Innovating for Tomorrow
The Convergence of Software Engineering and Green AI

Cruz, Luís; Franch, Xavier; Martínez-Fernández, Silverio

DOI
10.1145/3712007

Publication date
2025

Document Version
Final published version

Published in
ACM Transactions on Software Engineering and Methodology

Citation (APA)
Cruz, L., Franch, X., & Martínez-Fernández, S. (2025). Innovating for Tomorrow: The Convergence of Software Engineering and Green AI. *ACM Transactions on Software Engineering and Methodology*, *34*(5), Article 138. https://doi.org/10.1145/3712007

Innovating for Tomorrow: The Convergence of Software Engineering and Green AI

LUÍS CRUZ, Delft University of Technology, Delft, The Netherlands
XAVIER FRANCH, Universitat Politècnica de Catalunya, Barcelona, Spain
SILVERIO MARTÍNEZ-FERNÁNDEZ, Universitat Politècnica de Catalunya, Barcelona, Spain

The latest advancements in machine learning, specifically in foundation models, are revolutionizing the frontiers of existing software engineering (SE) processes. This is a bi-directional phenomenon, where (1) software systems are now challenged to provide AI-enabled features to their users, and (2) AI is used to automate tasks within the software development lifecycle. In an era where sustainability is a pressing societal concern, our community needs to adopt a long-term plan enabling a conscious transformation that aligns with environmental sustainability values. In this article, we reflect on the impact of adopting environmentally friendly practices to create AI-enabled software systems and make considerations on the environmental impact of using foundation models for software development.

CCS Concepts: • **Software and its engineering**; • **Computing methodologies** → **Artificial intelligence**;

Additional Key Words and Phrases: Green AI, Green Software, Sustainability, Software Engineering

ACM Reference format:
Luís Cruz, Xavier Franch, and Silverio Martínez-Fernández. 2025. Innovating for Tomorrow: The Convergence of Software Engineering and Green AI. *ACM Trans. Softw. Eng. Methodol.* 34, 5, Article 138 (May 2025), 13 pages.
https://doi.org/10.1145/3712007

1 Introduction

Software-related CO_2 emissions from the ICT sector currently account for 2.1–3.9% of global emissions [21]. In today's context with the widespread use of AI systems, there have been many calls from industry leaders and AI experts admitting a further increase of these emissions. OpenAI chief executive Sam Altman warned that the next wave of generative AI systems will consume vastly more power than expected [10]. Hugging Face Climate Lead Sasha Luccioni has shown the scalability impact of AI systems' inference: "While inference on a single example requires much less computation than that required to train the same model, inference happens far more frequently than model training—as many as billions of times a day for a model powering a popular user-facing product such as Google Translate" [33]. Focusing only on AI systems, in a middle-ground scenario, by 2027 AI servers could use between 85 and 134 terawatt hours (Twh) annually [14].

This work has been partially supported by the GAISSA project (TED2021-130923B-I00, which is funded by MCIN/AEI/10.13039/501100011033 and by the European Union "NextGenerationEU"/PRTR).
Authors' Contact Information: Luís Cruz (corresponding author), Delft University of Technology, Delft, The Netherlands; e-mail: L.Cruz@tudelft.nl; Xavier Franch, Universitat Politècnica de Catalunya, Barcelona, Spain; e-mail: xavier.franch@upc.edu; Silverio Martínez-Fernández, Universitat Politècnica de Catalunya, Barcelona, Spain; e-mail: silverio.martinez@upc.edu.

ACM 1557-7392/2025/5-ART138
https://doi.org/10.1145/3712007

That is similar to the annual electricity usage of countries such as Argentina, the Netherlands, and Sweden individually and constitutes approximately 0.5% of the world's current electricity consumption [19].

The convergence of **software engineering (SE)** and AI is a bi-directional phenomenon raising key sustainability concerns. On the one hand, software systems of a plethora of domains are now challenged to provide AI-enabled features to their users. To reduce the emissions from these AI systems from their development and training, to their usage and inference, and to its retirement, emerging AI software development lifecycles shall incorporate energy-aware capabilities. On the other hand, the potential of using AI and foundation models to automate tasks within the software development lifecycle is not just promising but becoming a reality, as reported in dedicated events such as AIware[1] and LLM4Code [2]. To fully embrace these technologies, the concerns regarding the emitted emissions from its usage shall be understood.

Therefore, the objective of this position article is twofold:

(1) Identifying the trans-disciplinary dimensions and dichotomies in which the research from the SE community shall contribute to build greener AI systems, as well as reasoning on the evolution of SE practices in such dimensions and dichotomies.
(2) Discussing the environmental sustainability of one application domain of AI systems: generative AI for SE tasks like generation of requirements, architecture, or code in which humans and intelligent agents jointly create software.

2 Overview of Green AI

We define *Green AI* as a trans-disciplinary field that aims to make AI systems environmentally sustainable. Environmental sustainability of software (including AI systems) refers to engineering systems having minimal impact in our planet throughout their whole lifecycle [4]. We distinguish this from AI for Sustainability or AI for Green, where AI is used to make different domains more environmentally friendly—e.g., using AI to make agriculture more sustainable. We argue that it is important to make a clear separation between Green AI and AI for Green as they require different approaches and different scientific backgrounds.

Given the complexity of AI systems, Green AI needs to be tackled from different angles, having each of these angles an important contribution to the overall carbon emissions of the systems. Data, AI models, and the code of the software systems are the foundations of AI systems [49]. As such, to build and maintain sustainable AI systems, we shall focus in each of these three aspects. Consequently, as depicted in Figure 1, we divide Green AI across three major dimensions: data-centric, model-centric, and system-centric. These dimensions are pinpointed below in Sections 2.1, 2.2, and 2.3. The figure also intersects Green AI with two dichotomies: hardware vs. software and reporting and monitoring vs. best practices.

