hardware, transparency, and community engagement were not part of this evaluation and should not be inferred from the presented data.

Therefore, this case study does not contradict the potential of open-source software to contribute to sustainable software development. Rather, it highlights the need to assess sustainability based on multiple measurable criteria beyond energy and resource consumption alone.

For a more in depth analysis, the use of alternative automation tools instead of Power Automate is recommended to minimize possible distortions caused by background load. Measurement over longer periods of time or with more complex interaction patterns such as parallel editing, live debugging, or the use of plugins could also provide deeper insights into the resource dynamics of the editors.

Future studies could additionally examine aspects such as update frequency, system compatibility, and energy usage during software installation and maintenance. Including other open-source editors such as Sublime Text or Neovim would also help to contextualize the findings in a broader Green IT discourse.

Acknowledgments

This work was funded by the Internet Society Foundation[5].

References

[Ab07] Aberdour, M.: Achieving Quality in Open Source Software. IEEE Software 24 (1), pp. 58–64, 2007.

[As21] Association for Computing Machinery: Computing and Climate Change. Tech Briefs 1, 2021.

[FM25] Fischer, L.; Messer, M.: Zenodo Repository: Green IT: Software Measurements of Text Editors Visual Studio Code and Pulsar, 2025, https://zenodo.org/records/15050857, (last accessed on 19.03.2025).

[Gi25] GitHub Staff: Sunsetting Atom, 2025, https://github.blog/news-insights/product-news/sunsetting-atom/, (last accessed on 27.02.2025).

[Gr18] Gröger, J. et al.: Entwicklung und Anwendung von Bewertungsgrundlagen für ressourceneffiziente Software unter Berücksichtigung bestehender Methodik. Umweltbundesamt, 2018.

[Ju23] Junger, D. et al.: Potentials of Green Coding -Findings and Recommendations for Industry, Education and Science. INFORMATIK 2023 - Designing Futures: Zukünfte gestalten, 2023.

[JW22] Junger, D.; Wohlgemuth, V.: Conception and test of a measuring station for the analysis of the resource and energy consumption of material flow-oriented environmental management information systems (EMIS). EnviroInfo 2022, 2022.

[JW23] Junger, D.; Wohlgemuth, V.: Design and implementation of a lecture for teaching current Green Coding approaches and practices at the HTW Berlin. INFORMATIK 2023 - Designing Futures: Zukünfte gestalten, 2023.

[Ka24] Kapoor, A.: Analysis of Code Editors: Features, Evolution, and Impact on Software Development. IJCAM, 2024.

[MR25a] Melvana Mauladendra, I.; Rauch, S.: Resource consumption based usability study for the comparison between customer segmentation with KMeans clustering using tensorflow and scikit-learn implementation, 2025, https://zenodo.org/records/14991212, (last accessed on 24.06.2025).

[MR25b] Melvana Mauladendra, I.; Rauch, S.: Resource consumption based usability study for the comparison between customer segmentation with KMeans clustering using tensorflow and scikit-learn implementation, 2025, https://zenodo.org/records/14991151, (last accessed on 24.06.2025).

[Pu25] Pulsar-Edit: Pulsar Homepage, 2025, https://pulsar-edit.dev/, (last accessed on 27.02.2025).

[5] https://www.isocfoundation.org/ [2025-06-06]

Open Data, Open Science and Open Society for a Sustainable Future

The Automotive Manufacturing Industry as a Potential Twin Transformer: A Research-In-Progress Perspective

Guillermo Vazquez Varela[1] and Hans-Knud Arndt[2]

Abstract: Twin transformation offers an orchestrated solution to the challenges encountered in digitalization and sustainability transformation endeavours. Despite current research efforts of this concept, there is a scarcity of empirical evidence related to its foreseen benefits, its underlying twin transformation maturity, and its implication among the manufacturing sector. We propose the automotive manufacturing industry as a promising case due to its global economic, societal, sustainability and business model relevance. Therefore, we provide theoretical context of twin transformation, its role in the manufacturing sector as well as in the automotive manufacturing industry. Additionally, we pair it with our ongoing research based on an echeloned design science research, including a preliminary problem analysis, and highlight potential future research directions on twin transformation.

Keywords: Twin Transformation, Automotive Manufacturing Industry, Maturity Model, Case Study

1 Introduction

Due to the constant need to stay competitive while, simultaneously, complying with climate-neutral policies, organizations keep their digital and sustainability transformations as a top strategic priorities [Br24]. While on one side, the digital transformation aims to improve organizations and their strategies through the use of digital technologies with a focus on economic value creation [Vi19], on the other side, the sustainability transformation aims to improve organizations through a positive impact on the society and the planet [DM16]. Aiming to avoid a collision of both digital and sustainable interests, the twin transformation idea arose [Gr23].

The twin transformation offers a value creation harmonization of both undertakings where, on one end, the digital transformation entitles the sustainability transformation and, on the other end, the sustainability transformation reshapes the digital transformation [Gr23]. Despite the acceptance of its novelty, including, IT service providers experts [Br24], within the research of the twin transformation, as in its related term “twin transition”, there are still various open questions that scholars ask for attention [TPC25]. In particular, there is

[1] Otto-von-Guericke-Universität Magdeburg, Fakultät für Informatik, Arbeitsgruppe Wirtschaftsinformatik - Managementinformationssysteme -, Universitätsplatz 2, 39106 Magdeburg, guillermo.vazquez@ovgu.de
[2] Otto-von-Guericke-Universität Magdeburg, Fakultät für Informatik, Arbeitsgruppe Wirtschaftsinformatik - Managementinformationssysteme -, Universitätsplatz 2, 39106 Magdeburg, hans-knud.arndt@iti.cs.uni-magdeburg.de

doi:10.18420/env2025_12

a need for empirical evidence to support the twin transformation [Ca24], further development of its maturity model [Br24], as well as its further study in the manufacturing sector [Br24; Ca24; TPC25].

In order to expand the body of knowledge related to twin transformation, we propose to take this important step by investigating the automotive industry due to its economic, political, and environment relevance for society [Pi21] as well to its ongoing digital and sustainability transformations [Ce23; LRV21]. Its significance is demonstrated by the fact that alone the automotive industry is responsible for the total of 27 % of the CO2 emissions in the EU [Pi21]. Since the manufacturing activities are the core of its business model [MF10], we advocate to delve into this sector as a research starting point having as a primary goal to expand the twin transformation capability maturity model (TTCMM) by [Br24] within its operations dimension. Thus, we ask the research question: *What is the twin transformation maturity of the automotive manufacturing industry?*

In this paper, we proceed as follows: First, we present an overview of the twin trans-formation state of the art and deep dive into the TTCMM aiming to bring light to potential research avenues. Second, we present the current status of related research areas to the twin transformation in the manufacturing sector and advocate for an approach to expand the TTCMM in the operations dimension. And third and finally, we present the automotive manufacturing industry as a future use case, highlighting the planned methodology for our ongoing study based on echeloned design science research (eDSR) by [TWV24] including an introduction to the first echelon, problem analysis. This first echelon includes the analysis of one automotive company which provides a preliminary answer to our research question. Overall, within this paper, chapter two and three offer a theoretical background while the third chapter presents our chose methodology for the next steps of our research with an initial conceptual analysis. Although the present research idea is not fully executed or yet implemented, we aim to contribute to the current body of knowledge of twin transformation, that has become increasingly relevant for both practice and academia.

Based on the nature of the topic and related fields, our present work can be initially, but not limited, be referred to the following sustainability development goals of the UN: (9) industry, innovation and infrastructure, (12) responsible consumption and production & (13) climate action [Un25].

2 Twin Transformation

Since the beginning of the decade, the call for an interplay of digital and sustainability transformation has been made searching for an answer to the need of organizations and government to become more competitive, sustainable, and ethical [TPC25; Vi19]. The study of the interaction of these two endeavors has been seen through multiple lenses, including but not limited to twin transition, digitally-enabled sustainable transition, green IT and twin transformation [Gr23; STT24; TPC25]. Twin transformation has emerged interest by practitioners [Br24; Ch24; Gr23] since it presented a promising approach to the digital-innovation compromise due to its harmonization nature and value proposal. Twin transformation is defined as “value-adding interplay between digital and sustainability transformation efforts that improve an organization by leveraging digital technologies for enabling sustainability and leveraging sustainability for guiding digital progress” [Ch24].

Within the core investigation of the twin transformation, various approaches have been presented with the idea to land the concept, create the basis for theory development and therefore turn on the flame of this promising field [Br24; Ch24; Gr23]. To begin with, a twin transformation model was developed aiming to illustrate the interrelation of digitalization and sustainability inside organizations [Gr23]. Next, a twin transformation capability framework presented dynamic capabilities for organizations interested in twin transform [Ch24]. Furthermore, a TTCMM was crafted to examine maturity of the twin transformation of organizations based according to various maturity stages [Br24]. And finally, a twin transformation resource orchestration pyramid was recently tailored to describe the resource management processes throughout twin transformation [Bu25].

The TTCMM stands out due to the tangible implications to evaluate the twin transformation maturity of organizations, as in digital transformations [Be20], with the help of capabilities and it stablishes an initial map to the summit called “true twin transformer” [Br24]. The idea is straightforward, represented with a pyramid whose levels are the maturity stages. The maturity stages are, with a bottom-up perspective: (1) awareness of combining digitalization and sustainability, (2) twin transformation development, (3) twin transformation implementation and (4) true twin transformer [Br24]. The maturity stages are divided into capability dimensions that help to identify in which state the organization is and what actions they should take to move to the next maturity stage, such as for example the operations dimension focused on the value creation processes. Each organization can start and follow its own pathway based on its own maturity of digitalization and sustainability [Br24]. The pathways are: Digital transformation (DT) expert, sustainability transformation (ST) expert, or twin transformation newcomer [Br24].

Despite the novelty presented in the TTCMM, further investigation and implementation is essential such as the further development of the capability dimensions and the application of the maturity model in various sectors [Br24]. The manufacturing sector promises a valuable use case that could fulfill the need for further research and evidence for the operations dimension [ABA25; Br24] due to its focus on value creation, along with providing the required solid empirical evidence to support the twin transformation construct [Br24; Ca24].

In the next chapter we present the state of the art of the twin transformation in the manufacturing sector, a potential improvement to the TTCMM through the operations dimension and an overview of manufacturing sectors.

3 Twin Transformation in the Manufacturing Sector

The manufacturing sector is facing various challenges regarding the implementation and acceleration of its sustainability transformations derived by continuous pressure to reduce their environmental impact, align with the market requests, follow stronger policies and increase its value creation [STT24]. As a solution to this need, the digital transformation has been studied and implemented in the form of industry 4.0 [Re23], circular economy strategies [Uh22], digitally-enabled sustainable transition [Uh22], among others [ABA25]. In the previously mentioned attempts, the sustainability and digitalization interactions are focused on specific topics and not as an orchestrated transformation.

Since there is a lack of literature regarding twin transformation in the manufacturing sector, we expanded our initial literature scope to the term twin transition. With the introduction of twin transition in a first stage, a new perspective was presented in the manufacturing sector. In this regard, research has focused on the identification of skills like resilience and digital, technical expertise [Ho24], capabilities identification in the value creation processes [Kü24], and the development of frameworks for manufacturing systems [VR23] to mention a few.

The design and operations (DesOps) framework presented by [VR23] aims to improve the efficiency of manufacturing systems based on the idea of twin transition. It represents major progress in this area since, currently, manufacturers do not accelerate their digital transformation due to a more conservative mindset [Ca18]. At this point, we would like to focus on the maturity models considered within the DesOps framework since they contribute to our research question.

The maturity models are divided into three domains, digitalization, automation, and circularity, and include the latest maturity models for manufacturing systems [VR23]. First, the digitalization domain focuses on the digitalization selection for manufacturing companies [Sc20]. Second, the automation domain specializes on the autonomy level with an artificial intelligence perspective [Ah19]. And third, the circularity domain takes care of the measurement of the circular economy [Ac21]. Based on its impact and relevance for the automotive industry [Ce23], the maturity models for manufacturing systems presented in the DesOps framework propose a practical solution to the need for further research in the operations dimension of the TTCMM [Br24].

In various sectors within the manufacturing sector, the twin transition has been studied mostly based on qualitative research [ABA25], covering from small and medium enterprises [ABA25], regions and products [TPC25] and specialized industries such as: farming [Da24], cosmetics [Pe23]. Based on an initial review, empirical research has not been made in relation to the twin transformation in the automotive industry. This sector could provide promising data due to its economic impact, size and regionality [Ca24; LRV21; Pi21].

In the next chapter, we present the status of the automotive manufacturing industry in relation to twin transformation, its maturity in the digital and sustainable transformation including an example from one automotive company.

4 The Automotive Manufacturing Industry as a Twin Transformer

In order to meet the sustainability targets, like the Paris Agreement, the automotive industry is in a pivotal situation in which they must renovate itself to become more sustainable, achieve the decarbonization targets and fulfill SDGs [Ce23; LKB22; Pi21]. One of the tools to achieve the sustainability demands and remain competitive, as well as the other manufacturing sectors [Ce23], is through the use of digital technologies [ABA25; LRV21].

The opportunity for the automotive industry to become twin transformers has already called the attention of researchers, investigating industrial policy [TG23] and economic concepts [Ca24]. The latter was conducted as a case study, stressing that empirical evidence is necessary to validate the effectiveness of a twin transformation in the automotive sector. In order to continue with the research of twin transformation in the automotive sector, it is necessary to divide it between the product and the manufacturing (production processes), based on an increased product complexity and segmentate value chains, to increase the performance of a potential twin transformation [Ca24]. The manufacturing industry within the automotive sector represents a more attractive field of study since this one represents the basis of its business model [MF10]. Therefore, we propose to focus on the automotive manufacturing industry at first glance.

Aiming to answer our initial research question that intends to bring light to the twin transformation maturity of the automotive manufacturer industry, we focus our research efforts on German automotive manufacturer that employs hundreds of thousands of people with a significant global market share. The first author currently has a working relationship with the referred manufacturer, which could facilitate significant data access and analysis.

The following subsections outline our selected methodology and a preliminary problem analysis which illustrates our initial research steps and paves the road for our problem statement.

