Invited: Toward an ML EDA Commons: Establishing Standards, Accessibility, and Reproducibility in ML-driven EDA Research

Vidya A. Chhabria
Arizona State University
Tempe, AZ, USA
vachhabr@asu.edu

Jiang Hu
Texas A&M University
College Station, TX, USA
jianghu@tamu.edu

Andrew B. Kahng
UC San Diego
La Jolla, CA, USA
abk@ucsd.edu

Sachin S. Sapatnekar
University of Minnesota
Minneapolis, MN, USA
sachin@umn.edu

Abstract

Machine learning (ML) is transforming electronic design automation (EDA), offering innovative solutions for designing and optimizing integrated circuits (ICs). However, the field faces significant challenges in standardization, accessibility, and reproducibility, limiting the impact of ML-driven EDA (ML EDA) research. To address these barriers, this paper presents a vision for an ML EDA Commons, a collaborative open ecosystem designed to unify the community and drive progress through establishing standards, shared resources, and stakeholder-based governance. The ML EDA Commons focuses on three objectives: (1) Maturing existing EDA infrastructure to support ML EDA research; (2) Establishing standards for benchmarks, metrics, and data quality and formats for consistent evaluation via governance that includes key stakeholders; and (3) Improving accessibility and reproducibility by providing open datasets, tools, models, and workflows with cloud computing resources, to lower barriers to ML EDA research and promote robust research practices via artifact evaluations, canonical evaluators, and integration pipelines. Inspired by successes of ML and MLCommons, the ML EDA Commons aims to catalyze transparency and sustainability in ML EDA research.

CCS Concepts

• **Hardware → Physical design (EDA)**; **Methodologies for EDA**.

Keywords

VLSI CAD, open-source tools and flows, datasets, machine learning

ACM Reference Format:
Vidya A. Chhabria, Jiang Hu, Andrew B. Kahng, and Sachin S. Sapatnekar. 2025. Invited: Toward an ML EDA Commons: Establishing Standards, Accessibility, and Reproducibility in ML-driven EDA Research. In *Proceedings of the 2025 International Symposium on Physical Design (ISPD '25), March 16–19, 2025, Austin, TX, USA.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3698364.3709131

1 Introduction

The rapid advancement of machine learning (ML) and artificial intelligence (AI) technologies, including the recent advances in large language models (LLMs), has unlocked opportunities for electronic design automation (EDA), enabling transformative shifts in the design, optimization, and verification of integrated circuits (ICs) and systems. ML techniques have shown promise in addressing complex problems across the RTL-to-GDSII design flow, including physical design problems such as layout optimization [1, 2], performance prediction [3, 4], and enabling automation [5, 6]. However, the successful integration of ML into the design flow requires a robust, open, and shareable ecosystem.

ML EDA Challenges. Despite increasing interest over the past decade, ML-driven EDA (ML EDA) faces significant and well-lamented challenges due to a disconnect between the open culture of ML research and the closed nature of the EDA community. ML research thrives on open access to datasets, benchmarks, metrics, and models. By contrast, EDA – particularly the RTL-to-GDSII flow – relies on components such as RTL, process design kits (PDKs), libraries, tools, and flows, all constrained by intellectual property (IP) restrictions. Non-disclosure agreements (NDAs) limit access to PDKs, end-user license agreements (EULAs) restrict interoperability and sharing of proprietary tools and flows, and unique tool-specific parameters further complicate reproducibility. RTL designs are often proprietary, hindering availability of reliable benchmarks; those available are often outdated or poorly maintained. The lack of cohesion within the EDA community compounds these issues, as fragmented efforts and the absence of a governing body prevent the development of standards, open formats, and equitable access. Without coordinated leadership to foster collaboration, establish benchmarks, and align initiatives, the field continues to struggle with reproducibility and transparency of innovation.