Hardware ↔ Software. With the current trend of specializing hardware for particular AI tasks, practitioners can no longer be agnostic of the hardware they use to run their software. The choice of hardware is essential to ensure that a particular AI pipeline runs properly with no issues downstream. Hence, practitioners now face the challenge of having to be hardware experts. Moreover, previous research demonstrates how different hardware configurations can significantly impact the energy consumption of training deep-learning models [15]. However, these differences are not obvious and require meticulous tradeoff analysis, which is sub-optimal in real-world scenarios. Hardware design for AI poses more concerns. Since AI chips are specialized to particular AI tasks,

[1]https://2024.aiwareconf.org/ (visited on 5 April 2024).
[2]https://llm4code.github.io/ (visited on 5 April 2024).

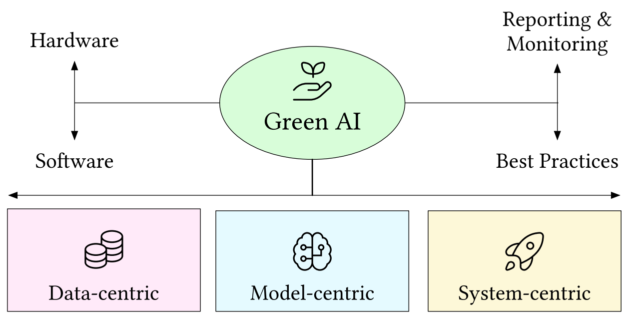

Fig. 1. Overview of Green AI.

updates to AI models might lead to different hardware requirements. This can significantly reduce the lifetime of hardware, posing challenges to the emerging "Right to repair" legislation efforts being discussed by the EU, the US, the United Nations and so on. Not surprisingly, the waste created by disposing hardware devices—coined as e-waste—is already a pressing environmental problem, as reported by the technology activist Murzyn.[3]

Reporting and Monitoring ↔ *Best Practices.* It is important that software systems are designed in a way that it is possible to report and monitor sustainability indicators [30]. Practitioners need to provide accurate information about the carbon footprint of their experiments and models. This can be done, e.g., via meta-data in repositories such as Hugging Face or using more elaborated and structured templates such as the concept of ID-card [3]. On the other side of the spectrum, we have best practices [25]. They provide practitioners with strategies that can be adopted to reduce the environmental footprint of their software. These two poles complement each other. Effective best practices can only be established when reliable measurement mechanisms are in place. Monitoring and reporting not only incentivize practitioners to adopt these best practices but also provide essential feedback on their implementation and effectiveness.

2.1 Data-Centric Green AI

Data-centric Green AI revolves around preparing data in a way that is expected to reduce the overall energy footprint of AI systems without hindering their performance. Previous work has addressed this challenge in multiple ways. Reducing data dimensionality, using feature selection or stratified random sampling already leads to linear gains in energy consumption [58].

Further strategies are being explored. Active learning and coreset extraction have promising potential by identifying the most valuable examples that contribute to the learning process and excluding redundant or non-informative ones with confidence [48].

Knowledge transfer/sharing consists of using knowledge rather than solely relying on data to train a model. For instance, existing pre-trained models can be used to train new models. Preliminary research showcases improvements in energy consumption by a factor of 15 with this technique [61].

Dataset distillation or dataset condensation involves synthesizing a smaller dataset derived from the original dataset, aiming to train a model on this reduced set while achieving test accuracy comparable to that of a model trained on the original dataset. This approach differs from coreset extraction by focusing on synthesizing informative samples rather than selecting existing ones.

[3]Murzyn is on a mission to report the hazards from handling electronic waste, a responsibility that has been delegated to Global South countries: https://murzyn.asia/en/mission-00-24-42/ (visited on 5 April 2024).

Data distillation has been used to reduce the size of the MNIST image set [32] into just 10 synthetic distilled images (one per class) and achieve close to original performance [62].

Curriculum learning consists of presenting training examples to the model in a specific order, starting with simpler examples and gradually increasing the difficulty of the examples as training progresses. This strategy provides the model with a structured learning experience that mimics how humans learn and requires less iterations to converge, with time reductions of 70% [41].

2.2 Model-Centric Green AI

Model-centric Green AI means developing experimental research to improve the AI model performance and energy efficiency. The target is to build and optimize AI models that can achieve similar outcomes while requiring fewer resources.

A notable example is the BLOOM model, whose carbon footprint have been studied during the training and inference stages to compete with other large language models [34]. As the design of the model architecture and its number of parameters affect the energy consumption [8], research on lighter, less data-intensive, and less energy-consuming AI models and architectures is increasing. Examples range from lightweight architectures, to resource-aware neural architecture search. A more recent example are small language models, like "Phi" that has shown performance comparable to models 5× larger [26]. Indeed, there is evidence that there is not always a tradeoff between green and performance metrics. A mining repository study on 1.417 models from Hugging Face could not find a correlation between model performance and carbon emissions of **machine learning (ML)** models [7]. This shows the potential for lightweight AI models architectures [8, 37].

Another model-centric strategy is the compression of existing models (aka AI model optimization): structured and unstructured pruning, quantization and binarization, efficient training and inference of models under different resource constraints. In this regard, Pytorch and TensowFlow are already offering AI models optimization libraries [45, 57].

2.3 System-Centric Green AI

System-centric green AI refers to the decisions we can take regarding the software architecture and serving of ML systems to make them environmentally sustainable. The ecological footprint of an AI model extends beyond its algorithmic design to encompass the entire system infrastructure that supports it. For instance, suboptimal hardware choices can significantly inflate energy consumption [15]), while implementing batched inference can enhance energy efficiency [64]. Sometimes, even the choice of the container base image, can have a dramatic impact on the overall energy footprint [46]. It is also essential to recognize the continuum from cloud to edge computing in considering energy footprints [16]. This continuum encompasses end-user devices (edge), cloud servers, and the network infrastructure that connects them. Each component plays a role in determining the overall environmental impact of AI systems, highlighting the need for holistic considerations in system design and deployment for sustainability.