4.1 Methodology

Our research is steered by the recently developed eDSR approach outlined by [TWV24] which dissects design science research (DSR) projects in smaller self-contained echelons to facilitate its development, validation and communication in an independent form. Furthermore, it stresses an iterative development and demands rigorous evaluation for the creation of IS artifacts. The eDSR paradigm partitions the DSR in five echelons including (1) problem analysis, (2) objectives and requirements definition, (3) design and development, (4) demonstration and (5) evaluation [TWV24]. (1) Problem analysis is represented by the problem statement [MH19] and validates the state of the art of the problem as well as its solvability [TWV24]. (2) Objectives and requirements definition validates them regarding applicability, coherence and completeness and is illustrated by the design objectives and requirements [Ma19; TWV24]. (3) Design and development aims to create a projectable design for functionality and its architecture evaluated through its applicability, consistency and elegance [TWV24]. (4) Demonstration highlights the instance of the artifact and should be assessed through a proof of concept [Nu15] via its ease of use, robustness and efficacy [TWV24]. (5) Evaluation is the final echelon and provides a contextualized artifact that fulfills the required consistency and completeness of the design [TWV24]. The validation techniques for each echelon will be selected at a later stage of our research taking into consideration the suggestions made by [TWV24]. At this stage, we select the validation techniques for the first echelon, problem analysis, described in the next section.

Additionally, we chose the qualitative method of case study, based on its popularity in the information systems research to study contemporary phenomenon in its real-world environment, in particular where the line between the phenomenon and its environment is ambiguous as it is the case of our research area [Re21].

4.2 Problem analysis (1st Echelon) – Preliminary problem statement

Starting with the first echelon, (1) problem analysis, we present an initial assessment of public and scientific information provided by our selected case study. As one of the biggest representatives of this industry, together with its European counterparts from Italy and France [Ca24], it aims to become greener in each of its manufacturing facilities developing ambitious strategies focused on the measurement of the absolute environmental impact of its manufacturing system [Ge23] and the strategy to achieve zero impact factories [Ge24]. Digitalization is part of the aspects of the “zero impact factory” strategy and it is described as follow: “Digitalization processes enable the transparent and optimized use of resources in factories through environmental data management from the machinery at the shopfloor level to the overall factory level.” [Ge24]. The content of digitalization within the strategy aligns with the ST-related capability for the operations dimension and therefore we can preliminarily conclude that our case organization fulfills at least the (1) maturity stage from the TTCMM [Br24] and potentially be considered as an ST expert.

At this stage, we would like to highlight that the previous conceptual analysis aims to represent preliminary results for the problem analysis since there is not information available about how in detail its strategy is implemented and how the digital and sustainability factors internally interact. We propose the consideration of the maturity models within the DesOps framework for manufacturing systems in the operations dimension of the TTCMM which simultaneously will contribute to the further corroboration and enhancement of the twin transformation. To support our proposal, we will continue our eDSR journey this year with the help of qualitative methods, initiating semi-structured interviews with sustainability and digitalization experts within our case study based on the literature from the selected methodology [TWV24]. Additionally, we plan to collect additional data during the proposed case study aiming to assess factors such as strategy effectiveness and process dimensions between digitalization and sustainability. Our results will provide privileged knowledge about how the digitalization and sustainability efforts are orchestrated within their strategy and will guide our forthcoming echelons [Re21; TWV24].

The presented example aims to illustrate the need for further empirical research in the automotive manufacturer industry, emphasizing to explain how (adaptations of structures, processes, and culture) large organizations embark on their twin transformation journey, rather than exclusively describing what of their transformation endeavor. Multiple large automotive organizations have their own sustainability strategies [LKB22] and experience their own digital transformation [Ce23; LRV21], therefore they could also be investigated on within their ongoing twin transformation. As a result of its economic impact, digital-sustainable development and regionality, the study of the automotive manufacturing industry could provide valuable insights aiming to answer the current challenges the twin transformation presents.

5 Conclusion and future research

We investigated the potential maturity of the automotive manufacturing industry to become a twin transformer based on the presented eDSR and methodologies. Further, we plan to approach the in-depth investigation related to an empirical case study of the maturity stemming from automotive manufacturing systems by the hand of the problem statement in the first place.

Within our planned research plans, we strive to enhance the empirical research of the twin transformations, specifically focusing on the automotive manufacturing sector that later could be extended to broader industry levels and organizations (e.g., manufacturers and suppliers). Further, our research-in-progress hopes to benefit researchers and practitioners with practical and theoretical insights from this endeavor.

Bibliography

[ABA25] Abilakimova, A.; Bauters, M.; Afolayan Ogunyemi, A.: Systematic literature review of digital and green transformation of manufacturing SMEs in Europe. Production & Manufacturing Research, 13(1), p. 2443166, 2025.

[Ac21] Acerbi, F.; Järnefelt, V.; Martins, J. T.; Saari, L.; Valkokari, K.; Taisch, M.: Developing a qualitative maturity scale for circularity in manufacturing. In (Dolgui, A.; Bernard, A.; Lemoine, D.; Von Cieminski, G.; Romero, D., eds): Advances in Production Management Systems. Artificial Intelligence for Sustainable and Resilient Production Systems. Springer International Publishing, Cham, pp. 377–385, 2021.

[Ah19] Ahlborn, K.; Bachmann, G.; Biegel, F.; Bienert, J.; Falk, S.; Fay, A.; Gamer, T.; Garrels, K.; Grotepass, J.; Heindl, A.; Heizmann, J.; Hilger, C.; Hoffmann, M.; Hoffmeister, M.; Jochem, M.; Kalhoff, J.; Kamp, M.; Kramer, S.; Kosch, B.; Legat, C.; Michels, J. S.; Mildner, A.; Nettsträter, A.; Pant, R.; Pittschellis, R.; Schauf, T.; Schlinkert, H.-J.; Ulrich, M.; Zinke, G.: Technology Scenario 'Artificial Intelligence in Industrie 4.0'. Working Paper. Federal Ministry for Economic Affairs and Energy (BMWi), 2019.

[Be20] Berger, S.; Bitzer, M.; Häckel, B.; Voit, C.: Approaching digital transformation – development of a multi-dimensional maturity model. In Proceedings of the 28th European Conference on Information Systems (ECIS), An Online AIS Conference, June 15-17, 2020.

[Br24] Breiter, K.; Crome, C.; Oberländer, A. M.; Schnaak, F.: Dynamic capabilities for the twin transformation climb: a capability maturity model. Information Systems Frontiers, 26(6), pp. 2205–2226, 2024.

[Bu25] Burghard, F.; Heim, L.; Kreuzer, T.; Wozar, J.: Twin to win: A resource orchestration perspective on twin transformation. In: ECIS 2025 Proceedings, 10, 2025.

[Ca18] Calvino, F.; Criscuolo, C.; Marcolin, L.; Squicciarini, M.: A taxonomy of digital intensive sectors. OECD Science, Technology and Industry Working Papers 2018/14, 2018.

[Ca24] Carbonell, J. S.: In search of the twin transition: The limited performativity of the 'green and digital' transitions in the European automotive industry. Working Paper 2024/05. JRC Working Papers Series on Labour, Education and Technology, 2024.

[Ce23] Cetrulo, A.; Dosi, G.; Moro, A.; Nelli, L.; Virgillito, M. E.: Automation, digitalization and decarbonization in the European automotive industry: A roadmap towards a just transition. Working Paper 2023/36. LEM Working Paper Series, 2023.

[Ch24] Christmann, A.-S.; Crome, C.; Graf-Drasch, V.; Oberländer, A. M.; Schmidt, L.: The twin transformation butterfly: capabilities for an integrated digital and sustainability transformation. Business & Information Systems Engineering, 66(4), pp. 489–505, 2024.

[Da24] Dal Mas, F.; Ndou, V.; Massaro, M.; Bagnoli, C.: Combining deep and digital technologies as a path toward twin transition: the "future farming" case study. IEEE Transactions on Engineering Management, 71, pp. 14271–14281, 2024.

[DM16] Dyllick, T.; Muff, K.: Clarifying the meaning of sustainable business: introducing a typology from business-as-usual to true business sustainability. Organization & Environment, 29(2), pp. 156–174, 2016.

[Ge24] Gebler, M.; Warsen, J.; Meininghaus, R.; Baudis, M.; Cerdas, F.; Herrmann, C.: Implementing zero impact factories in Volkswagen's global automotive manufacturing system: a discussion of opportunities and challenges from integrating current science into strategic management. Sustainability, 16(7), p. 3011, 2024.

[Ge23] Gebler, M.; Witte, S.; Blume, S. A.; Muhl, M.; Finkbeiner, M.: The "Impact Points"-method: A distance-to-target weighted approach to measure the absolute environmental impact of Volkswagen's global manufacturing system. Journal of Cleaner Production 386, p. 135646, 2023.

[Gr23] Graf-Drasch, V.; Kauffeld, L.; Kempf, L.; Oberländer, A. M.; Teuchert, A.: Driving twin transformation—The interplay of digital transformation and sustainability transformation. ECIS 2023 Proceedings. European conference on information systems. Kristiansand, Norway, 2023.

[Ho24] Hofmann Trevisan, A.; Acerbi, F.; Dukovska-Popovska, I.; Terzi, S.; Sassanelli, C.: Skills for the twin transition in manufacturing: A systematic literature review. Journal of Cleaner Production, 474, p. 143603, 2024.

[Kü24] Kürpick, C.; Schreiner, N.; Krauß-Kodytek, L.; Kühn, A.; Plass, S.; Scholz, T.: Capabilities for the strategic alignment of sustainability and digitalization in manufacturing: insights from theory and practice. In: 2024 IEEE 65th International Scientific Conference on Information Technology and Management Science of Riga Technical University (ITMS), Riga, Latvia: IEEE, pp. 1–6, 2024.

[LRV21] Llopis-Albert, C.; Rubio, F.; Valero, F.: Impact of digital transformation on the automotive industry. Technological Forecasting and Social Change, 162, p. 120343, 2021.

[LKB22] Lukin, E.; Krajnović, A.; Bosna, J.: Sustainability strategies and achieving SDGs: a comparative analysis of leading companies in the automotive industry. Sustainability, 14(7), p. 4000, 2022.

[MF10] MacDuffie, J. P.; Fujimoto, T.: Why dinosaurs will keep ruling the auto industry. Harvard Business Review, 88(6), 2010.

[Ma19] Maedche, A.; Gregor, S.; Morana, S.; Feine, J.: Conceptualization of the problem space in design science research. In (Tulu, B.; Djamasbi, S.; Leroy, G., eds): Extending the boundaries of design science theory and practice (DESRIST 2019). Springer, pp. 18–31, 2019.

[MH19] Mullarkey, M. T.; Hevner, A. R.: An elaborated action design research process model. European Journal of Information Systems, 28(1), pp. 6–20, 2019.

[Nu15] Nunamaker Jr., J. F.; Briggs, R. O.; Derrick, D. C.; Schwabe, G.: The last research mile: Achieving both rigor and relevance in information systems research. Journal of Management Information Systems, 32(3), pp. 10–47, 2015.

[Pe23] Perossa, D.; Acerbi, F.; Rocca, R.; Fumagalli, L.; Taisch, M.: Twin Transition cosmetic roadmapping tool for supporting cosmetics manufacturing. Cleaner Environmental Systems, 11, p. 100145, 2023.

[Pi21] Pichler, M.; Krenmayr, N.; Schneider, E.; Brand, U.: EU industrial policy: Between modernization and transformation of the automotive industry. Environmental Innovation and Societal Transitions, 38, pp. 140–152, 2021.

[Re21] Recker, J.: Scientific research in information systems: a beginner's guide. Cham: Springer International Publishing (Progress in IS), 2021.

[Re23] Rehman, S. U.; Giordino, D.; Zhang, Q.; Alam, G. M.: Twin transitions & industry 4.0: Unpacking the relationship between digital and green factors to determine green competitive advantage. Technology in Society, 73, p. 102227, 2023.

[Sc20] Schuh, G.; Hicking, J.; Jordan, F.; Stroh, M.-F.; Saß, S.-A.: Strategic target system to select digitalization measures in manufacturing companies. In (Camarinha-Matos, L. M.; Afsarmanesh, H.; Ortiz, A., eds): Boosting Collaborative Networks 4.0. Cham: Springer International Publishing, pp. 227–236, 2020.

[STT24] Spaltini, M.; Terzi, S.; Taisch, M.: Development and implementation of a roadmapping methodology to foster twin transition at manufacturing plant level. Computers in Industry, 154, p. 104025, 2024.

[TPC25] Tabares, S.; Parida, V.; Chirumalla, K.: Twin transition in industrial organizations: Conceptualization, implementation framework, and research agenda. Technological Forecasting and Social Change, 213, p. 123995, 2025.

[TG23] Topuria, S.; Gräf, H.: German industrial policy and the twin transition: Pre- and post-Covid trajectories in the automotive and IT services sectors. Working Paper 213/2023, 2023.

[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.

[Uh22] Uhrenholt, J. N.; Kristensen, J. H.; Rincón, M. C.; Adamsen, S.; Jensen, S. F.; Waehrens, B. V.: Maturity model as a driver for circular economy transformation. Sustainability, 14(12), p. 7483, 2022.

[Un25] United Nations: Sustainable Development Goals. Department of Economic and Social Affairs, 2025. Available at: https://sdgs.un.org/goals (accessed: 10/03/2025).

[VR23] Van Erp, T.; Rytter, N. G. M.: Design and operations framework for the Twin Transition of manufacturing systems. Advances in Production Engineering & Management, 18(1), pp. 92–103, 2023.

[Vi19] Vial, G.: Understanding digital transformation: A review and a research agenda. The Journal of Strategic Information Systems, 28(2), pp. 118–144, 2019.

Management of geological data and notifications in accordance with the Geological Data Act in Schleswig-Holstein

Friedhelm Hosenfeld [1], Jan Willer[2], Laura Dzieran [3], and Johannes Tiffert[4]

Abstract: To cope with the tasks resulting from the Geological Data Act (GeolDG), the Schleswig-Holstein State Agency for the Environment (LfU) designed a Geological Notification and Management System (GAMS), which is intended to provide suitable support for the administration and processing of geological notifications and investigation data. The system presented, which is currently under development, is based on the data analysis and business intelligence platform Disy Cadenza Workbooks, into which a management application for controlling workflow processes, data checking and categorisation is integrated. Particular challenges lie in the implementation of numerous interfaces to components and external systems that are important for the editing process, such as the front-end web application for notifications of geological investigations. Dashboards have been designed in Cadenza Workbooks, which provide a convenient overview of essential data and enable the processing of notifications to be controlled.

