TRAIL TRUSTED AI LABS
BY DIGITALWALLONIA / SPW-RECHERCHE

TRAIL Summer Workshop' 25
Project Proposal

Full Name of Team Leader	Stefano Forti (University of Pisa, Italy)
Project Title	**GreenAItelier: Sustainable AI Systems from Open Data to Policy Integration**
Profile of the Team Leader(s)	Stefano Forti holds a PhD in Computer Science (2020) from the University of Pisa, Italy, where he is currently a tenure-track assistant professor. His research focuses on software engineering and distributed systems, addressing the development, deployment, and orchestration of multi-service applications, with an emphasis on predicting and suitably balancing application performance, operational costs, and environmental footprint. He has co-authored more than 50 peer-reviewed publications, including journal articles, international conference papers, and book chapters. He is associate editor of the ACM SIGSOFT Software Engineering Notes (SEN) and participated or participates in the program committee of various international conferences (CCGRID, SOSE, ESOCC). He participated in several collaborative research projects at the local, national, and international level. Among these, he served as work package leader in the project "Lightweight Self-adaptive Cloud-IoT Monitoring across Fed4FIRE+ Testbeds", cascade-funded by Fed4Fire+ under Horizon 2020, and was awarded a six-month DAAD research scholarship to conduct research on continuous reasoning for managing next-generation distributed systems. He currently delivers a BSc module on Green Computing (6 ECTS) and PhD courses on sustainable ICT. His research explores how next-generation software systems, such as data-intensive AI systems, can be engineered sustainably from software design to system deployment.
Abstract	The GreenAItelier project aims at bootstrapping interdisciplinary collaboration among King's College London, UniVienna, and the University of Pisa to design and prototype methods for assessing and reducing the environmental footprint of AI-based and data-intensive software applications, while accounting for emerging regulatory frameworks. The project focuses on measuring the sustainability impact of software systems through standardised metrics, such as the Software Carbon Intensity (SCI) score, and exploring engineering strategies to reduce carbon emissions while maintaining functional and operational quality. By devising new software engineering and application management techniques, GreenAItelier will develop a first working proof-of-concept demonstrating sustainability-aware AI development practices. Public datasets, such as those provided by ElectricityMaps for carbon intensity of energy grids, and open-source tools like CodeCarbon for emissions tracking, will possibly be adapted and reused within the scope of the project. The two-week research sprint at TRAIL Summer Workshop will deliver a preliminary environmental impact assessment prototype and a case study application, together with a roadmap for integrating sustainability metrics into future AI engineering pipelines, aligned with emerging national and international

	regulatory frameworks. A key aspect of our proposal is its highly inter-disciplinary nature. Despite being rooted in software engineering aspects, participants will be able to work on a fast-evolving research area at the intersection of sustainability science, data science and public policy.
Project Objectives	During the TRAIL Summer Workshop, GreenAItelier intends to lay the foundations to design and prototype a framework for assessing and reducing the environmental footprint of AI-based and data-intensive software applications. The project will focus on selecting, adapting and integrating standardised sustainability metrics into the evaluation of selected AI tasks. This will involve developing a first proof-of-concept tool (TRL 2-3) capable of measuring carbon emissions associated with model training and inference phases, while exploring software engineering and application management strategies to reduce emissions without compromising functional/operational quality. Experts in sustainability regulatory frameworks will ensure that the approach is aligned with emerging national and international policies on green ICT. This will enable offering a regulatory perspective to participants and going towards developing decision-making framework for policy makers, and an interesting point of discussion with other participants in TRAIL workshop. The current status of the project is conceptual. Therefore, by the end of the Summer Workshop, GreenAItelier aims to achieve the following objectives: • Develop a first methodology to assess the environmental footprint of AI-based multi-service applications, • Deliver a preliminary environmental impact assessment prototype and a case study application, based on open datasets and open-source tools, and • Define a roadmap for integrating sustainability metrics into AI development pipelines and identify future research directions on this topic, strengthening collaboration among King's College London, UniVienna, and the University of Pisa, within the CircleU alliance. To achieve its goals, GreenAItelier seeks for participants with expertise or strong interest in software engineering, data science and AI, and environmental sustainability. Some familiarity with (Python) programming, data processing/integration and working with REST APIs is beneficial and strongly recommended. During the project, participants will develop skills related to green software engineering, particularly those related to identify which elements of a software architecture and deployment impact on an application's carbon footprint and will develop new models to assess it. Besides, they will learn how to align sustainable AI workflow development with legal requirements and policy frameworks, understanding where relevant information can be found and how to use/interpret it. Naturally, teamwork, communication and collaboration skills will be strongly exploited during the workshop. The research nature of the workshop will require participants to engage in discussion and collaborative problem-solving so as to co-design the proof-of-concept methodology and its implementation.