We anticipate that sustainability challenges will initially be tackled with a system-centric approach, given the overlap between AI systems and traditional software systems. However, we argue that emerging challenges lie in the data-centric and model-centric domains of AI, where SE plays a critical role [51, 52]. The significance of data quality has already been acknowledged for its impact on the reliability, robustness, efficiency, and trustworthiness of modern software systems [39]. From a model-centric perspective, there is an unexplored potential to employ SE to challenge the *status quo*. For example, adopting energy-conscious practices for model adaptation [43], hyperparameter tuning [65], and so on.

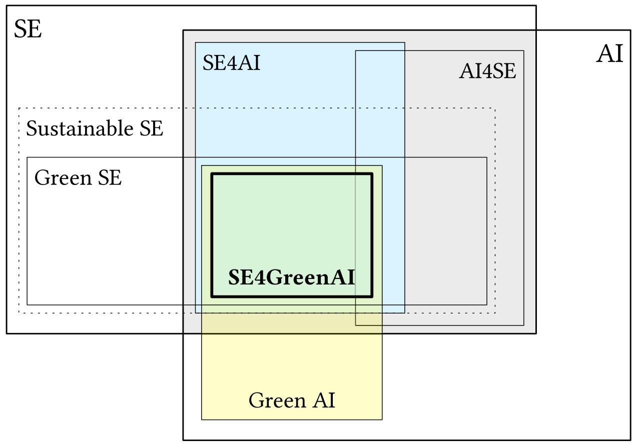

Fig. 2. Diagram of research fields related to SE and Green AI. Subfields are represented by shapes nested within larger shapes, indicating their relationship to the broader field. The size of the shapes is not indicative of any specific meaning.

2.4 Intersection between SE and Green AI

We foresee the importance of bringing the discipline of Green Software to the forefront when defining and operating Green-AI requirements in AI projects. SE methods are essential for addressing Green AI across all three dimensions—data, model, and system. The SE community, in particular, should take a leading role in addressing the software side of the hardware- vs. -software dichotomy, ensuring that sustainable practices are embedded throughout the software lifecycle.

A general view of SE fields related to Green AI is mapped in Figure 2. It begins with the combination of AI and SE, leading to the emergence of the subfields AI4SE and SE4AI. The latter, highlighted in blue, focuses on practices, methodologies, and techniques within the field of SE specifically tailored for the development and deployment of AI-enabled systems. The combination of this field with Green AI results in the main topic of this article, labeled in the picture as SE4GreenAI—highlighted in green. Notably, this intersection inherits practices and concerns from Green Software and its parent, Sustainable SE, as depicted by the figure.

3 Reflections on the Future of SE for Green AI

When reflecting about the future of SE, the emergence of Green AI introduces several challenges that warrant reflection. In this section, we delve into these potential challenges, providing *context* and outlining their *implications for SE.*

3.1 Consideration of the Business Case

Context. Not all problems are alike. Diverse business cases require different positioning with respect to environmental sustainability. An image processing model for cancer diagnose shall prioritize precision no matter what is the carbon footprint required for training. On the contrary, a film affinity recommender system will hardly be considered life-critical, therefore little gains of accuracy at the cost of dramatic increments of energy consumption are not justified; instead, digital sobriety[4] should start to be the norm, and not the exception.

[4]https://www.vox.com/climate/2024/3/28/24111721/ai-uses-a-lot-of-energy-experts-expect-it-to-double-in-just-a-few-years (visited on 5 April 2024).

Implications for SE. Business case elaboration and analysis lies at the heart of the requirements engineering area [20]. Its application to AI systems, often denoted by RE4AI, is currently a mainstream in SE, yielding to a number of actionable findings, e.g., definition of new scopes for requirements [54] or new quality requirement types [24]. However, as Habibullah et al. remark, there is a research gap in establishing tradeoffs among quality requirements [24]. Understanding and specifying these tradeoffs would help to translate the business case into quantifiable requirements, associated metrics, and optimization functions allowing for a proper validation of the AI system.

3.2 Consolidation of Fundamental Concepts

Context. Design of AI systems is still a young research area. As a consequence, fundamental concepts are still not fully understood and lead to inconsistent or even wrong terminology. For instance, we may find papers using the terms "energy efficiency" with the meaning of "energy consumption," or defining metrics with an incorrect measurement unit. This situation hampers communication and interdisciplinary collaboration [38].

Implications for SE. Although there are several excellent works that focus on establishing a consolidated terminology for environmental sustainability [22], the truth is that they have not made their way in the SE community. Observing what happens in other areas, a standard playing the same role as for instance ISO/IEC 25010 in the field of software quality [2] could be a significant asset towards this consolidation. Eventually, such a standard could be the upcoming ISO/IEC 20226 on AI environmental sustainability.[5]

3.3 Monitoring Sustainability

Context. Collecting energy data is not a trivial task. Even after collecting it, practitioners have to process and analyze energy data and tracing it back to the AI pipelines, to hotspots and make necessary adjustments if possible. To make matters worse, improving sustainability is often a tradeoff problem: we want to reduce energy consumption without hindering other requirements of the system. For example, improving energy efficiency at the cost of privacy is likely unacceptable, depending on the business case. On another note, the impacts of AI systems on the environment can hardly be simplified by looking at energy consumption alone. For the same energy consumption of executing a AI software, the carbon footprint can vary depending on the time, region, and the type electricity used to power the servers at that particular. Besides carbon footprint, we also have the problematic of water footprint—i.e., the water being evaporated due to the software execution; mostly due to cooling down servers. Moreover, embodied carbon footprint also plays an important role. GPU usage needs to be optimized to make sure hardware is not stalling idle most of the time during their lifetime [63].