Keywords: Geological Data Act, geological investigations, data management and data sharing, Cadenza Workbooks, dashboards

1 Introduction

The Geological Data Act (GeolDG, [Gz20], [RSH23]), which came into force in 2020, requires the State Agency for the Environment (LfU), as the responsible state authority for Schleswig-Holstein, to receive, review and publish notifications of geological investigations within specified deadlines. In addition, the geological data submitted following the investigations must be categorised according to their properties and their need for protection for public availability.

In terms of making data available, the Geological Data Act goes significantly beyond other existing legal regulations in environmental information, open data and access 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] Landesamt für Umwelt Schleswig-Holstein, Hamburger Chaussee 25, D-24220 Flintbek , jan.willer@lfu.landsh.de
[3] Landesamt für Umwelt Schleswig-Holstein, Hamburger Chaussee 25, D-24220 Flintbek , laura.dzieran@lfu.landsh.de, https://orcid.org/0000-0002-7370-9980
[4] DigSyLand - Institute for Digital System Analysis & Landscape Diagnosis, Zum Dorfteich 6, D-24975 Husby, tiffert@digsyland.de

doi:10.18420/env2025_13

information laws, as not only governmental data is affected, but the publication of data collected by the private sector is also prescribed: clients and contractors of geological investigations like drillings must report these before they are carried out and make the investigation data available afterwards [Gz20].

The aim is to permanently secure geological data and make it publicly available in order to ensure the sustainable management of the geological subsurface [Gz20].

In order to be able to guarantee legally compliant processing of the notifications and data deliveries within the tight time frame specified, continuous digital data processing and documentation is required from the notification of a geological investigation to the transfer of the associated technical and assessment data and their categorisation, combined with automated publication of the categorisation decision and the publication of the verification data.

This resulted in the need for the LfU to develop a management system that would provide suitable support for those responsible in the fulfilment of these tasks.

This article presents the Geological Notification and Management System (GAMS: Geologisches Anzeige- und Managementsystem), the development of which was initiated on the basis of the aforementioned task. Chapter 2 describes the requiremens and the architecture of the system followed by Chapter 3 dealing with the development steps and implementation. Chapter 4 concludes with a summary and an outlook.

2 Requirements and Architecture

2.1 Requirements for the geological notification management system

According to the Geological Data Act, notifications must be submitted at least 14 days before the start of a geological investigation. The relevant information on the planned investigation is transmitted with a notification, so that the management of the notifications essentially comprises the import of the notifications with structured, analysable storage of the information (*verification data*), support in checking the notifications and monitoring the subsequent transmission deadlines.

Once the geological investigations have been completed, both *technical data* and *assessment data* must be submitted to the LfU within certain deadlines. This leads to requirements for the management of these investigation results for the administration of this data and for support in checking the data. This includes in particular the *categorisation of the data*, which is to be carried out by the LfU and supported by GAMS in a suitable manner: All transmitted data must be assigned to categories in accordance with the Geological Data Act. The categorisation must be documented in an administrative act and made available to the public as well as to the notifying parties in a timely manner.

The overarching requirements therefore include monitoring the relevant deadlines, but also the provision of a message portal to enable communication with the involved parties and an interface to the electronic records system for audit-proof filing of the categorisation.

The areas of responsibility must be adequately mapped via a user and rights management system in order to be able to guarantee all data protection and data security requirements.

2.2 Overview of the basic components

The system was to be located within the central operating infrastructure of the Ministry of the Environment called ZeBIS (**Ze**ntrale **B**etriebs**i**nfra**s**truktur, see also [Ho21b]) in the TDC (Twin Data Centre) data centre of the state IT service provider Dataport[5] and also build on the standard components provided there as far as possible. These include central PostgreSQL database servers with PostGIS, on which the GAMS database is set up to manage all data, including geodata, and the ETL tool FME (Feature Manipulation Engine[6]) for transferring data and files between the platforms involved (see Figure 1). The use of certain specialised applications also requires the operation of an Oracle RDBMS within the LfU, so that data from the GAMS database must be transferred to Oracle as well as files must be transferred to the LfU using FME ETL processes.

Other ZeBIS components include a data warehouse database (DWH) for public data, which serves as a data source for OGC geodata services that are realised with the GDI software deegree[7].

The core management system consists of a PHP management application that is integrated into the Disy Cadenza Workbooks platform. Both components are explained in detail in the following sections (see sections 3.1 and 3.2).

Within the TDC data centre, further basic services such as secure file transfer, file service, mail server and Keycloak[8] are provided. Keycloak serves as a central identity and access management tool with a connection to the Active Directory (AD) of the state authorities offering single sign-on.

Outside the TDC, the Schleswig-Holstein Environmental Portal[9] is operated, which on the one hand forms the meta-information system for all environmental data, but on the other hand also makes all geodata accessible to the public via a WebGIS on the basis of the OGC services offered by deegree, including the drilling data [LfU25].

[5] Dataport AöR: https://www.dataport.de/who-we-are/, TDC: https://www.dataport.de/services-produkte/rechenzentrum/ accessed: 30/07/2025
[6] Safe Software: https://fme.safe.com/, accessed: 30/07/2025
[7] deegree: https://www.deegree.org/, accessed: 30/07/2025
[8] Keycloak: https://www.keycloak.org/, accessed: 30/07/2025
[9] Schleswig-Holstein Umweltportal: https://umweltportal.schleswig-holstein.de/, accessed: 30/07/2025

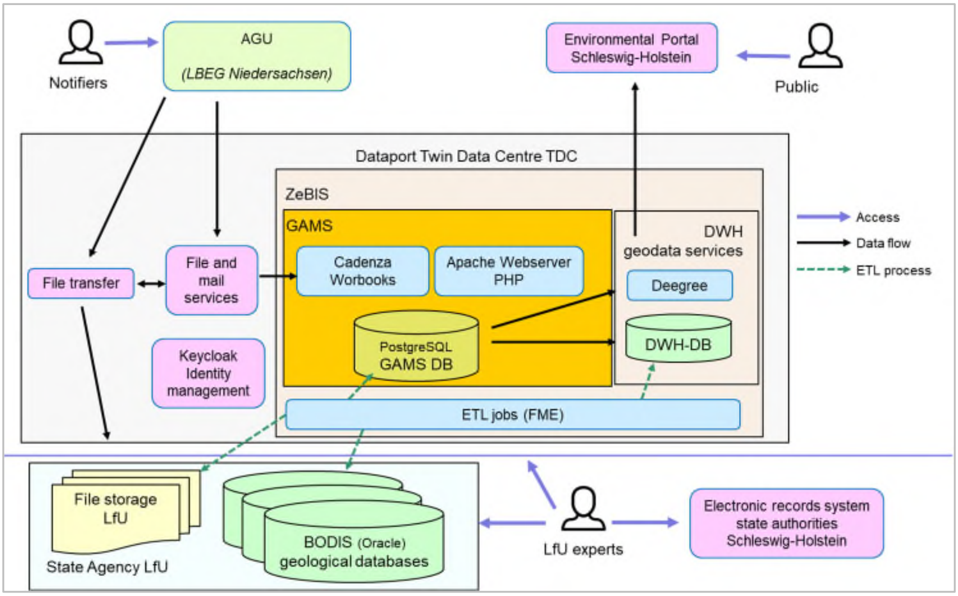


Figure 1:Architectural sketch of the most important components used for GAMS

3 Development steps and implementation

3.1 Navigation, analyses and WebGIS with Disy Cadenza Workbooks

The Disy Cadenza platform [DI25] is already being used for various tasks and issues in the Schleswig-Holstein environmental department. This includes specialised information systems based on Cadenza as well as map and factual data-based information and evaluation systems in the ZeBIS environment.

Disy Cadenza is a software which has been developed by the company Disy for many years on behalf of a co-operation of numerous federal states and other partners and adapted to current requirements and technical possibilities. It combines flexible evaluations, reporting and GIS functionalities in a software platform that is optimised for environmental data in particular [DI25].

In contrast to the previous - still widely used - Cadenza Classic, the current *Cadenza Workbooks* platform relies consistently on web technology with database-based configuration via a web interface. Supplementary web applications can be integrated directly into the Cadenza user interface in order to extend the evaluation platform in particular for editing data.

3.2 Integrated PHP management application to control the processes

While Cadenza Workbooks can be seen as a comprehensive construction kit for the flexible compilation of analyses and GIS functions, GAMS also requires a component that can implement more complex input functionalities as well as process controls developed individually for the management of notifications and geological data. This management application is implemented using the PHP[10] scripting language, which also allows flexible adaptations to technical requirements to be realised at short notice.

The combination of Cadenza with an integrated PHP specialised application to supplement application-specific functions has already proven to be a well suited solution in the Schleswig-Holstein environmental department in areas such as geology and soil, e.g. with the development of a consulting database geology and soil [Ho17] and a geotope register [Ho21a] as well as with other web applications for other departments of the LfU.

The PHP management application for GAMS is implemented on the basis of a framework developed by DigSyLand, which provides basic functionalities with regard to data security, catalogue processing, input validation and configurable data entry masks.

GAMS is particularly innovative because the management application is one of the first specialised applications to be integrated directly into Cadenza Workbooks via the Web API Workbooks, as the application framework and user management of Cadenza Classic were previously always used. Therefore, significant new developments had to be created with the connection via the Web API Workbooks and with the OpenID Connect interface to Keycloak as the central identity provider for GAMS. On the basis of this connection, which realises a single sign-on (SSO) of Cadenza Workbooks and the management application, also direct linkage between dashboards and input masks is possible.

3.3 Interface to the notification web application AGU

The submission of the notifications and data of the geological investigations is performed by the notifiers using the web application AGU[11] (*Anzeige Geologischer Untersuchungen*: notification of geological investigations) operated by the State Agency for Mining, Energy and Geology (LBEG[12]) of Lower Saxony (s. Figure 1). The LBEG provides this front-end application for the states of Lower Saxony, Bremen, Hamburg, and Schleswig-Holstein as well as for the Federal Institute for Geosciences and Natural Resources (BGR). Submitted notifications including all data files are sent by the AGU to the responsible state authorities.

---

[10] PHP: https://www.php.net/, accessed: 30/07/2025

[11] Anzeige Geologischer Untersuchungen (AGU): https://nibis.lbeg.de/agu/startseite, accessed: 30/07/2025

[12] Landesamt für Bergbau, Energie und Geologie (LBEG) des Landes Niedersachsen: https://www.lbeg.niedersachsen.de/, accessed: 30/07/2025

It is planned that the GAMS application will receive the notifications by the AGU by e-mail or file transfer as XML files. GAMS carries out checks for formal consistency and stores the notification data in the PostgreSQL GAMS database (see Figure 1). This also includes the transfer of the transmitted geometries of the investigation area, so that a spatial analysis by the specialists is possible with GAMS (e.g. boreholes in so-called risk areas). Once the data has been imported, subsequent processes such as e-mail notifications to the responsible employees at the LfU and the notifying parties are automatically triggered via the message portal implemented in the GAMS management PHP application. The basic routines of these processes have already been developed, mainly used for a predecessor application of AGU, and are now being adapted to the current AGU and GAMS applications.

A secure file transfer from the LBEG to the GAMS data centre is planned for the transfer of additional files with the investigation data, so that the transferred files can be analysed using the PHP management application, which supports the examination by the experts. Finally, the files with the examination data are transferred to the internal LfU network for storage using an ETL process (see Figure 1).

3.4 Evaluation dashboards

Comprehensive, customisable evaluations of the data are required in order to optimally carry out the technical assessment of the reports and investigation data. This includes map visualisations as well as the selection, analysis and further processing of notification data sets selected on the map. The evaluation criteria and the compilation of the result presentations as tables, maps and diagrams as well as reports should be customisable by the specialist managers themselves as required.

Therefore dashboards are implemented in Cadenza Workbooks offering the essential information in a suitable display form at a glance, so that, for example, depending on the processing deadlines, the data records to be processed as a matter of priority are highlighted with the option of calling them up directly in the GAMS management application for processing. Lists of the queried records are added by map views and diagrams.

3.5 Further current development steps

The input and editing masks of the GAMS management application are currently being completed, in particular with functionalities for realising the notication management workflows. This includes automatic checking, e. g. for risk areas based on geometries, unrealistic depths, etc. as well as support for the manual processing steps for formal checking, confirmation and publication, including the option of requesting additional information from the notifier. It will provide suitable support for the legally prescribed classification of the investigation data and the determination of the data category by offering processing masks and checklists.

In addition to maps of the location of geological investigations the information provided on the Internet also includes the categorisation notices, which are generated from the management application as accessible (barrier-free) PDF files. The categorisation notice documents the submitted and checked verification data as well as the list of the submitted data with categorisation as *verification data*, *technical data* or *assessment data* including further classifications like reasons for protection. As the categorisation notices are made publicly available as part of the official administrative act, it was an essential requirement to create the corresponding PDF files in an accessible format. It is also planned to store these files belonging to the corresponding official administrative act in the state's e-file (electronic records system) in an audit-proof manner (see Figure 1).

4 Summary and outlook

The GAMS Geological Notification and Management System is designed to efficiently support the State Agency for the Environment in Schleswig-Holstein as the responsible authority in the management of notifications and geological investigation data.

The automatic import into the GAMS database with structuring of the supplied metadata and basic consistency checks facilitates the technical review for those responsible, who are also supported by customised evaluation options in Cadenza Workbooks.

Centrally curated dashboards with tables, diagrams and maps as well as views and evaluations customised by the users themselves enable flexible responses to current requirements. The GAMS database thus provides a standardised database that is required to fulfil the obligations of the Geological Data Act by the state agency.

The IT components in the central operational infrastructure ZeBIS provided for by the framework concept of the Ministry of the Environment can be suitably combined in GAMS, so that GAMS can be considered a good example of the application of the framework concept.

Main challenges include the linkage with internal and external systems and components providing a number of interfaces, such as for the ID provider Keycloak and the AGU web frontend.

The first installation on the test platform in the ZeBIS environment in the data center, with connection to the basic components, is planned for mid-2025. Going live is scheduled for the end of 2025.

Bibliography

[DI25] Disy Informationssysteme GmbH: Software for Data Analytics, Business & Location Intelligence, https://www.disy.net/en/products/disy-cadenza/overview/, accessed: 30/07/2025.