Project Dataset	GreenAItelier does not involve training new AI models. It focuses on assessing and reducing the carbon footprint of existing AI workflows (training and inference). Specifically, it will adapt and apply the Software Carbon Intensity (SCI) metric developed by the Green Software Foundation. This includes: • Estimating emissions from training/inference workloads • Evaluating model execution in different hardware/grid contexts • Optionally integrating with open-source estimation tools or methodologies such as CodeCarbon GreenAItelier will need data about the carbon intensity of the energy that powers up different computational resources that run the considered AI workflows. To this end, GreenAItelier will leverage (dumps of) publicly available datasets. In particular, the project will use the ElectricityMaps dataset (https://www.electricitymaps.com/), which provides real-time and historical data on carbon intensity and electricity production mix, available through an API and downloadable in JSON and CSV formats. This dataset includes timestamps, geolocation, emission factors, and energy source breakdowns, enabling accurate estimation of emissions based on grid conditions. Additionally, the project plans to reuse/adapt CodeCarbon (https://mlco2.github.io/codecarbon/), an open-source Python library that estimates CO2 emissions from software execution based on hardware usage and local electricity carbon intensity. Both resources will be possibly integrated into the project's prototype to support the measurement of the Software Carbon Intensity (SCI) score and the environmental impact assessment of a case study AI application workflow. Ideally, we will discuss and select at least a use case from those proposed by participants and/or provide use cases from our research areas and/or the state of the art, e.g. UC Irvine Machine Learning Repository- Health datasets; CUSP London data repository (under KCL's TRE can share with participants through dedicated link) - Mobility & GHG datasets; other contextual data on CUSP Github (https://github.com/cusp-london/) openly accessible - demographic, socio-economic, ENCODE genomic datasets (https://www.encodeproject.org).
Background Information	The environmental footprint of software systems is an increasing concern, as the ICT sector accounts for approximately 2% of global carbon emissions and could grow to 20% of global electricity demand by 2030 (Danushi et al., 2025). With the slowdown of Moore's Law and Dennard Scaling, improving ICT sustainability now requires interventions at the software level. This growth is especially apparent in AI systems (Tabbakh et al., 2024; Patterson et al., 2021; Patterson et al., 2022; Faiz et al., 2024). Recent initiatives, such as the Software Carbon Intensity (SCI) specification (Green Software Foundation, 2024), provide standardised methods to assess and reduce software-related emissions, but systematic approaches for AI systems remain limited. Research on sustainable software engineering has produced various but fragmented methodologies, covering green coding practices, carbon-aware deployment strategies, and lifecycle-based footprint assessments. However, few initiatives integrate environmental metrics into the design, development, and operational management of AI-based software applications. Moreover, alignment with emerging regulatory frameworks on green ICT, such as those under the European Green Deal, remains an open challenge.

	GreenAItelier aims at addressing these gaps by developing a framework to combine sustainability measurement and regulatory compliance. By relying on public datasets (e.g., ElectricityMaps) and open-source tools (e.g., CodeCarbon), the project will deliver a first proof-of-concept for sustainability-aware AI development. Its interdisciplinary focus will contribute to advancing the broader research landscape of sustainable, carbon-aware and carbon-efficient software, while fostering new collaborations across King's College London, UniVienna, and the University of Pisa.
Bibliographic References	- Danushi, O., et al. (2025). Carbon-Efficient Software Design and Development: A Systematic Literature Review. ACM Computing Surveys. https://doi.org/10.1145/3728638 - Tabbakh, A., et al. (2024). Towards sustainable AI: A comprehensive framework for Green AI. Discover Sustainability, 5, Article 408. https://doi.org/10.1007/s43621-024-00641-4 - Green Software Foundation. (2024). Software Carbon Intensity Specification v1.1.0. https://sci.greensoftware.foundation/ - Patterson et al. (2021), Carbon Emissions and Large Neural Network Training - Patterson et al. (2022), The Carbon Footprint of Machine Learning Training Will Plateau, Then Shrink - Luccioni et al. (2022), Estimating the Carbon Footprint of BLOOM, a 176B Parameter Language Model - Faiz et al. (2024), LLMCarbon: Modeling the end-to-end Carbon Footprint of Large Language Models
Detailed Work Plan	GreenAItelier will be structured in four phases over the two-week TRAIL Summer Workshop. Stefano Forti (University of Pisa) will give two to three introductory lectures on models to estimate the energy consumption and carbon emissions of ICT systems and software. Next, the team will coordinate development, focussing on adapting SCI metrics to AI workflows. Francesca Vantaggiato (King's College London) will align the methodology with emerging environmental policy frameworks. Yi-Jing Li (King's College London) will contribute expertise on urban digital systems and potential smart city applications. Andrea Tanzer (UniVienna) will support data preparation and footprint measurement for large-scale datasets. Overall, the work plan is structured into four phases: 1. Introductory lectures and framework requirements definition 2. Prototype development and testing 3. Impact assessment and regulatory alignment 4. Final presentation and roadmap for future research Particularly, in the initial days, the team will finalise the methodological framework and define a case study. Participants will be asked to propose and analyse a use case AI workflow and identify which data about the workflow execution and

	deployment is needed so as to compute the associated carbon footprint (e.g. amount of data to be stored, datacentre location, hardware used for training/inference, ...). This will allow to collaboratively define a general model for determining the environmental footprint of AI workflows (i.e. the emissions it produces and, if relevant, the emissions it reduces so to get a net figure). A first proof-of-concept prototype will be developed and tested during the second phase. This will be done by adapting the SCI computation to the defined model and to code an actual service that is capable of processing the relevant input data and output the results, also enabling what-if analyses (e.g. "what if I use this datacentre to deploy my service, instead of the current one? Would I reduce the carbon footprint of my deployment?"). The third phase will focus on a preliminary environmental impact assessment and regulatory compliance analysis of the proposed model. This will enable introducing regulatory aspects in our framework and understand how to make it useful/suitable to support (local, national, international) policy decision-makers. The final phase will consolidate results into a presentation and a roadmap for future interdisciplinary research. Key deliverables will include a conceptual framework (Day 3), a proof-of-concept (Day 8), a preliminary impact assessment over at least a use case (Day 11), and a final project roadmap and presentation (Day 14).
Other Remarks	The project will make use of standard computational resources, suitable for developing and testing Python code: lightweight AI models or workflows and monitoring and assessment tools like CodeCarbon. To this end, we ask participants to bring their own laptops, and we will use free Cloud resources (e.g., Google's CoLab), if needed. We need Internet connection for all participants.