Implications for SE. There is an open challenge of developing user-friendly energy monitoring tools that seamlessly integrate with various environments, including Edge AI devices, and virtual ones like Docker containers. These tools should not only enable detailed analysis but also offer a comprehensive overview of the energy consumption associated with software products. Additionally, they need to accommodate diverse resource metrics and provide actionable insights to facilitate environmentally conscious decision-making throughout the development and maintenance phases of these systems [22].

[5] https://www.iec.ch/blog/importance-sustainable-ai (visited on 5 April 2024).

3.4 Compliance

Context. Beyond environmental sustainability, the rise of AI technologies has raised significant concerns in society, such as privacy, security, fairness, and transparency. In response, task forces worldwide have been formed to establish regulations that AI models in production must comply with. One prominent example is the European Union's AI Act.

Regarding sustainability, Recital 46 (amendments 81–83) of the AI Act states: "*AI systems should [...] reduce energy consumption, resource use, and waste, as well as improve their energy efficiency and overall system performance. [...] The design of AI systems should enable the measurement and logging of energy and resource consumption at each stage of development, training, and deployment. [...] The Commission should develop a common specification for the methodology.*"

However, a major limitation in this regulation is that, to the best of our knowledge, the specific **key performance indicators (KPIs)** and methodologies for reporting sustainability have not yet been defined, leaving this as an open challenge for organizations.

Implications for SE. The discipline of SE has a long history of addressing compliance in critical domains. While the specific low-level techniques are different (e.g., measuring concept drift is relatively new to SE), the system and practice of presenting KPIs to stakeholders and providing alerts and having a team ready to act on incidents is something that has been addressed by SE for many years. For example, incident management is a requirement for compliance in Fintech software organizations [27, 28] and is not necessarily different when it comes to AI products. It is therefore important to avoid reinventing the wheel and create a synergy between AI and SE where existing SE practices are tuned to AI-based software systems.

3.5 Clarification of Roles' Involvement

Context. The emergence of AI demands knowledge and skills that go beyond those traditionally required in classical SE. The consideration of new roles and the interactions among them have been subject of investigation. Specially, the fit of data scientists in the SE team has been thoroughly investigated [29, 60]. In the context of Green AI, knowledge on environmental sustainability is a must for the optimal consideration of carbon footprint in system design. However, experts on sustainability are hardly involved in the development of AI systems.

Implications for SE. Same as sometimes software engineers criticize data scientists for building code not adhering to SE principles, software engineers may be criticized by addressing environmental sustainability without having profound expertise in the team. The role of environmental sustainability expert should be explicitly recognized in the Green AI development team, same as data scientists or domain experts are. This role can be played either by real sustainability professionals or by software engineers who have acquired along time the necessary knowledge and skills.

3.6 Changes in the ML Lifecycle

Context. There have been ongoing efforts to reshape the development lifecycle of modern software systems to cover the inclusion of AI-enabled features. However, along with this push to redefine the development lifecycle of AI-enabled software, environmental sustainability is often overlooked [23, 31, 35, 47].

Implications for SE. It is important to add sustainability concerns across the whole lifecycle, from the very early stages until the retirement of software and/or respective AI models. There is an unexplored potential to employ SE to challenge the status quo. For example, model retraining is still the default approach for model adaptation, leaving out other strategies that may be more cost-effective and sustainable [42–44]. Moreover, adhering to the three R's principle—reduce, reuse, and recycle (in that specific order)—is essential in fostering sustainability practices within AI systems [5]. We need frameworks that help practitioners (1) opt for competitive alternatives to AI when available (reduce), (2) consume existing models instead of training their own (reuse), and (3) monitor model degradation and consider model adaptation strategies (recycle) before disposing a model and training a new one [43].

3.7 Quest for Open Science

Context. Nowadays, the SE community is fully aware of the importance of open science as a basic principle supporting transparency and replication [53]. Initiatives at all levels (journals, conferences, national research programs, ...) push the general adoption of open science by researchers. In the AI field, it has become customary to give access to (preprint of) papers, data, software and models in public repositories such as arxiv, Zenodo, GitHub and Hugging Face. However, there is a lack of clear guidelines on what type of information to store in these repositories. For instance, in the case of Green AI, Castaño et al. report that carbon emission-related information is only marginally reported in the most popular model repository, Hugging Face [9]. This means that in spite of community willingness to go open, it is still difficult to replicate experiments or to conduct long-term cohort studies.

Implications for SE. The research community shall produce consolidated and agreed guidelines to allow software engineers to be systematic and rigorous in the documentation of the sustainability dimension of their experiments. Two different actions in this direction are needed: (i) determine what is the information related to environmental sustainability to be described, (ii) creating domain-specific description sheets that can be used to consolidate such information. An example in this direction is the ID-card proposed by Abualhaija et al. in the natural language processing domain [3]. Another example focusing on environmental sustainability is GAISSALabel, which gathers a repository of training and inference emissions footprints of ML models [17]. Alternatively, the information can be coded as meta-data in the repositories, although usually, this will result in less comprehensive descriptions.

3.8 Democratic Sustainable AI

Context. Large software enterprises are in a frenetic race to pioneer groundbreaking AI solutions while concurrently striving to enhance the environmental sustainability of their operations. However, the considerable upfront costs associated with implementing environmentally friendly solutions can dissuade businesses from making the transition, despite the potential for long-term cost savings. Smaller organizations, such as **small and medium-sized enterprises (SMEs)**, often constrained by tight budgets and limited resources, encounter heightened barriers to entry in adopting green AI technologies.