[Gz20] Gesetz zur staatlichen geologischen Landesaufnahme sowie zur Übermittlung, Sicherung und öffentlichen Bereitstellung geologischer Daten und zur Zurverfügungstellung geologischer Daten zur Erfüllung öffentlicher Aufgaben (Geologiedatengesetz - GeolDG) - Geologiedatengesetz vom 19. Juni 2020 (BGBl. I S. 1387), 2020, https://www.gesetze-im-internet.de/geoldg/, accessed: 30/07/2025

[Ho17] Hosenfeld, F. et al.: Consulting Database Geology and Soil in Schleswig-Holstein. In (Otjacques, B. et al., eds.): From Science to Society: The Bridge provided by Environmental Informatics, Adjunct Proceedings of the 31st EnviroInfo conference, Luxembourg, pp. 173-179, 2017.

[Ho21a] Hosenfeld, F. et al.: Neuentwicklung einer Intranet-Web-Anwendung für das Geotopkataster Schleswig-Holstein. In (Freitag, U. et al., eds.): Umweltinformationssysteme - Wie verändert die Digitalisierung unsere Gesellschaft? Tagungsband des 27. Workshops „Umweltinformationssysteme (UIS2020)“ des Arbeitskreises „Umweltinformationssysteme“ der Fachgruppe „Informatik im Umweltschutz‘‘ der Gesellschaft für Informatik (GI), Springer Vieweg, Wiesbaden, pp. 13-27, 2021, https://doi.org/10.1007/978-3-658-30889-6_2, accessed: 30/07/2025.

[Ho21b] Hosenfeld, F. et al.: Governmental Information Systems Supporting Nature Conservation in Schleswig-Holstein. In (Wohlgemuth, V. et al., 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.

[LfU25] Landesamt für Umwelt Schleswig-Holstein: Bohrungen in Schleswig-Holstein. Metadatensatz im Umweltportal Schleswig-Holstein, https://umweltportal.schleswig-holstein.de/trefferanzeige?docuuid=ffa99fdc-8548-41fd-b10a-a165ff905aee, accessed: 30/07/2025.

[RSH23] Reiche, S.; Schröder, L. A.; Hermann, D.: The collection, digitisation, interpretation and publication of geological data in the German site selection procedure – status and challenges. In (Dietl, C. et al., eds.): Interdisciplinary Research Symposium on the Safety of Nuclear Disposal Practices safeND 2023, Berlin, 13–15 September 2023, Safety of Nuclear Waste Disposal, Volume 2, pp. 141–141, 2023, https://doi.org/10.5194/sand-2-141-2023, accessed: 30/07/2025.

Green Delivery Analytics: Advancements in Open Data for Sustainable Last Mile Logistics

Richard Schulte [1], Mattes Leibenath [2], Franziska Giese [3], Uta Kühne [4], Senad Hasanspahic [5], Somrita Banerjee[6], Aina Andriamananony[7] and Benjamin Wagner vom Berg [8]

Abstract: In last mile logistics the implementation of electric vehicles and associated infrastructure are key elements to increase sustainability and efficiency. The Green Delivery Analytics project focuses on the systematic use of open data sources and compliance with open standards in data formats, structures and communication protocols to emphasize interoperability and adaptability. These standards facilitate seamless data access, exchange and publication, particularly to strengthen reusability, for example as contributions to national open data portals, such as the Mobilithek. This paper showcases how the project incorporates open data and associated interoperability architecture aspects to achieve more sustainability within the analytics and planning of last mile logistics in German cities. The aim is to advance the knowledge base on open data and interoperability in the field of last mile logistics and contribute to the UN SDG Targets 9.4, 11.6 and 13.2.

Keywords: Open Data, Last Mile Logistics, Data Architecture, Spatial Analytics, Smart City, Sustainable Logistics

1 Introduction

The ‘last mile’ in distribution logistics refers to the final segment of the supply chain, which is characterized by a high degree of complexity, partly due to the delivery of goods to a diverse range of recipients [SLWW13]. This distribution stage accounts for a significant portion of both the overall distribution costs and the associated emissions [Brab20]. The optimization of last mile logistics (LML) for electric vehicles, such as

[1] Hochschule Bremerhaven, Smart Mobility Institute, An der Karlstadt 8, Bremerhaven, 27568, rschulte@bremerhaven.de, https://orcid.org/0009-0008-6650-1832

[2] Hochschule Bremerhaven, mleibenath@hs-bremerhaven.de, https://orcid.org/0009-0000-7209-0592

[3] Hochschule Bremerhaven, fgiese@hs-bremerhaven.de, https://orcid.org/0009-0004-9052-9684

[4] Hochschule Bremerhaven, ukuehne@hs-bremerhaven.de, https://orcid.org/0009-0007-2199-6096

[5] Hochschule Bremerhaven, shasanspahic@hs-bremerhaven.de, https://orcid.org/0009-0009-6281-0416

[6] Hochschule Bremerhaven, sbanerjee@hs-bremerhaven.de, https://orcid.org/0000-0001-7474-7981

[7] Cosmo Mobility UG Customer-Oriented Sustainable Mobility, Butteldorf 10, Elsfleth, 2693, info@cosmo-mobility.org

[8] Hochschule Bremerhaven, bwagnervomberg@hs-bremerhaven.de, https://orcid.org/0000-0001-6976-9625

doi:10.18420/env2025_14

electric cargo bikes, is considered a central solution, as they contribute to both emission reduction and compliance with legal requirements [GBKP19]. The utilization of cargo bikes requires micro-hub structures for commissioning and consolidation of goods within cities [Trip19], as well as optimization of networks in regard to location, transportation, and route planning [HuHG20]. The economic and ecological implementation of cargo bikes and associated infrastructure requires a strategic combination of operational parameters [SBGM19] as well as the linking and analysis of patterns within large amounts of data from the contexts of city, sustainability and logistics [CBVV17].

The research project "Green Delivery Analytics" (GDA) takes on these challenges. Within the project a software tool will be developed, that enables municipalities and logistics service providers to plan and optimize the LML for the 80 biggest cities, based on a data-driven approach. [Bund24]

Comprehensive data focus forms the basis for data-driven decisions in logistics [GuMR21] and therefore the development of the data architecture is considered essential for the effective integration and processing of diverse data sources [Shei13] which can be used to realize the planning of sustainable LML. The high relevance and the large amounts of data in such software solutions require a particularly large focus on data architecture aspects such as data model, structure, format and management [Shei13].

While (open) data on logistics is rarely available, context data on e.g. demographics and city infrastructure is often more likely to be openly available based on data collection by municipalities [CVMT24]. This offers opportunities to address the challenges of data unavailability [AlSF25]. The usage and redistribution of open data and derived data products strives for the aforementioned focus on architectural aspects to fulfill the requirement of interoperability of open data and in this manner contribute to the implementation and monitoring of SDGs [Stat19].

RQ. How can an open data architecture be designed and implemented to integrate heterogeneous municipal, traffic and socio-economic datasets so that low-emission LML networks can be planned and evaluated for German cities?

This paper contributes to research on German LML through the development of an open data catalogue, a data architecture prototype and by enabling research on the implementation of UN SDG Targets 9.4, 11.6 and 13.2.

2 Background and Related Work

Due to the continuous growth in e-commerce activity and home deliveries, LML is playing an increasingly important role in freight logistics, especially in urban areas. Within the courier, express and parcel market, LML is the most expensive and least efficient segment of the logistics chain and accounts for a large proportion of emissions. [SLWW13]

The decisive influencing factors here are the logistics infrastructure, the consignment structure and the structure of the recipients. The EU is increasing pressure through climate policies and urban regulations to promote alternatives in LML [Euro00]. The layout of the delivery districts, the stop density and the proportion of standard, courier and express consignments also have an impact on efficiency and costs. Sustainable LML systems rely on the use of light commercial electric vehicles, cargo bikes, micro-hubs and urban consolidation centers, which require new approaches to infrastructure planning. [Keco25]

Network design for micro-vehicles must account for constraints such as limited range, lower payload, traffic regulations and topography. Traditional heuristics or planning models based on diesel vans are inadequate for LML [FeEs25]. Emerging tools support routing, fleet composition and hub placement using GIS and optimization models but often rely on proprietary software or siloed data [Elen23]. As [NeCA20] emphasizes, “the growing volume and variety of data produced in the urban ecosystem are crucial for obtaining the city's insights and building knowledge-based solutions for a smarter and more sustainable urban development”. The authors also state that open data can be considered an important enabler for such data-driven solutions [NeCA20]. The necessary data for LML solutions includes data on socio-demographics, geography as well as logistics/traffic data [GuMR21]. Some of these can be retrieved as open data.

Open data can be defined as data that can be freely used, reused and redistributed by anyone, typically under conditions that at most require attribution and adherence to share-alike licenses. The key principles of open data emphasize availability, accessibility, interoperability and clarity in licensing [NeCA20, Okf00a]. The Open Knowledge Foundation’s (OKF) definition on open data, adopted by related literature such as [NeCA20], underlines key attributes including universal participation and the freedom to access and redistribute data. The OKF’s full “Open Definition” extends this definition onto other domains, such as open knowledge or open formats [Okf00a, Okf00b]. In the context of geospatial standardization, the role of the Open Geospatial Consortium (OGC), as the authoritative body for open geospatial standards, is especially significant [Ogc00].

To ensure legal compliance with these principles, several licenses are recommended by the OKF, including Creative Commons Zero (CC0-1.0), Creative Commons Attribution (CC-BY-4.0), Open Data Commons Attribution (ODC-By-1.0) and the Open Database License (ODbL-1.0), among others. These licenses are evaluated for compatibility with the OKF’s principles in the Open Definition. [Okf00c]

The importance of open data is further supported by the FAIR Guiding Principles for scientific data management and stewardship, which aim to ensure that data should be FAIR (“Findable, Accessible, Interoperable and Reusable”). [WDAA16]

3 Catalogue of Open Data for LML

As shown in Chapter 2, LML optimization requires data from different domains, including socio-demographic, geographical, logistical/traffic, as well as geospatial reference data for statistical analytics [Bund00a]. Unfortunately, logistics and traffic data are usually not published as open data, due to economic and data protection reasons [TmEA19].

Table 1: Catalogue of Open Data for LML in Germany

Domain	Example Dataset	License	Source
Socio-demographic data	Official Census, Global Human Settlement Layer	dl-by-de/2.0, CC-By-4.0	[PSPF24, Stat00]
Geographical data	Official Federal Data (ATKIS, ALKIS), OpenStreetMap	dl-zero-de/2.0 or dl-by-de/2.0, CC-By-4.0, ODbL-1.0	[Advd00, Lgln00, MoMi17]
Logistical data	NA	NA	NA
Geospatial Reference data	GeoGitter INSPIRE (ETRS89-extended LAEA)	dl-by-de/2.0	[Bund00a]

As seen in Table 1, open data can be derived from various sources, ranging from official federal cadastral offices or official European projects to volunteered geographic information in OpenStreetMap. Problems arise through the complexity in used licenses (even within datasets, due to state-level organization), which may require name attribution or even share-alike.

4 Prototype Architecture

Due to the project's objective to publish its own data to the German open data portal for mobility data – Mobilithek [Bund00b] – it is needed to design the projects data architecture in a way that ensures reusability of produced data [WDAA16]. The FAIR Guiding Principles showcase why data architecture plays a big role in the FAIRness of publishable data. Subgoals of the FAIR Guiding Principles include ensuring that the used "protocol is open, free and universally implementable", data is described in "formal (…) and broadly applicable language" or that data meets "domain-relevant community standards" [WDAA16]. In general, data architecture can be understood as the blueprint for how all data is collected, stored, integrated and made accessible for analysis [Shei13]. To ensure the project's data architecture is designed accordingly, used file formats, data structures and communication protocols need to be coherent and comprehensible. A well-designed architecture can not only establish the FAIR Guiding Principles but also enables a seamless integration of the different datasets that vary greatly in structure and granularity.

Figure 1 gives a brief overview of the data architecture used within the project GDA to integrate open data and contribute to interoperability of the project's approach. It shows the architecture divided into different data processing layers. Within the data storage layer an object-relational database including a spatial extension is used as a central and harmonized data store of the logistical and contextual data.

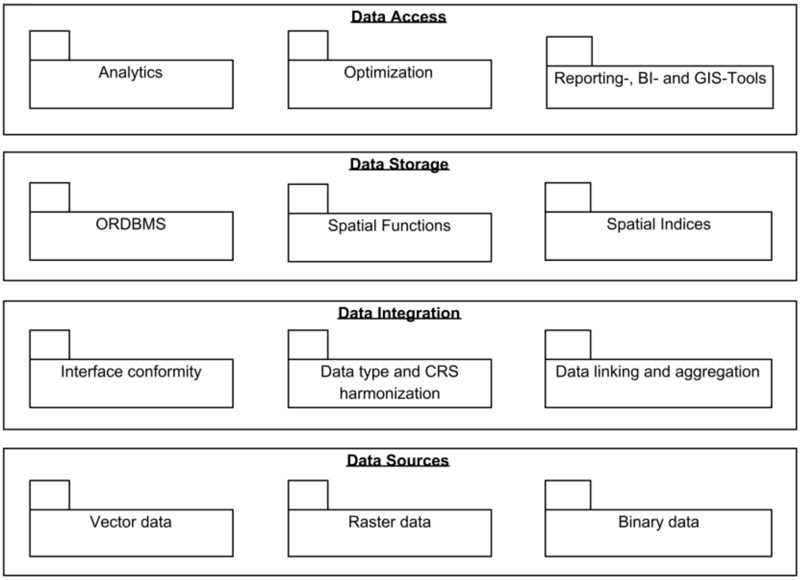


Fig. 1: Overview of the Prototype Architecture of the GDA Project

Two data models exist for the storage of spatial information, vector and raster. Vector data consists of points, which might form polygons or lines, and is useful for the rendering of data, as the shape of an object can be accurately displayed. Raster or grid data is stored in geolocated grid cells or pixels. The examples can range from high-resolution satellite imagery to population densities in square kilometers. Due to the rasterization of data some form of aggregation usually takes place, while the vector data, which explicitly stores geometries for individual objects, usually needs to simplify the number of points used to shape a polygon or line. [CoKa92]

Both data models utilize different data formats, with formats describing "a specific, pre-established structure for the organization of a digital file or bitstream" [Prem15]. The formatting of information has been a focus of standardization processes, as big data analytics and data exchange requires consistent and unique conventions [GCLM17].

In the current draft of the Open Knowledge Foundation's Open Format Definition an open format is defined as "a freely available published specification which places no restrictions, monetary or otherwise, upon its use" or at the very least as a format that "can be processed with at least one free/libre/open-source software tool" [Okf00d]. The GDA architecture uses open file formats wherever possible to avoid vendor lock-in effects and enhance interoperability.

The Project uses the following open formats [Gisg19]:

1. Raster, such as GeoTIFF.
2. Vector, such as Shapefile, GeoPackage, GML, GeoJSON.
3. Binary, such as OSM PBF.

In this project, as shown in the previous sections, raster data and vector data have to be handled simultaneously to enable a complex analysis of geospatial and socio-demographic data. Therefore, the project must be able to harmonize and unify the heterogeneous formats in terms of resolution, attribution and reference systems for example through the rasterization of vector data or the vectorization of raster data. [SPLL20]

The harmonization of coordinate reference systems (CRS) plays a crucial role in the integration of geospatial datasets, as these datasets are often based on heterogeneous reference systems. A CRS provides a standardized framework to describe locations on Earth's surface by using coordinate definitions. This may include geographic (e.g., latitude/longitude), projected (planar), or geocentric (3D) coordinate systems. [KSSP23]

Since many datasets are published using varying CRS, inconsistencies may arise in spatial analyses if no harmonization is conducted. Particularly for applications involving distance, area or direction calculations, projected CRS with minimal regional distortion are preferred. In this context, the use of EPSG:3035 (ETRS89 / LAEA Europe) was chosen due to its suitability for pan-European statistical and area analyses, as well as its compatibility with INSPIRE-recommended standards, in contrast to the official geodetic CRS EPSG:4258, which is not projected and less suitable for metric operations. [ALGI01]

To enable necessary spatial transformations a flexible and robust data management system is essential. The project utilizes PostgreSQL, an open source object-relational database system, to store and process analytical datasets, which include fact tables (e.g. stop density or cost metrics), dimension tables (e.g. road types, surface characteristics) and shared spatial and temporal dimensions (e.g. city grid cells, time intervals). PostgreSQL uses a "BSD-style license and is thus free and open source software" [AgRa16] and has been one of the first databases to allow for spatial analysis. [AgRa16]

PostGIS, one of the many available extensions for PostgreSQL, offers specialization for spatial analysis and supports spatial joins (e.g. intersect, within, distance-based joins), spatial indexing (e.g. GiST indexes) and operations such as buffering, clipping and coordinate transformations. [AgRa16]

Besides efficient data storage and processing, the interoperability and accessibility of spatial datasets also depend significantly on the utilization of standardized and open protocols. In general, open protocols are a specific type of open standard, just like open formats, and should ensure the interoperability between systems, services and stakeholders

[CeFu07]. Protocols are central to the FAIR Guiding Principle of Accessibility, which states that (meta)data should be retrievable through standardized communication protocols that are open, free, and universally implementable, while also supporting authentication and authorization mechanisms. [WDAA16]

In the project, protocols are necessary for accessing the data sources, communication between internal services or components, as well as publishing results to the Mobilithek. Even though many open protocols for spatial formats have been developed, many services still rely on HTTP Downloads (like ATKIS, Census, Global Human Settlement Layer) or OGC Web Feature Service or Web Map Service (like ALKIS). For futureproof interoperability, tools should be able to use OGC API (currently establishing), REST APIs and GraphQL.

5 Conclusion

The study has shown that a robust open data architecture for sustainable last-mile logistics has to address a variety of standards to be able to systematically integrate spatial structure layers, socio-demographic data and analytical grids into a PostGIS-centered stack that complies with FAIR principles. The GDA project delivers a systematic overview of open data sources in the context of LML analytics. Furthermore, open formats, open protocols as well as open relational and spatial data structures that are associated with these data sources are described to briefly present the requirements of processing this data. The presented content is particularly useful for researchers, urban planners and logistics stakeholders who aim to develop or enhance data-driven solutions for sustainable LML. It also serves as a valuable foundation for software developers, data engineers and system architects who intend to process or analyze such data.

At the same time, the paper has highlighted several challenges, e.g. data sparsity on logistics-specific data and the heterogeneity of existing open standards. Nevertheless, the modular architecture and adherence to open protocols aim for flexibility and expandability. Next steps include the definition of specific data integration components, the ingesting of carrier-provided logistics data, the utilization of GraphQL/REST endpoints for logistics/traffic data (not open data) as well as deploying AI-enabled analytics that transform curated datasets into predictive demand and enable the optimization of location, transportation and route planning.

Acknowledgements

GDA is funded by the Bundesministerium für Verkehr (BMV, German Federal Ministry of Transport) as part of the mFUND program (project number 19F2262A) with a funding amount of around 2.1 Mio. Euro. As part of mFUND the BMV supports research development projects in the field of data based and digital mobility innovations. Part of

the project funding is the promotion of networking between the stakeholders in politics, business, administration and research as well as the publication of open data on the Mobilithek portal.

Bibliography

[Advd00] AdV der Länder: *OPEN DATA*. URL https://basemap.de/open-data/. - abgerufen am 2025-05-27. — basemap.de

[AgRa16] Agarwal, Sarthak; Rajan, K. S.: Performance analysis of MongoDB versus PostGIS/PostGreSQL databases for line intersection and point containment spatial queries. In: *Spatial Information Research* Bd. 24 (2016), Nr. 6, S. 671–677

[ALGI01] Annoni, A; Luzet, C; Gubler, E; Ihde, J ; EUROPEAN COMMISSION, JOINT RESEARCH CENTRE (Hrsg.): *Map Projections for Europe*. Italy : Institute for Environment and Sustainability, 2001

[AlSF25] Al-Rahamneh, Anas; Serrano-Hernandez, Adrian; Faulin, Javier: The Impact of Integrating Open Data in Smart Last-Mile Logistics: The Example of Pamplona Open Data Catalog. In: *Sustainability* Bd. 17 (2025), Nr. 2, S. 408

[Brab20] Brabänder, Christian: *Die Letzte Meile: Definition, Prozess, Kostenrechnung und Gestaltungsfelder*, *essentials*. Wiesbaden : Springer Fachmedien Wiesbaden, 2020

[Bund00a] Bundesamt für Kartographie und Geodäsie: *Geographische Gitter für Deutschland in Lambert-Projektion (GeoGitter Inspire)*. URL https://gdz.bkg.bund.de/index.php/default/geographische-gitter-fur-deutschland-in-lambert-projektion-geogitter-inspire.html. - abgerufen am 2025-05-27

[Bund00b] Bundesministerium für Digitales und Verkehr: *Über - Mobilithek.info - Mobilitätsdaten Deutschland*. URL https://mobilithek.info/about. - abgerufen am 2025-06-10. — Über

[Bund24] Bundesministerium für Verkehr: *Datenbasierte Ausgestaltung einer CO2-neutralen Last-Mile-Logistik für die 80 größten deutschen Städte - Green Delivery Analytics*. URL https://www.bmv.de/SharedDocs/DE/Artikel/DG/mfund-projekte/green-delivery-analytics.html. - abgerufen am 2025-05-19. — BMV.de

[CBVV17] Cardenas, Ivan; Borbon-Galvez, Yari; Verlinden, Thomas; Van De Voorde, Eddy; Vanelslander, Thierry; Dewulf, In: *Competition and Regulation in Network Industries* Bd. 18 (2017), Nr. 1–2, S. 22–43

[CeFu07] Cerri, Davide; Fuggetta, Alfonso: Open standards, open formats, and open source. In: *Journal of Systems and Software* Bd. 80 (2007), Nr. 11, S. 1930–1937

[CoKa92] Congalton, Russell; Kass Green: The ABCs of GIS. In: *Journal of Forestry* Bd. 90 (1992), Nr. 11

[CVMT24] Correia, Diogo; Vagos, Cristiano; Marques, João Lourenço; Teixeira, Leonor: Fulfilment of last-mile urban logistics for sustainable and inclusive smart cities: a case study conducted in Portugal. In: *International Journal of Logistics Research and Applications* Bd. 27 (2024), Nr. 6, S. 931–958

[Elen23] Elena García: *D1.2 - Data Management Plan (DMP) v1* (Data Management Plan Nr. 1), 2023

[Euro00] European Commission: *Urban mobility and accessibility*. URL https://commission.europa.eu/eu-regional-and-urban-development/topics/cities-and-urban-development/priority-themes-eu-cities/urban-mobility-and-accessibility_en. - abgerufen am 2025-06-18. — Urban mobility and accessibility

[FeEs25] Ferreira, Joao C.; Esperança, Marco: Enhancing Sustainable Last-Mile Delivery: The Impact of Electric Vehicles and AI Optimization on Urban Logistics. In: *World Electric Vehicle Journal* Bd. 16, Multidisciplinary Digital Publishing Institute (2025), Nr. 5, S. 242

[GBKP19] Gina Chung; Ben Gesing; Keya Chaturvedi; Philipp Bodenbenner Logistics Trend Radar, DHL Customer Solutions & Innovation (Hrsg.).

[GCLM17] Ghiringhelli, Luca M.; Carbogno, Christian; Levchenko, Sergey; Mohamed, Fawzi; Huhs, Georg; Lüders, Martin; Oliveira, Micael; Scheffler, Matthias: Towards efficient data exchange and sharing for big-data driven materials science: metadata and data formats. In: *npj Computational Materials* Bd. 3 (2017), Nr. 1, S. 46

[Gisg19] GISGeography: *The Ultimate List of GIS Formats and Geospatial File Extensions*. URL https://gisgeography.com/gis-formats/. - abgerufen am 2025-07-26. — GIS Geography

[GuMR21] Gutierrez-Franco, Edgar; Mejia-Argueta, Christopher; Rabelo, Luis: Data-Driven Methodology to Support Long-Lasting Logistics and Decision Making for Urban Last-Mile Operations. In: *Sustainability* Bd. 13 (2021), Nr. 11, S. 6230

[HuHG20] Huang, Zhihong; Huang, Weilai; Guo, Fang: Integrated sustainable planning of micro-hub network with mixed routing strategy. In: *Computers & Industrial Engineering* Bd. 149 (2020)

[Keco25] KE-Consult Kurt&Esser GbR: *Nachhaltigkeitsstudie 2025*. Köln : Bundesverband Pkaet- und Expresslogistik e. V., 2025

[KSSP23] Kumar, Manish B.; Singh, Anju; Pravesh, Ram; Majid, Syed Irtiza; Tiwari, Akash: Referencing and Coordinate Systems in GIS. In: Kumar, M. ; Singh, R. B. ; Singh, A. ; Pravesh, R. ; Majid, S. I. ; Tiwari, A. (Hrsg.): *Geographic Information Systems in Urban Planning and Management*. Singapore : Springer Nature, 2023, S. 25–46

[Lgln00] LGLN: *ALKIS-Komponenten*. URL https://www.lgln.niedersachsen.de/startseite/geodaten_karten/afis_alkis_atkis/alkis/alkis-komponenten-101056.html. - abgerufen am 2025-06-19

[MoMi17] Mooney, Peter; Minghini, Marco: A review of OpenStreetMap data. In: *Mapping and the citizen sensor*, Ubiquity Press (2017), S. 37–59

[NeCA20] Neves, Fátima Trindade; de Castro Neto, Miguel; Aparicio, Manuela: The impacts of open data initiatives on smart cities: A framework for evaluation and monitoring. In: *Cities* Bd. 106 (2020)

[Ogc00] OGC: *OGC Standards*. URL https://www.ogc.org/standards/. - abgerufen am 2025-06-17

[Okf00a] OKF: *What is Open Data?* URL https://opendatahandbook.org/guide/en/what-is-open-data/. - abgerufen am 2025-05-21. — Open Data Handbook

[Okf00b] OKF: *What is open?* URL https://okfn.org/en/library/what-is-open/#:~:text='Open%20knowledge'%20is%20any%20content,building%20blocks%20of%20open%20knowledge. - abgerufen am 2025-05-21. — Open Knowledge Library

[Okf00c] OKF: *Conformant Licenses*. URL https://opendefinition.org/licenses/. - abgerufen am 2025-05-21

[Okf00d] OKF: *Open Format Definition*. URL https://opendefinition.org/ofd/. - abgerufen am 2025-06-11

[Prem15] PREMIS Editorial Committee: *PREMIS Data Dictionary for Preservation Metadata, Version 3.0*, 2015

[PSPF24] Pesaresi, Martino Schiavina ,Marcello; Politis ,Panagiotis; Freire ,Sergio; Krasnodębska ,Katarzyna; Uhl ,Johannes H.; Carioli ,Alessandra; Corbane ,Christina; u. a.: Advances on the GHSL by joint assessment of Earth Observation and population survey data. In: *International Journal of Digital Earth* Bd. 17, Taylor & Francis (2024), Nr. 1

[SBGM19] Sheth, Manali;Butrina, Polina; Goodchild, Anne; McCormack, Edward: Measuring delivery route cost trade-offs between electric-assist cargo bicycles and delivery trucks in dense urban areas. In: *European Transport Research Review* Bd. 11 (2019), Nr. 1, S. 11

[Shei13] Sheikh, Nauman: *Implementing analytics: A blueprint for design, development, and adoption* : Newnes, 2013

[SLWW13] Schnedlitz, Peter; Lienbacher, Eva; Waldegg-Lindl, Barbara; Waldegg-Lindl, Marianne: Last Mile: Die letzten – und teuersten – Meter zum Kunden im B2C ECommerce. In: Crockford, G. ; Ritschel, F. ; Schmieder, U.-M. (Hrsg.): *Handel in Theorie und Praxis*. Wiesbaden : Springer Fachmedien Wiesbaden, 2013, S. 249–273

[SPLL20] Silva-Coira, Fernando Ladra, Susana; López, Juan R.; Gutiérrez, Gilberto: Efficient processing of raster and vector data. In: Rao, P. (Hrsg.) *PLOS ONE* Bd. 15 (2020), Nr. 1

[Stat00] Statistische Ämter des Bundes und der Länder: *Der Zensus-Atlas*. URL https://www.zensus2022.de/DE/Aktuelles/Hinweis_Zensusatlas.html. - abgerufen am 2025-05-27. — Statistisches Bundesamt

[Stat19] Statistics Division United Nations: *Introduction to data interoperability across the data value chain*. Dhaka, 2019

[TmEA19] T. Moschovou; E.I. Vlahogianni; A. Rentziou: Challenges for data sharing in freight transport.