Implications for SE. It is essential to foster an inclusive AI ecosystem that empowers not only tech giants but also SMEs to embrace sustainability—make sustainable AI a democratic ambition. While it is anticipated that software organizations will increasingly gravitate towards sustainable technologies, as a community we ought to ensure that entry into sustainable AI is readily accessible and not leveraged as a strategy to monopolize the AI ecosystem. Development frameworks should inherently support sustainability practices without imposing significant additional costs or effort.

3.9 The role of Education

Context. Equally important is the imperative to educate the next generation of AI developers in sustainable AI literacy. Green SE is still a niche topic [11], with very few available materials to support educators. Moreover, the rapidly evolving landscape of AI technologies presents challenges in adapting AI and Computer Science curricula to incorporate Green AI practices. In fact, many repositories of Green AI practices are very recent and experimental [25], but evidence shows that adoption of Green AI practices is emerging [13].

Implications for SE. The education of SE must explicitly incorporate topics on responsible AI and, more specifically, the environmental considerations of AI systems. This is viable, as SE education has made strong advances on the quality attributes for AI systems, such as maintainability and trustworthiness. Educational materials on sustainability as quality requirement [6] should be openly accessible to empower aspiring software engineers to prioritize it within their organizations.

3.10 Construction of Theories

Context. There is a large corpus of research works that report experiments of any kind with an environmental sustainability dimension. In fact, the original set of 98 studies considered by Verdecchia et al. in their 2023's paper [59] is falling short today after only a couple of years. While every individual experiment may be valuable *per se*, their results still remain mainly de-aggregated and are difficult to be integrated into a holistic body of knowledge [36], i.e., a theory [55].

Implications for SE. Software engineers shall gradually adopt what Stol and Fitgerald call "*a theory-focused research approach*" [56]. Building theories is not only important because of the outcome, but also a theory construction process brings the need for critical thinking to researchers and makes them aware of the need to better report their experiments so as to allow building theories on a solid basis. A critical point is the consideration of context [18], determined by the independent and confounding variables of the experiment. Modeling context in the theory may facilitate, for instance, integrating the source of energy in the calculation of carbon emissions.

4 Green Foundation Models for SE

The usage of large language code models is a powerful tool that is becoming more and more popular. The context of SE is not an exception, with many interesting applications on program repair, vulnerability detection, code generation, and so on. It has been reported recently by Microsoft that, among GitHub Copilot users, 40% of the code they commit is "AI-generated and unmodified"[6] Foundation models open the stage for developing software systems without developers creating a single line of code. Whole software systems can soon be generated from natural language prompts.

A wide adoption of foundation models in coding activities poses an important question on the impact it has on the environmental sustainability of SE [50]. On one hand, there is a clear reduction of number of developers needed for equivalent tasks, which can lead to smaller carbon footprints inherited from having teams working together. On the other hand, there is a concerning energy overhead from continuously prompting code generation models to converge to a version of the software that can go to production. We need paradigms that minimize trial-and-error iterations to prevent unnecessary effort and resources from being wasted on highly complex AI models. One potential solution is the use of a context-aware mixture of experts [40], which can help avoid defaulting to large models when simpler ones could effectively accomplish the task.

Furthermore, it is not clear whether generated code will abide by energy efficiency coding practices [12]. SOTA code generation models are purely statistical: we argue that the most frequent coding patterns are not necessarily the most energy efficient.

[6] Scott Guthrie, Executive of the Cloud and AI group at Microsoft, as of 5 April 2024: https://www.microsoft.com/en-us/Investor/events/FY-2023/Morgan-Stanley-TMT-Conference.

Hence, there is a pressing need for a more in-depth examination of the technical and sustainability stemming from AI-generated software code. Foundation models ought to provide transparency in terms of sustainability indicators at different stages of the software lifecycle. Such indicators should paint a clear picture on the ecological footprint of training those models, using them, but also using their outputs. Strategies should be adopted to make sure these models are more likely to learn energy-efficient coding practices than regular ones.

5 Conclusion

In this article, we explore the key implications of Green AI for the SE community and highlight several areas where it can have a significant impact, such as compliance, monitoring, stakeholder engagement, education, and among others. This impact is a real need for the industry, with examples with a high involvement from practitioners such as the Green Software Foundation [1], promoting sustainability to become a core priority to software teams, just as important as performance, security, cost and accessibility.

The above reflections highlight the need for the SE research community to take a leading role in Green AI and advocate for a research agenda that embraces the challenges of developing environmentally friendly AI software systems.

We argue that more research in SE is essential to support Green AI practices and establish them as a standard in AI-driven software development. However, we also anticipate a rebound effect, where Green AI may struggle to reduce the overall environmental footprint in the short term, as it will likely meet the increasing demand for AI services. Despite this challenge, promoting Green AI remains a crucial mission for researchers, practitioners, and society as a whole.

References

[1] SSIA Joins the Green Software Foundation. 2024. Green Software Foundation—greensoftware.foundation. Retrieved October 18, 2024 from https://greensoftware.foundation/

[2] ISO/IEC 25010:2023. 2023. Systems and Software Engineering—Systems and Software Quality Requirements and Evaluation (SQuaRE). Retrieved from https://www.iso.org/standard/78176.html

[3] Sallam Abualhaija, Fatma Basak Aydemir, Fabiano Dalpiaz, Davide Dell'Anna, Alessio Ferrari, Xavier Franch, and Davide Fucci. 2023. Replication and verifiability in requirements engineering: The NLP for RE case. arXiv:2304.10265. Retrieved from https://doi.org/10.48550/ARXIV.2304.10265

[4] Christoph Becker, Ruzanna Chitchyan, Leticia Duboc, Steve Easterbrook, Birgit Penzenstadler, Norbert Seyff, and Colin C. Venters. 2015. Sustainability design and software: The Karlskrona manifesto. In *Proceedings of the IEEE/ACM 37th IEEE International Conference on Software Engineering*, Vol. 2. IEEE, 467–476.

[5] Sadia Sultana Bristy, Tooba Azam, Mukta Marion Islam, Raisur Rahman, Ahmed Wasif Reza, and Mohammad Shamsul Arefin. 2023. Green Cloud computing: A sustainable energy-efficiency approach for business rapidity and the environment. In *Proceedings of the International Conference on Intelligent Computing & Optimization*. Springer, 312–327.

[6] Coral Calero, Ma Ángeles Moraga, Manuel F. Bertoa, and Leticia Duboc. 2015. Green software and software quality. In *Green in Software Engineering*. C. Calero and M. Piattini (Eds.), Springer, 231–260.

[7] Joel Castaño, Silverio Martínez-Fernández, Xavier Franch, and Justus Bogner. 2023. Exploring the carbon footprint of hugging face's ML models: A repository mining study. In *Proceedings of the ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM '23)*. IEEE, 1–12. DOI: https://doi.org/10.1109/ESEM56168.2023.10304801

[8] Roger Creus Castanyer, Silverio Martínez-Fernández, and Xavier Franch. 2024. Which design decisions in AI-enabled mobile applications contribute to Greener AI? *Empirical Software Engineering* 29, 1 (2024), 2.

[9] Joel Castaño, Silverio Martínez-Fernández, Xavier Franch, and Justus Bogner. 2023. Exploring the carbon footprint of hugging face's ML models: A repository mining study. In *Proceedings of the ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM '23)*, 1–12. DOI: https://doi.org/10.1109/ESEM56168.2023.10304801

[10] Kate Crawford. 2024. Generative AI Is Guzzling Water and Energy. Retrieved April 5, 2024 from https://www.nature.com/articles/d41586-024-00478-x

[11] Luís Cruz. 2022. The Five Dimensions of Sustainable Software Engineering and How Education Can Help! Retrieved from http://luiscruz.github.io/2022/01/01/sustainable-se-intro.html.https://doi.org/10.6084/m9.figshare.25036772

[12] Vlad-Andrei Cursaru, Laura Duits, Joel Milligan, Damla Ural, Berta Rodriguez Sanchez, Vincenzo Stoico, and Ivano Malavolta. 2024. A controlled experiment on the energy efficiency of the source code generated by code Llama. In *Proceedings of the International Conference on the Quality of Information and Communications Technology*. Springer, 161–176.

[13] Vincenzo De Martino, Silverio Martínez-Fernández, and Fabio Palomba. 2024. Do developers adopt green architectural tactics for ML-enabled systems? A mining software repository study. arXiv:2410.06708. Retrieved from https://arxiv.org/abs/2410.06708

[14] Alex de Vries. 2023. The growing energy footprint of artificial intelligence. *Joule* 7, 10 (2023), 2191–2194.

[15] Santiago del Rey, Silverio Martínez-Fernández, Luís Cruz, and Xavier Franch. 2023. Do DL models and training environments have an impact on energy consumption? In *Proceedings of the 49th Euromicro Conference on Software Engineering and Advanced Applications (SEAA '23)*. IEEE, 150–158.

[16] Santiago del Rey, Silverio Martínez-Fernández, and Xavier Franch. 2023. A review on green deployment for edge AI-abstract. In *ICT4S (Doctoral Symposium, Demos, Posters)*, 64–67.

[17] Pau Duran, Joel Castaño, Cristina Gómez, and Silverio Martínez-Fernández. 2024. GAISSALabel: A tool for energy labeling of ML models. arXiv:2401.17150. Retrieved from https://arxiv.org/abs/2401.17150

[18] Tore Dybå, Dag I. K. Sjøberg, and Daniela S. Cruzes. 2012. What works for whom, where, when, and why? On the role of context in empirical software engineering. In *Proceedings of the ACM-IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM '12)*. ACM, New York, NY, 19–28. DOI: https://doi.org/10.1145/2372251.2372256

[19] Delger Erdenesanaa. 2023. A.I. Could Soon Need as Much Electricity as an Entire Country. Retrieved April 5, 2024 from https://www.nytimes.com/2023/10/10/climate/ai-could-soon-need-as-much-electricity-as-an-entire-country.html

[20] Xavier Franch, Andreas Jedlitschka, and Silverio Martínez-Fernández. 2023. A requirements engineering perspective to AI-based systems development: A vision paper. In *Proceedings of the International Working Conference on Requirements Engineering: Foundation for Software Quality*. Springer, 223–232.

[21] Charlotte Freitag, Mike Berners-Lee, Kelly Widdicks, Bran Knowles, Gordon S. Blair, and Adrian Friday. 2021. The real climate and transformative impact of ICT: A critique of estimates, trends, and regulations. *Patterns* 2, 9 (2021). DOI: https://doi.org/10.1016/j.patter.2021.100340

[22] Achim Guldner, Rabea Bender, Coral Calero, Giovanni S. Fernando, Markus Funke, Jens Gröger, Lorenz M. Hilty, Julian Hörnschemeyer, Geerd-Dietger Hoffmann, Dennis Junger, et al. 2024. 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 (2024), 402–418.

[23] Mark Haakman, Luís Cruz, Hennie Huijgens, and Arie Van Deursen. 2021. AI lifecycle models need to be revised: An exploratory study in Fintech. *Empirical Software Engineering* 26, 5 (2021), 95.