[Trip19] Tripp, Christoph: *Distributions- und Handelslogistik: Netzwerke und Strategien der Omnichannel-Distribution im Handel* : Springer Fachmedien Wiesbaden, 2019

[WDAA16] Wilkinson, Mark D.; Dumontier, Michel; Aalbersberg, Ijsbrand Jan; Appleton, Gabrielle; Axton, Myles; Baak, Arie; Blomberg, Niklas; Boiten, Jan-willem; u. a.: The FAIR Guiding Principles for scientific data management and stewardship. In: *Scientific Data* Bd. 3. London, United States, Nature Publishing Group (2016)

An innovative educational format for the heating transition: The Heat Transition School as a practical teaching and learning concept

Frederic Urbschat [1], Sven Lampe [2], Jantje Halberstadt [1], Jorge Marx Gomez [2]

Abstract: The heat transition describes the shift from fossil fuels to sustainable, predominantly renewable heat supply systems – a central element of the energy transition. In the *WärmewendeNordwest* (WWNW) project, the Heat Transition School, an innovative teaching and learning format, was developed for this purpose. This format integrates project topics and conveys key challenges while promoting the development of viable solutions to real-life problems. The target groups – prospective master craftspeople in building systems integration and students from various disciplines – work on practical tasks while combining their different perspectives. An accompanying evaluation provided important impetus for further development. This article highlights the design process, the actors involved, the project phases and challenges, and recommendations for the use of these kinds of formats with heterogeneous target groups in research projects.

Keywords: innovative teaching and learning formats, heat transition, education format, transfer

1 Introduction

Combating climate change requires profound change, particularly in the energy sector. As part of the energy transition, the heating transition aims to decarbonise heat generation, distribution, and use. Buildings are a key area of action, as they account for around 32% of global final energy demand and 30% of energy-related CO_2 emissions [Cl22; Ür20]. Achieving the goals of the Paris Agreement will require not only technological innovations, but structural and behavioural changes as well [Cl18; IE23; UN15]. Successful transition integrates energy efficiency, renewable energies, and conscientious energy use [IE21; La19]. Implementing this transformation is complex and requires coordinated political, planning, and social integration efforts [Lu14; Re18]. Education also plays an essential role. Innovative formats such as the Heat Transition School, developed in the *WärmewendeNordwest* project [WW21], address this challenge by bringing together theory and practice while promoting interdisciplinary cooperation and providing targeted training for future skilled workers [Ba14a; Ba14b; WWR11]. The Heat Transition School

[1] University of Vechta, Faculty I, Department of Sustainability, Entrepreneurship & Innovation, Driverstraße 22, Vechta, 49377, Germany, frederic.urbschat@uni-vechta.de, https://orcid.org/0009-0002-7646-5886; jantje.halberstadt@uni-vechta.de, https://orcid.org/0000-0002-6647-532X
[2] Carl von Ossietzky University, Faculty II, Department of Computing Science, Ammerländer Heerstraße 114-118, Oldenburg, 26129, Germany, sven.lampe1@uni-oldenburg.de, https://orcid.org/0000-0001-9550-7100; jorge.marx.gomez@uol.de

doi:10.18420/env2025_15

is aimed at prospective master craftspeople in the field of building system integration, as well as masters students from various disciplines such as transformation management in rural areas (geography) and energy science and engineering. This article presents the design process and challenges of this educational approach, discusses its pedagogical framework, and reflects on the insights gained from it with a focus on future applications.

2 Background and Related Work

2.1 Background

The heating transition (German: *Wärmewende*) describes the decarbonisation of heat supply systems, and is a central pillar of the energy transition in Germany and many other countries. Consistent with the Paris Agreement targets [Cl18; UN15], it builds on three core principles: (1) Efficiency - optimising energy use to achieve the same or greater benefits with less energy input [La19], (2) Substitution - replacing fossil fuels with renewable sources such as solar thermal, biomass, or ambient heat, and (3) Sufficiency - reducing demand through behavioural change [Gr18; St06].

The building sector accounts for 32% of global final energy demand and around 30% of energy-related CO_2 emissions [IE21; Ür15], with heating and cooling representing the largest share. Key measures include energy-efficient refurbishment and renovation, integration of renewable heating technologies (e.g. heat pumps, solar thermal), and such as smart heating controls, building automation, and energy monitoring systems. Digitalisation enables continuous performance tracking, data-driven optimisation, and the integration of decentralised energy generation [BAG+12; KKK+12].

Practical implementation requires coordinated action across multiple spatial levels:

- Building: improved insulation, low-temperature heating systems, and digital control technologies to reduce demand and manage renewable integration.
- Campus: universities can serve as living labs to test heat recovery concepts, e.g. from data centres [Jo18] and intelligent HVAC control [LBH+23].
- District: decentralised networks combining generation, storage, and smart distribution can reduce losses and costs, especially in low-density areas [Lu14; PW11].
- Municipal: local heat transition strategies benefit from participatory planning, integration of district heating, and digital management platforms to coordinate stakeholders [BB05; Re18].

Technological evolution such as the heating transition require the alignment of technical, social and institutional change [Ge02]. This makes effective knowledge transfer essential;

this needs to be supported by inclusive training formats and collaborative innovation networks [LR10]. The research project *WärmewendeNordwest* addresses these challenges by integrating digitalisation, technical innovation, and knowledge transfer across all levels of heat supply. Within this framework, the Heat Transition School was developed to link academic and vocational expertise in a practical, interdisciplinary setting.

2.2 Related Work

A targeted literature search (June 2025) in ERIC, Scopus, FIS Bildung, and Google Scholar focused on educational settings combining university students and vocational trainees. With the exception of four publications from the healthcare sector [Fl24; He19; Mi22; Re16], no relevant studies in technical or general education fields were identified.

Examples from the healthcare sector confirm the benefits of joint learning: Fleig et al. [Fl24] evaluated a one-day interprofessional training (SICKO Junior) for medical and nursing trainees, while Mink et al. [Mi22] analysed a four-week collaboration at the Heidelberg Interprofessional Training Centre. Even though comparable concepts exist in veterinary education [He19] and extracurricular tutorials [Re16], all of them are limited in disciplinary scope.

To our knowledge, there are no documented cases of interdisciplinary, practice-oriented formats combining engineering students and prospective master craftspeople in the context of the heating transition. The Heat Transition School therefore represents a novel approach in technical education, explicitly integrating digital methods, applied engineering knowledge, and cross-sector collaboration.

3 Purpose of the Heat Transition School

The aim of the Heat Transition School is to develop innovative, practical teaching and learning formats. The focus is on a flexible concept that appeals to both prospective master craftsmen and students, and is closely linked to real project topics from the WWNW project.

1. Project transfer: The School promotes knowledge transfer between project partners and participants by working on real project topics and developing practical solutions. Under the guidance of the partners, participants learn about the key challenges of the heat transition and, ideally, develop transferable approaches. This strengthens their analytical, creative and problem-solving skills – with long-term benefits for their education and career. At the same time, networking between science and skilled workers is expanded upon.

2. Participation: The School enables participants to actively engage with project topics. The integration of different perspectives generates new impetus for the project partners, linking and promoting a dynamic exchange between transfer and participation.

3. Cooperation for the heat transition: The interdisciplinary composition of the participants promotes the exchange between theory and practice. Project partners accompany the process and ensure professional quality. Working on real-life problems leads to innovative solutions. At the same time, mutual understanding between the target groups grows, contributing to the sustainable establishment of the format.

4. Transition to the format: The aspects mentioned above form the basis of the Heat Transition School concept and are reflected in the structure and design of its format. Chapter 4 discusses these guidelines and describes its concrete development.

4 Format Development Workshops

As part of transversal activity 2, the Heat Transition School was developed as a flexible teaching format for a wide range of target groups. The aim was to develop a needs-based concept with innovative didactic approaches for students and prospective master craftspeople. The originally planned 1.0 format was not implemented; instead, following additional developments and work, the adapted 2.0 format was created.

Format 1.0: The original concept was developed in several workshops with project partners and students. Content and topic suggestions were closely aligned with the six research fields and the transversal activity. Initial drafts envisaged interactive formats such as excursions and simulation games. However, feedback revealed challenges in terms of addressing and reaching the target groups, especially trainees. Finding a date, duration, and location proved to be laden with organisational hurdles. The low number of binding commitments ultimately led to a revision of the format.

Format 2.0: Further development of the previous format: In March 2025, a further development workshop was held to design a viable format based on previous experience. The target groups were adjusted, content was streamlined, and organisational processes were simplified. The aim was to create a feasible, connectable format with clearly recognisable added value.

Key adjustment:

- Target groups: Prospective master craftspeople were involved instead of trainees – as part of their full-time training with a project partner, craftspeople were more easily accessible and did not need to take time off work. And the university students participating in the project remained “in place”, i.e. their participation took place via university teaching.
- Key topics: The four core areas – generation, use, building technology, future – remained unchanged. The allocation was based on existing project content to reduce the amount of preparation required.

- Duration and location: The duration was reduced to two days to facilitate integration into existing training and study contexts. The venue was in Oldenburg, Germany and the surrounding area.
- Participation and approach: Participation was mandatory for master trainees/students. An adapted approach to encourage early application is currently being developed for students.

The Heat Transition School was held as a three-day event in Hamburg, Germany, with the location chosen in consultation with the master students involved in planning the excursion. It was a compulsory part of the master training programme. The first day was spent travelling to the venue, getting to know each other, and introducing the topics. On the following days, key aspects of the heat transition were addressed through practical tasks. The highlight was the excursion to HafenCity in Hamburg, allowing insights into different urban heat transition approaches.

The content of the participating research fields (RF) and transversal activity (TA) was the thematic breadth of the heat transition. These ranged from the application of digital platforms in everyday craftsmanship (RF1); to issues of building automation, data security, and interface management (RF2); to building technology system analyses and the development of decision-making rules using artificial intelligence (RF3). Other topics included intelligent control strategies for CO_2 reduction in local energy systems (RF4), the scaling of technical potential at the neighbourhood level (RF5), and municipal action planning for the socially acceptable implementation of the heat transition (RF6). In addition, aspects regarding the use of open data and data-based platforms were addressed (TA1). The work was carried out in interdisciplinary groups via the use of practical questions. The training institution of the prospective master craftspeople did not play an active role in generating this content.

Day 1	Day 2	Day 3
1. Arrival of target groups/organisation team (starting midday)	1. Arrival of project partners	1. Arrival & welcome
2. First meeting	2. Arrival & welcome/agenda	2. Reflection on Day 2: results from the heat transition café
3. Intro to the heat transition: board game “Changing the Game”	3. Input lecture: heat transition	3. Group work with LEGO®-based creative method
4. Dinner together	4. Presentation: *WärmewendeNordwest* project	4. HafenCity tour in Hamburg: energy/heat transition in neighbourhoods
5. Additional goals	5. Heat transition café: research fields & transversal activities (40 minutes per table)	5. Conclusion (late afternoon)
	6. Evening social event	

Tab. 1: Three-day workshop schedule for stakeholder engagement in the heat transition

5 Format Description

What’s new? The Heat Transition School brought together two target groups: prospective master craftspeople with practical experience, and students with theoretical backgrounds. In a shared learning environment, they worked together on key issues relating to the heat transition in an interdisciplinary manner. The content was closely aligned with the research fields and transversal activity of the WWNW project.

What’s different? Students could only earn credit points within the framework of regular modules; otherwise, participation was voluntary with no credits being earned. Participation was free of charge for both groups, as the format was financed by project funds – an indication of the research-oriented and experimental nature of this project.

What added value does the Heat Transition School offer? The School teaches the basics of the heat transition, while also focussing on practical formats such as a simulation game with real-life problems. The exchange between theory and practice promotes mutual understanding and creative solutions. Participants benefit from personal and professional development, interdisciplinary learning, and sustainable networking.

Evaluation as part of the heat transition school. A key element was the accompanying evaluation by students from the University of Vechta, Germany, who actively participated to systematically document the process and carry out an overall evaluation. Expectations, daily feedback, and assessments of individual programme items were collected, including the simulation game, specialist input, the Lego Serious Play session, and the excursion to HafenCity. The aim was to qualitatively assess and optimise the format.

6 Challenges

After an initial planning phase with a different format and target group, the concept was redesigned at the beginning of 2025. In May 2025, a three-day event (format 2.0) was successfully held with a focus on prospective master craftspeople. Various challenges arose during the preparation and implementation:

1. Involvement of project partners: The participation of the project partners proved difficult in some cases, as not all of them had structurally integrated innovative teaching and learning formats. The short notice given in May 2025 also made coordinated participation more difficult.

2. Target group acquisition: Despite early notice, reaching students proved challenging. This may have been due to the chosen time period in the middle of their semesters, and the lack of credit points available.

3. Accommodation: The tight time frame made it difficult to find acceptable accommodation in terms of facilities, location, and capacity.

7 Evaluation Results

A full evaluation was not yet available as of June 2025. However, initial feedback provides indications of the effectiveness of the format, as well as starting points for further development. The evaluation was carried out in a standardised manner using a five-point Likert scale and open feedback.

Day 1: The board game Changing the Game was generally considered thematically relevant and entertaining, even while it was perceived as too short and limited in content.

Day 2: Technical input, project presentations, and the interactive heat transition café were rated positively overall – especially the exchange between the research fields. Differences in the evaluation of individual contributions indicated different interests and presentation formats. The daily structure was perceived as coherent, with the café standing out as particularly effective.

Day 3: Reflection with Lego Serious Play was mostly met with criticism; the method was widely perceived as unsuitable. The subsequent excursion to Hamburg's HafenCity was rated more positively – especially in terms of relevance and practical clarity.

General feedback: The general conditions (accommodation, catering, time structure) were generally rated positively, while the programme and organisation were assessed more critically. There were isolated indications of an increase in competence, even though these could not be verified across the board.

Classification of results: Due to the small number of participants and the unbalanced group composition (ratio of master students to students approximately 3:1), the results can only be generalised to a limited extent.