[24] Khan Mohammad Habibullah, Gregory Gay, and Jennifer Horkoff. 2023. Non-functional requirements for machine learning: Understanding current use and challenges among practitioners. *Requirements Engineering* 28, 2 (2023), 283–316. DOI: https://doi.org/10.1007/S00766-022-00395-3

[25] Heli Järvenpää, Patricia Lago, Justus Bogner, Grace Lewis, Henry Muccini, and Ipek Ozkaya. 2023. A synthesis of green architectural tactics for ML-enabled systems. arXiv:2312.09610. Retrieved from https://arxiv.org/abs/2312.09610

[26] Mojan Javaheripi and Sébastien Bubeck. 2023. Phi-2: The surprising power of small language models. Retrieved April 5, 2024 from https://www.microsoft.com/en-us/research/blog/phi-2-the-surprising-power-of-small-language-models

[27] Eileen Kapel, Luís Cruz, Diomidis Spinellis, and Arie Van Deursen. 2023. Incident management in a software-defined business: A case study. *SSRN Electrical Journal*. DOI: https://doi.org/10.2139/ssrn.4333515

[28] Eileen Kapel, Luis Cruz, Diomidis Spinellis, and Arie Van Deursen. 2024. On the difficulty of identifying incident-inducing changes. In *Proceedings of the 46th International Conference on Software Engineering: Software Engineering in Practice*, 36–46.

[29] Miryung Kim, Thomas Zimmermann, Robert DeLine, and Andrew Begel. 2018. Data scientists in software teams: State of the art and challenges. *IEEE Transactions on Software Engineering* 44, 11 (2018), 1024–1038. DOI: https://doi.org/10.1109/TSE.2017.2754374

[30] Michael Kläs, Jens Heidrich, Jürgen Münch, and Adam Trendowicz. 2009. CQML scheme: A classification scheme for comprehensive quality model landscapes. In *Proceedings of the 35th Euromicro Conference on Software Engineering and Advanced Applications*. IEEE, 243–250.

[31] Dominik Kreuzberger, Niklas Kühl, and Sebastian Hirschl. 2023. Machine learning operations (MLOps): Overview, definition, and architecture. *IEEE Access* 11 (2023), 31866–31879.

[32] Yann LeCun, Corinna Cortes, and C. J. Burges. 2010. MNIST handwritten digit database. *ATT Labs [Online]*. Retrieved from http://yann.lecun.com/exdb/mnist 2.

[33] Alexandra Sasha Luccioni, Yacine Jernite, and Emma Strubell. 2023. Power hungry processing: Watts driving the cost of AI deployment? arXiv:2311.16863. Retrieved from https://arxiv.org/abs/2311.16863

[34] Alexandra Sasha Luccioni, Sylvain Viguier, and Anne-Laure Ligozat. 2023. Estimating the carbon footprint of bloom, a 176b parameter language model. *Journal of Machine Learning Research* 24, 253 (2023), 1–15.

[35] Lucy Ellen Lwakatare, Ivica Crnkovic, and Jan Bosch. 2020. DevOps for AI–Challenges in development of AI-enabled applications. In *Proceedings of the International Conference on Software, Telecommunications and Computer Networks (SoftCOM '20).* IEEE, 1–6.

[36] Silverio Martínez-Fernández. 2024. Environmental sustainability of machine learning systems: Reducing the carbon impact of their lifecycle process (keynote). In *Proceedings of the International Conference on Product-Focused Software Process Improvement (PROFES '24)*. Springer.

[37] Silverio Martínez-Fernández, Xavier Franch, and Francisco Durán. 2023. Towards green AI-based software systems: An architecture-centric approach (GAISSA). In *Proceedings of the 49th Euromicro Conference on Software Engineering and Advanced Applications (SEAA '23).* IEEE, 432–439.

[38] N. Nahar, H. Zhang, G. Lewis, S. Zhou, and C. Kastner. 2023. A meta-summary of challenges in building products with ML components—Collecting experiences from 4758+ practitioners. In *Proceedings of the IEEE/ACM 2nd International Conference on AI Engineering – Software Engineering for AI (CAIN '23)*. IEEE Computer Society, Los Alamitos, CA, 171–183. DOI: https://doi.org/10.1109/CAIN58948.2023.00034

[39] Phu H Nguyen, Sagar Sen, Beatriz Bretones-Cassoli, Nicolas Jourdan, and Maria Chiara Magnanini. 2023. Software engineering and AI for data quality in cyber-physical systems/Internet of things-SEA4DQ'22 report. *ACM SIGSOFT Software Engineering Notes* 48, 1 (2023), 108–111.

[40] Nienke Nijkamp, June Sallou, Niels van der Heijden, and Luís Cruz. 2024. Green AI in action: Strategic model selection for ensembles in production. In *Proceedings of the 1st ACM International Conference on AI-Powered Software*, 50–58.

[41] Emmanouil Antonios Platanios, Otilia Stretcu, Graham Neubig, Barnabas Poczos, and Tom M. Mitchell. 2019. Competence-based curriculum learning for neural machine translation. arXiv:1903.09848. Retrieved from https://arxiv.org/abs/1903.09848

[42] Lorena Poenaru-Olaru, Luis Cruz, Jan S. Rellermeyer, and Arie Van Deursen. 2023. Maintaining and monitoring AIOps models against concept drift. In *Proceedings of the IEEE/ACM 2nd International Conference on AI Engineering–Software Engineering for AI (CAIN '23).* IEEE, 98–99.

[43] Lorena Poenaru-Olaru, Natalia Karpova, Luis Cruz, Jan Rellermeyer, and Arie van Deursen. 2024. Maintenance techniques for anomaly detection AIOps solutions. In *Proceedings of the IEEE/ACM 2nd International Conference on AI Engineering–Software Engineering for AI (CAIN '24).*

[44] Lorena Poenaru-Olaru, June Sallou, Luis Cruz, Jan S Rellermeyer, and Arie Van Deursen. 2023b. Retrain AI systems responsibly! Use sustainable concept drift adaptation techniques. In *2023 IEEE/ACM 7th International Workshop on Green And Sustainable Software (GREENS).* IEEE, 17–18.

[45] PyTorch. 2024. Quantization—PyTorch 2.2. Retrieved April 5, 2024 from https://pytorch.org/docs/stable/quantization.html

[46] Enrique Barba Roque, Luis Cruz, and Thomas Durieux. 2025. Unveiling the energy vampires: A methodology for debugging software energy consumption. In *Proceedings of the 47th International Conference on Software Engineering (ICSE '25).*

[47] Dhia Elhaq Rzig, Foyzul Hassan, and Marouane Kessentini. 2022. An empirical study on ML DevOps adoption trends, efforts, and benefits analysis. *Information and Software Technology* 152 (2022), 107037.

[48] S. Salehi and A. Schmeink. Data-centric green artificial intelligence: A survey. *IEEE Transactions on Artificial Intelligence* 5, 5 (Sep. 4), 1973–1989. DOI: https://doi.org/10.1109/TAI.2023.3315272

[49] Danilo Sato, Arif Wider, and Christoph Windheuser. 2019. Continuous delivery for machine learning. Retrieved April 5, 2024 from https://martinfowler.com/articles/cd4ml.html

[50] Jieke Shi, Zhou Yang, and David Lo. 2024. Efficient and green large language models for software engineering: Vision and the road ahead. arXiv:2404.04566. Retrieved from https://arxiv.org/abs/2404.04566

[51] Arumoy Shome, Luis Cruz, and Arie Van Deursen. 2022. Data smells in public datasets. In *Proceedings of the 1st International Conference on AI Engineering: Software Engineering for AI*, 205–216.

[52] Arumoy Shome, Luis Cruz, and Arie van Deursen. 2024. Data Vs. model machine learning fairness testing: An empirical study. In *Proceedings of the 5th IEEE/ACM International Workshop on Deep Learning for Testing and Testing for Deep Learning (DeepTest '24)*, 1–8.

[53] Forrest Shull, Jeffrey C. Carver, Sira Vegas, and Natalia Juristo Juzgado. 2008. The role of replications in empirical software engineering. *Empirical Software Engineering* 13, 2 (2008), 211–218. DOI: https://doi.org/10.1007/S10664-008-9060-1

[54] Julien Siebert, Lisa Joeckel, Jens Heidrich, Koji Nakamichi, Kyoko Ohashi, Isao Namba, Rieko Yamamoto, and Mikio Aoyama. 2020. Towards guidelines for assessing qualities of machine learning systems. In *Quality of Information*

and Communications Technology. Martin Shepperd, Fernando Brito e Abreu, Alberto Rodrigues da Silva, and Ricardo Pérez-Castillo (Eds.), Springer International Publishing, Cham, 17–31.

[55] Dag I. K. Sjøberg, Tore Dybå, Bente C. D. Anda, and Jo E. Hannay. 2008. *Building Theories in Software Engineering*. Springer London, 312–336. DOI: https://doi.org/10.1007/978-1-84800-044-5_12

[56] Klaas-Jan Stol and Brian Fitzgerald. 2013. Uncovering theories in software engineering. In *Proceedings of the 2nd SEMAT Workshop on a General Theory of Software Engineering (GTSE '13)*, 5–14. DOI: https://doi.org/10.1109/GTSE.2013.6613863

[57] TensorFlow. 2024. TensorFlow model optimization. Retrieved April 5, 2024 from https://www.tensorflow.org/model_optimization?hl=en

[58] Roberto Verdecchia, Luís Cruz, June Sallou, Michelle Lin, James Wickenden, and Estelle Hotellier. 2022. Data-centric Green AI an exploratory empirical study. In *2022 International Conference on ICT for Sustainability (ICT4S)*. IEEE, 35–45.

[59] Roberto Verdecchia, June Sallou, and Luis Cruz. 2023. A systematic review of Green AI. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 13, 4 (2023), e1507. DOI: https://doi.org/10.1002/WIDM.1507

[60] Andreas Vogelsang and Markus Borg. 2019. Requirements engineering for machine learning: Perspectives from data scientists. In *Proceedings of the IEEE 27th International Requirements Engineering Conference Workshops (REW '19)*, 245–251. DOI: https://doi.org/10.1109/REW.2019.00050

[61] Paul Walsh, Jhilam Bera, Vibhu Saujanya Sharma, Vikrant Kaulgud, Raghotham M. Rao, and Orlaith Ross. 2021. Sustainable AI in the cloud: Exploring machine learning energy use in the cloud. In *Proceedings of the 36th IEEE/ACM International Conference on Automated Software Engineering Workshops (ASEW '21)*. IEEE, 265–266.

[62] Tongzhou Wang, Jun-Yan Zhu, Antonio Torralba, and Alexei A. Efros. 2018. Dataset distillation. arXiv:1811.10959. Retrieved from https://arxiv.org/abs/1811.10959

[63] Carole-Jean Wu, Ramya Raghavendra, Udit Gupta, Bilge Acun, Newsha Ardalani, Kiwan Maeng, Gloria Chang, Fiona Aga Behram, James Huang, Charles Bai, et al. 2021. Sustainable AI: Environmental implications, challenges and opportunities. arXiv:2111.00364. Retrieved from https://arxiv.org/abs/2111.00364

[64] Tim Yarally, Luís Cruz, Daniel Feitosa, June Sallou, and Arie van Deursen. 2023. Batching for Green AI—An exploratory study on inference. In *Proceedings of the 49th Euromicro Conference on Software Engineering and Advanced Applications (SEAA '23)*. IEEE, 112–119.

[65] Tim Yarally, Luís Cruz, Daniel Feitosa, June Sallou, and Arie Van Deursen. 2023. Uncovering energy-efficient practices in deep learning training: Preliminary steps towards Green AI. In *Proceedings of the IEEE/ACM 2nd International Conference on AI Engineering–Software Engineering for AI (CAIN '23)*. IEEE, 25–36.

Received 5 April 2024; revised 18 October 2024; accepted 5 November 2024