8 Recommendations

Based on previous experience, the following key recommendations can be derived for successful implementation: Teaching and learning formats should be anchored in the structures of the project partners in an effort to effectively use resources. It is essential to select the venue early on, taking into account conference rooms, technology, and catering. Relevant stakeholders, especially companies, should be involved at an early stage and provided with specific information about the added value of the School. It's also advisable to address prospective master craftspeople and university students in different ways, for example by emphasising the practical or academic relevance of their participation. Linking the programme to curricular offerings – for example through credit points – can provide additional motivation. During implementation, clear structures, professional moderation, and active expectation management are crucial for learning success in heterogeneous groups. Practical content and concrete project tasks promote the connection between theory and practice. A long lead time supports planning reliability, while waiving participation fees promotes low thresholds and diversity. Finally, structured follow-up is recommended to achieve systematic feedback evaluation. Examining digital or hybrid formats can help reach new target groups and increase scheduling flexibility.

9 Conclusion

In both its original and revised formats, The Heat Transition School aims to provide practical insights and job-related skills. The interdisciplinary collaboration between students and prospective master craftspeople, as well as the project's tackling of real-life challenges, promote systemic thinking and joint approaches to solutions. With format 2.0, the concept was specifically adapted to the needs of the participants: modularised content enabled integration for students, while practice-oriented content supported the transfer to the professional work of prospective master craftspeople. The Heat Transition School is thus establishing itself as an innovative learning space for the skilled workers of tomorrow – most notably as a transferable model for energy transition professional qualification. Initial evaluation results confirm that the format successfully combines theory and practice, promotes interdisciplinary cooperation, and contributes to the targeted qualification of future skilled workers. Its goals have thus been achieved.

Acknowledgements

We would like to thank the Federal Ministry of Research, Technology and Space for its financial support for the *WärmewendeNordwest* project (03SF0624). Our very special thanks go out to all project partners for their active and constructive participation in the design and further development of the Heat Transition School.

References

[Ba14a] Barth, M.: Implementing Sustainability in Higher Education: Learning in an age of transformation. Routledge, 2014.

[Ba14b Barth, M. et al.: Learning to change universities from within: A service-learning perspective on promoting sustainable consumption in higher education. Journal of Cleaner Production 62, pp. 72–81, 2014.

[BAG+12] Batty, M.; Axhausen, K.W.; Giannotti, F., et al.: Smart cities of the future. The European Physical Journal Special Topics 214 (1), pp. 481–518, 2012.

[BB05] Bulkeley, H.; Betsill, M.: Rethinking Sustainable Cities: Multilevel Governance and the ‘Urban’ Politics of Climate Change. Environmental Politics 14 (1), pp. 42–63, 2005.

[Cl18] on Climate Change (IPCC), I. P.: Global Warming of 1.5°C. Accessed: 2025-03-15, Intergovernmental Panel on Climate Change, 2018.

[Cl22] on Climate Change (IPCC), I. P.: Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the IPCC. Cambridge University Press, 2022.

[Fl24] Fleig, K. et al.: Interprofessionelles Team-Training für Pflegeauszubildende und Medizinstudierende in der Pädiatrie: analoge und digitale Formate im Vergleich, Interprofessional Team Training for Nursing Trainees and Medical Students in Pediatrics: Comparing Analog and Digital Formats. German, Klinische Pädiatrie, Epub ahead of print, 2024.

[Ge02] Geels, F.W.: Technological transitions as evolutionary reconfiguration processes: a multi-level perspective and a case-study. Research Policy 31 (8), pp. 1257–1274, 2002.

[Gr18] Grubler, A. et al.: A low energy demand scenario for meeting the 1.5 °C target and sustainable development goals without negative emission technologies. Nature Energy 3, pp. 515–527, 2018.

[He19] Heise, C.: Interprofessionelles Lernen zwischen Studierenden der Veterinärmedizin und Auszubildenden in Tierpflegeberufen, Unpublished Master’s thesis, Masterarbeit, Stiftung Tierärztliche Hochschule Hannover, 2019, https://elib.tiho-hannover.de/servlets/MCRFileNodeServlet/tiho_derivate_00002808/Heises-ws24.pdf.

[IE21] (IEA), I. E. A.: Net Zero by 2050, https://www.iea.org/reports/net- zero-by-2050, Accessed: 2025-03-10, 2021.

[IE23] (IEA), I. E. A.: Energy Efficiency 2023, https://www.iea.org/reports/energy-efficiency-

2023, Accessed: 2025-03-20, 2023.

[Jo18] Jouhara, H. et al.:Waste Heat Recovery Technologies and Applications. Thermal Science and Engineering Progress, 2018.

[KKK+12] Klein, L.; Kwak, J.-Y.; Kavulya, G., et al.: Coordinating occupant behavior for building energy and comfort management using multi-agent systems. Automation in Construction 22, pp. 525–536, 2012.

[La19] Lachal, B.: Energy transition. In: Energy transition. Wiley, Hoboken, NJ, USA, 2019.

[LBH+23] Lesnyak, E.; Belkot, T.; Hurka, J., et al.: Applied Digital Twin Concepts Contributing to Heat Transition in Building, Campus, Neighborhood, and Urban Scale. Big Data and Cognitive Computing 7 (3), p. 145, 2023.

[LR10] Loorbach, D.; Rotmans, J.: The practice of transition management: Examples and lessons from four distinct cases. Futures 42 (3), pp. 237–246, 2010.

[Lu14] Lund, H. et al.: 4th generation district heating (4GDH): Integrating smart thermal grids into future sustainable energy systems. Energy 68, pp. 1–11, 2014.

[Mi22] Mink, J. et al.: Interprofessionelle Sozialisation und Zusammenarbeit auf einer interprofessionellen Ausbildungsstation – eine rekonstruktive Analyse. Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen 169, pp. 94–102, 2022, https://www.sciencedirect.com/science/article/pii/S1865921722000010.

[PW11] Persson, U.; Werner, S.: Heat Distribution and the Future Competitiveness of District Heating. Applied Energy 88, pp. 568–576, 2011.

[Re16] Reichel, K. et al.: Interprofessional peer-assisted learning as a low-threshold course for joint learning: Evaluation results of the interTUT Project. GMS Journal for Medical Education 33 (2), Doc30, 2016, https://www.egms.de/static/en/journals/zma/2016-33/zma001029.shtml.

[Re18] Reckien, D. et al.: How are cities planning to respond to climate change? Assessment of local climate plans from 885 cities in the EU-28. Journal of Cleaner Production 191, pp. 207–219, 2018.

[St06] Stern, N.: The Economics of Climate Change: The Stern Review. Cambridge University Press, 2006.

[UN15] UNFCCC: The Paris Agreement, https://unfccc.int/process-and-meetings/the-parisagreement, Accessed: 2025-03-10, 2015.

[Ür15] Ürge-Vorsatz, D. et al.: Heating and cooling energy trends and drivers in buildings. Renewable and Sustainable Energy Reviews 41, pp. 85–98, 2015.

[Ür20] Ürge-Vorsatz, D. et al.: Advances toward a net-zero global building sector. Annual Review of Environment and Resources 45, pp. 227–269, 2020.

[WW21] (WWNW), W. P. H.: Wärmewende Nordwest, https://www.waermewende-nordwest.de/, Accessed: 2025-03-20, 2021.

[WWR11] Wiek, A.; Withycombe, L.; Redman, C. L.: Key competencies in sustainability: A reference framework for academic program development. Sustainability Science 6 (2), pp. 203–218, 2011.

Frameworks for Sustainable Digital Infrastructures

Empowering Communities through IoT²: A Citizen Science Approach to Flood Monitoring and AI-supported Resource Optimization

Digital Resilience, Maker Education, and Energy-Aware IoT Development

Franziska Mai[1], Guido Burger[2] and Klaus-Uwe Gollmer [3]

Abstract: This paper presents an integrated educational and technical framework for strengthening climate resilience and digital empowerment through citizen-driven environmental sensing and algorithmic literacy. The IoT²-Werkstatt[4] is an open platform that enables students, teachers and communities to build and program IoT-based monitoring devices using open, visual and AI-supported tools. As a central use case, we examine a co-developed flood monitoring project carried out as part of the project "DigiSelfTrans"[5]. The platform fosters technical and civic empowerment in STEM education (Science, Technology, Engineering, Mathematics), particularly in rural areas. This paper outlines the conceptual foundations and implementation strategy of the IoT²-Werkstatt, presents selected outcomes (e.g. Energy savings for field devices), and discusses practical barriers such as system maintenance and equitable participation. It contributes to emerging hybrid models of sustainability education and participatory technology.

Keywords: Internet of Things and Thinking, Makerspace, Artificial Intelligence, Citizen Science, Climate Resilience, STEM Education, Energy Efficiency, Digital Inclusion

1 Introduction

Climate change and digital transformation are reshaping how education and public infrastructure must evolve. While institutions draft top-down policies, many rural communities lack the tools and capacity to act locally, resulting in growing gaps in participation. Rural regions often suffer from limited access to hands-on STEM (Science, Technology, Engineering, Mathematics) education [RM25], while also facing disproportionate exposure to climate-related threats such as flash floods or biodiversity loss. What is needed is an approach that enables especially young people to monitor,

[1] Environmental Campus Birkenfeld, f.mai@umwelt-campus.de
[2] Maker, mail@fab-lab.eu
[3] Environmental Campus Birkenfeld, k.gollmer@umwelt-campus.de, https://orcid.org/0000-0002-1330-499X
[4] www.iot-werkstatt.de
[5] https://www.umwelt-campus.de/iss/projekte/laufende-projekte/digiselftrans

doi:10.18420/env2025_16

understand, and shape their environment through accessible, meaningful digital tools.

The IoT²-Werkstatt is an educational innovation that combines open hardware, visual programming, AI-assisted learning, and community-based sensing. Students, educators, and volunteers build real-world systems such as air quality indicators, irrigation tools, or flood monitors using plug-and-play sensors and microcontrollers. [GB23] A central example is a flood monitoring system from the DigiSelfTrans project, developed with schools, municipalities and the German Red Cross.

Rather than proposing a fixed solution, the IoT²-Werkstatt provides a learning ecosystem where people explore and iterate. This paper outlines the platform's structure, selected project results, and key open challenges such as device maintenance, sustained use, and data ethics—situating the work in broader debates around digital resilience and community empowerment. [Go24][Go19a][Go19b]

2 Informatics Education – From Algorithms to Applications

Digital education requires more than access to devices. It must foster algorithmic thinking and digital creation. The IoT²-Werkstatt introduces learners to coding through visual tools such as Ardublock.

Graphical programming allows learners to create sensor-based applications by connecting logic blocks - lowering entry barriers and removing syntax errors. Students work with variables, conditions, and loops, and later move on to finite state machines or event-based programming. A basic example might involve reading a CO_2 value and activating a traffic light system (green/yellow/red) depending on thresholds[6].

Key informatics concepts include:

- Input–Process–Output logic,
- Data abstraction,
- Conditional execution and loops,
- Modularization and system decomposition,
- IoT²-Superblocks.

[6] https://www.heise.de/select/make/2020/5/2022015381334973804

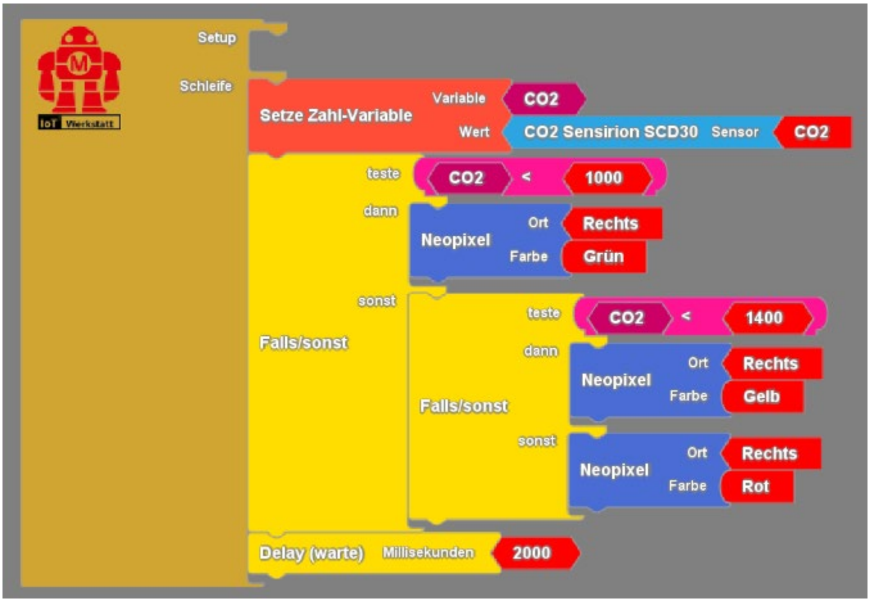


Figure 1: Ardublock screenshot – simplified CO_2 traffic light logic

People are not only coding; they are building systems that respond to real environmental data. They measure, decide, and act - closing the loop between sensor data and tangible action. This promotes a systems-thinking mindset and strengthens problem-solving skills.

For those who wish to go further, the platform enables a smooth transition to text-based programming in Arduino C/C++. The visual blocks correspond directly to standard code structures, supporting progression from beginner to intermediate level.

In summary, the IoT²-Werkstatt offers an intuitive yet powerful entry point[7] into informatics education, rooted in real-world relevance and designed for accessibility, creativity, and long-term learning.

3 Citizen Science in Practice: Flood Monitoring

As flash floods become more frequent and less predictable, decentralized and low-cost early-warning systems gain importance – particularly in regions where conventional infrastructure is either unavailable or underdeveloped. As part of the "DigiSelfTrans" project, the IoT²-Werkstatt has been set up to enable citizens to work together on developing and deploying flood sensors, among other things, under the motto

[7] https://www.umwelt-campus.de/iot-werkstatt/tutorials/schnellstart-octopus-anschliessen-und-einrichten

"Understanding digitalization, promoting self-efficacy, mastering the climate crisis". These robust, battery-powered devices are designed, assembled, and programmed in schools, communes and blue light associations, forming a grassroots infrastructure for local resilience.

Technical tasks include:

- Soldering components,
- 3D printing and laser cutting of building components,
- Programming energy-efficient DeepSleep cycles in Ardublock,
- Integration of Wi-Fi or LoRaWAN communication,
- Synchronizing with cloud platforms like ThingSpeak, Thingsboard or local MQTT brokers.

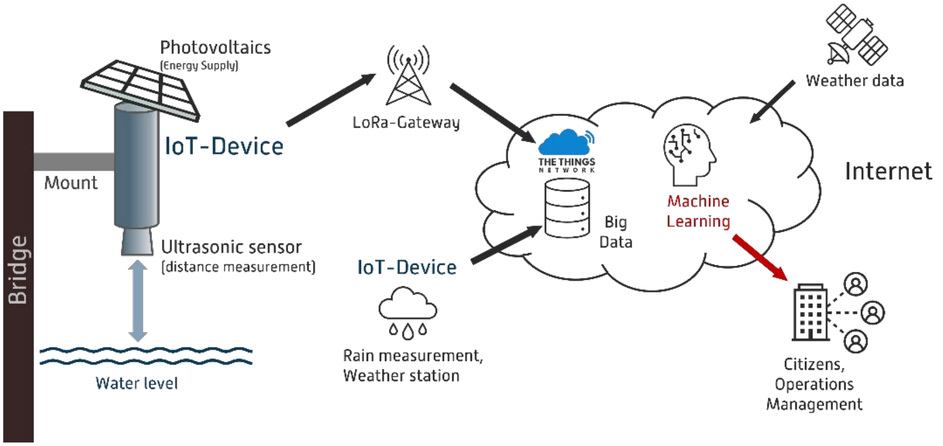


Figure 2: Functioning and networking of the flood gauge

The core device is a microcontroller with batteries/solar and an ultrasonic distance sensor, which we install in a PVC-tube on a bridge or above a small river. It measures the distance from the sensor to the water surface. Since we mount the sensor at a fixed height, a smaller distance means a higher water level - and in many locations, this directly correlates with potential flooding of nearby homes or infrastructure.

Once the measured water level rises above a critical threshold, the system triggers an early-warning notification. This information is used by citizens, schools, or first responder organizations such as the German Red Cross to take precautionary action. We are transmitting the sensor data via LoRaWAN, using a local LoRa gateway to reach cloud dashboards. There, we visualize water levels and compare them with predefined danger zones.

To enhance the reliability and relevance of this data, we intend to combine them with additional sources:

- Weather data, such as rainfall intensity and radar forecasts,
- Other IoT devices, such as humidity or soil sensors,
- Human observations and manual readings,
- And in the future, machine learning models for event prediction and pattern detection.

Together, this creates a participatory, multi-source warning system, empowering local actors not only to receive alerts, but to understand and act upon the situation. We can expand this form of digital resilience infrastructure modularly and adapt to different local contexts, especially in areas affected by past disasters such as the Ahr Valley, where citizens have an urgent need for transparent, local data to rebuild trust and reduce uncertainty during rainfall events. The process is guided by teachers, volunteers, and partners like the German Red Cross (DRK Kreisverband Vulkaneifel e.V.), making it a shared civic-tech endeavor.[8]

Figure 3: Dashboard prototype realized in Thingsboard[9] for visualization and analysis of measurement data with water level thresholds for generating alarm systems

Each device can be programmed visually, including logic such as “transmit water level every 10 minutes unless battery < 3.4V” or “send alert if threshold X is crossed.” However, open questions remain. „How should data integrity be maintained if the device powers off or resets from Deep Sleep?”, “How can intermittent connectivity be handled, especially in

[8] https://messpegel.de/
[9] https://thingsboard.io/

remote areas?”, “What are the best physical mounting strategies, e.g., for attaching gauges to bridges or culverts?”, “Which communication technology (LoRaWAN vs. Wi-Fi) is most viable given local constraints?”.

Currently, deployments serve not only as learning environments, but also as testbeds for applied research. Feedback loops from real-world use help refine both software and hardware design-raising awareness for practical challenges like battery reliability, sensor calibration drift, and firmware robustness.

3.1 Digital Resilience as Public Infrastructure

The project expands the notion of disaster preparedness: resilience today includes sensor literacy, data fluency, and system thinking. Additional modules (e.g., air quality or humidity sensors) can transform school buildings into climate-aware learning spaces. By using Makerspaces as community labs, resilience-building becomes participatory, inclusive, and locally adaptable.

A key aspect, especially in communities affected by past disasters such as the Ahr Valley flood of 2021, is trust in data and local control. Many residents now experience existential anxiety during heavy rain events. It is therefore vital that they are not passive recipients of external alerts but can access and interpret environmental data themselves. The ability to install a sensor at the creek behind one’s home and check live data on a mobile dashboard creates not only technical transparency but psychological security and self-efficacy.

These infrastructures help build not only technical competencies but also social capital. Community workshops, student-led hackathons, and multi-generational repair events promote ownership, responsibility, and collaboration. Once implemented, such sensor networks can scale bottom-up, forming a low-cost, open-source backbone for environmental observatories and community alert systems.

3.2 Outlook: From Monitoring to Self-Empowered Resilience

The long-term vision goes beyond teaching how to build IoT-Systems. It seeks to enable citizens to identify risks of all kinds, deploy and maintain their own solutions, and manage local data autonomously. Upcoming developments include:

- Solder-free kits developed with *Make:magazine* and *Fab-Lab.eu*,
- Plug-and-play IoT[2] Dashboards for non-programmers,
- AI guidance tools for setup, learning and optimization,
- Modular extensions like rain-, soil moisture- or solar sensors.

Over time, the distinction between school project and civic infrastructure will fade. Each

IoT station becomes part of a living, learning network where environmental awareness, digital tools, and community resilience converge. Only speaking of the flood monitoring project, we built and installed more than fifteen flood gauges with schools. Until April 2026 there are at least six more devices planned with the optimized program code (5.1).

4 100 Ideas – A Platform for Creative Exploration

The true potential of the IoT²-Werkstatt lies not in any single use case, but in its ability to unlock creativity and technical exploration - especially among young learners. In hackathons, workshops and labs, hundreds of project ideas have emerged, often surprising educators with their originality and relevance. Unlike conventional development settings, these ideas are rarely overengineered or constrained by legacy thinking. Young people are less "locked in" to academic patterns and often propose intuitive, unexpected solutions to everyday problems. Examples[10] include:

- A smart doorbell that notifies residents via smartphone,
- A connected pill box that reminds users and alerts their doctor on missed doses,
- An ice warning alarm that detects slippery conditions and triggers alerts,
- A DIY ECOTRON greenhouse that simulates climate conditions for plant research using cooling packs, fans, and cloud data input,
- A cocktail-mixing machine controlled by voice commands via Alexa and IoT protocols,
- A smart irrigation system using soil sensors and neural network logic for optimized water use.

These projects display the flexibility of the IoT² platform and its capacity to adapt to diverse learning goals and contexts. Each one is built with plug-and-play sensors and programmed via visual blocks by students as young as fourteen. Importantly, these learners also engage with questions of data ethics: Who owns the data? Should it be shared? What happens to personal or environmental information when it enters the cloud? Through APIs and dashboards, users begin to understand the social layer of digital systems and their role in shaping them. In this sense, the IoT²-Werkstatt is not defined by a specific project like flood monitoring It is an open invitation to explore, design, and act on the world through code, sensors, and imagination.

[10] https://charta-digitale-vernetzung.de/deutschland-intelligent-vernetzt/div-konferenzen/iot-hackathon/

5 Artificial Intelligence for Resource-Efficient IoT

The integration of AI into the IoT²-Werkstatt opens new possibilities - both in optimizing embedded systems and in supporting learners throughout the programming process. Generative AI tools such as ChatGPT are increasingly being used by students, teachers, and developers to reduce energy consumption, improve code quality, and democratize access to digital literacy.

5.1 Energy Optimization through AI

Microcontroller-based IoT devices are often constrained by limited power budgets, especially when battery- or solar-powered. In the DigiSelfTrans project, we used ChatGPT to analyse and refactor the monitoring-code with a specific goal: maximizing battery life. We present ChatGPT the C code generated by the code generator and ask for suggestions for optimization.[11]

By prompting the AI to suggest optimizations, learners were introduced to techniques such as:

- DeepSleep modes that shut down the microcontroller between measurements,
- Dynamic sensor cycles that adjust based on weather conditions,
- Modular power management, e.g. disabling unused peripherals.

The result was a significant improvement in energy efficiency: Average current consumption dropped by 72%, from ~47 mA to ~13 mA.

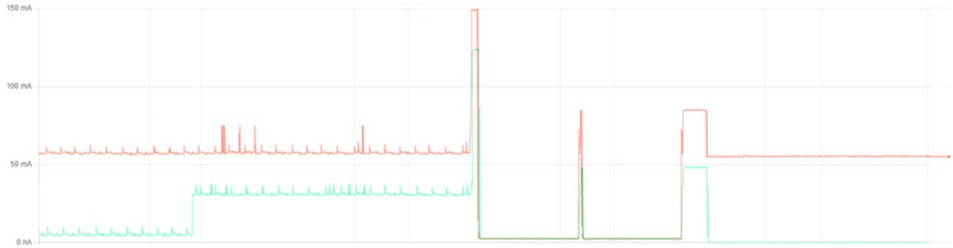


Figure 4: Energy usage overlay-comparison (duration 10s) of the flood gauge before (red line) and after (green line) AI-assisted optimization

This process not only made the systems more practical but also offered a real-world context for learning about software optimization, resource management, and AI-assisted development. The fact that we achieve these results through natural-language interaction with a chatbot demonstrates the accessibility and impact of AI tools in the maker context.

[11] https://www.umwelt-campus.de/iss/projekte/laufende-projekte/digiselftrans/ki-assistent-ressourceneffizienz

5.2 AI in Education and Resilience Empowerment

AI also transforms the way learners approach programming. Superblocks are a middle ground between HLL (High-Level Language - with guaranteed frustration) and Vibe Programming (without learning anything). The focus for the user is entirely on innovation and new ideas. At the same time, it trains algorithmic thinking. In the IoT²-Werkstatt, AI tools serve as digital tutors[12], helping students understand control structures, fix errors, and extend their own ideas. In natural language they can ask the system how to start with projects and ideas with basic questions like "How do I measure temperature every 10 minutes?", "What does this condition mean?" or "How can I define thresholds for my sensor values?" These digital dialogues lower entry barriers by replacing code syntax with conversation. Learners are no longer afraid to make mistakes, and they gain confidence as they watch their questions transform into working code.

Teachers, in turn, use AI to generate custom worksheets and challenges, step-by-step video scripts, podcast outlines or differentiated tasks for students at varying skill levels. AI thus supports inclusive education: it adapts to the learner's pace, language, and prior knowledge. It enables access to informatics even for those who struggle with abstract logic or complex syntax. By making computing conversational, it shifts the experience from memorization to exploration. AI in the IoT²-Werkstatt becomes more than a tool. It is a bridge between creativity and computation, supporting a new generation of learners in becoming not just consumers of technology, but co-creators of their digital future.

6 Conclusion and Outlook

The IoT²-Werkstatt links environmental monitoring, informatics education, and civic engagement. Participants build systems that reflect real-world complexity - gaining skills in problem-solving, systems thinking, and sustainability.

Since its origin at the 2016 IT Summit of the German Federal Government in Saarbrücken, with members of the IoT expert group, committed makers and the Environmental Campus of Trier University of Applied Sciences, the concept was evaluated in hackathons with over one hundred school teams. We integrate the initiative into curricula (e.g., in Rhineland-Palatinate), teacher training, and university courses like "Applied Computer Science & AI" at the Environmental Campus Birkenfeld.

Reliable figures are difficult because these are distributed DIY projects. By spring 2021, around three thousand CO_2 traffic lights were certainly in use in Germany alone (600 in classrooms, 2.000 Heise boards, plus many local projects); worldwide, a scale in the high four- to low five-digit range can be assumed.[13] The IoT²-Werkstatt joined the MINT

[12]https://chatgpt.com/g/g-67ff6858ea088191a94500af1066151c-iot-werkstatt-der-erste-mint-und-make-r-tutor
[13] https://www.umwelt-campus.de/iot-werkstatt/news-detailansicht/5-jahre-diy-co2-ampeln

Alliance of the BMFTR in 2021.[14] Events such as hackathons, workshops and training are especially well received by young people and encourage further creative projects [DG17]. Ongoing dissemination happens through Maker Faire 2025 in Hannover, Make:magazine, and hands-on workshops inviting educators, volunteers, and civic actors to adopt and adapt the tools ensuring that the IoT²-Werkstatt continues to grow as a model for participatory digital resilience.

Acknowledgements

We gratefully acknowledge that the project “DigiSelfTrans” is funded by the German Federal Ministry for Research, Technology and Space (BMFTR) within the program “DATIpilot – Support and Learning for Innovation and Transfer: An Experimental Space” in the Context of DATI.

Bibliography

[GB23] Gollmer, K.-U.; Burger, G.: IoT²-Werkstatt – MINT macht Mut - https://www.umwelt-campus.de/fileadmin/Umwelt-Campus/IoT-Werkstatt/Tutorials_Oktopus/QuickStart_Anleitungen/IoT_MINT_Making_Zukunft_Gestalten.pdf

[Go24] Gollmer et al.: Informatik- Balkonkraftwerk und DIY-Energiewende – Ein Plädoyer für eine Informatik im Kontext in MNU-Journal- Ausgabe 03.2024 – ISSN 0025-5866. Verlag Klaus Seeberger, Neuss.

[Go19a] Gollmer, K.-U.; Burger, G.; Krieger, R. & Wahl, M.: IoT-workshop: Application examples for adult education. Chapter in Big Data Analytics for Cyber-Physical Systems, Machine Learning for the Internet of Things, 1st Edition, Elsevier, Paperback ISBN:9780128166376

[Go19b] Gollmer, K.-U.; Kreten, S.; Stolz, F.; Dartmann, G. & Burger G.: IoT-Workshop: Blueprint for pupils education in IoT, Chapter in Big Data Analytics for Cyber-Physical Systems, Machine Learning for the Internet of Things, 1st Edition, Elsevier, Paperback ISBN:9780128166376

[RM25] Rajuroy, Adam; Johnson, Mary: The Influence of Infrastructure Deficiencies on STEM Education Outcomes and Its Impact on Future Engineering Innovations, https://www.researchgate.net/publication/389073301, 02/2025

[DG17] Digital-Gipfel: Positionspapier IoT-Werkstatt – Expertengruppe M2M / Internet der Dinge, Juni 2017, https://charta-digitale-vernetzung.de//dokumente/deutschland-intelligent-vernetzt/publikationen/Positionspapier%20IoT-Werkstatt.pdf

[14] https://www.bildung-forschung.digital/digitalezukunft/de/bildung/iot