Success of ML. Donoho [7] defines *frictionless reproducibility* as arising from three foundations: (1) availability and sharing of data; (2) availability and sharing of code that processes this data; and (3) competitive testing as a means of evaluation. The success of ML can largely be attributed to these principles, which have fostered transparency, accessibility, and rigorous validation, enabling rapid progress. Open datasets and benchmarks such as ImageNet [8] accelerated early ML research by enabling consistent evaluation and fostering innovation through challenges such as Kaggle [9]. Standardized libraries such as TensorFlow, PyTorch, and scikit-learn further reduced barriers, enabling rapid experimentation and deployment. Similarly, MLCommons [10], an AI consortium, advances AI systems through open collaboration, standardized benchmarks such as MLPerf [11], and metrics (accuracy, efficiency, and energy performance), exemplifying the impact of a *Commons*.

Need for an ML EDA Commons. The need for a shareable ML EDA research infrastructure has long been recognized [12]. The work in [13] demonstrated the potential of ML in physical design (PD), and outlined several challenges related to infrastructure for

ACM ISBN 979-8-4007-1293-7/25/03
https://doi.org/10.1145/3698364.3709131

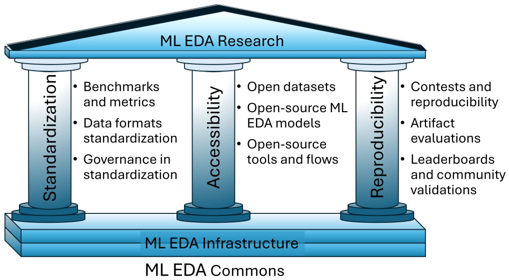

Figure 1: The three pillars of an ML EDA Commons.

ML for PD that are still unaddressed today. Realizing the infrastructure challenge, our community has started moving toward an open culture with the availability of open-source tools [14], PDKs [15–17], researchers open-sourcing their models [5, 18], datasets [19–22], and scripts. While these initiatives contribute to ML EDA infrastructure, they are scattered or fragmented due to IP challenges, and often inadequate absent buy-in from key stakeholders. While IP-related reasons create one category of fragmentation, geographical reasons are another. Research and development efforts are unevenly distributed across regions due to disparities in access to funding, geopolitical considerations, infrastructure, and educational resources. These differences often lead to isolated advancements, which fail to integrate into a cohesive global framework. These challenges result in a highly siloed ecosystem, with research efforts dispersed across academia and industry. Such fragmentation creates significant barriers for new entrants, along with reproducibility challenges.

A viable solution requires a large-scale and well-concerted community effort. ML EDA, with many fewer actors, cannot afford redundant efforts. The community must coordinate to identify and fill gaps in existing infrastructure, generate the most critical content, and shape external contributions through contests and community engagement. Inspired by the MLCommons, there is a need for an ML EDA Commons to mature existing infrastructure, standardize formats, establish quality standards for datasets, benchmarks, metrics, leaderboards, etc., and generally advance the accessibility and reproducibility of ML EDA research.

Toward establishing an ML EDA Commons, this paper presents four perspectives. (1) It calls for the establishment of an ML EDA Commons to build ML EDA infrastructure and propel ML EDA research for the RTL-to-GDSII flow. (2) It highlights existing initiatives that can function as foundational ML EDA infrastructure. (3) It identifies critical gaps in ML EDA infrastructure efforts that the ML EDA Commons must address. (4) It outlines the essential components of an ML EDA Commons, including maturing existing infrastructure, governance, standardization, accessibility, and reproducibility.

The rest of the paper is organized as follows. Section 2 defines what an ML EDA Commons is within the context of physical design. It explains key terminologies, including the definition of a Commons; the ML EDA infrastructure; and how these components are related. Section 3 reviews current contributions to building an ML EDA infrastructure, which forms part of the ML EDA Commons. It also examines what has been accomplished so far, how existing efforts fall short, and identifies gaps that need to be addressed. Section 4 details how an ML EDA Commons can be established. It specifies the required components, what enhancements are needed to improve current infrastructure, who would be responsible for development and maintenance, and who are the key stakeholders. It discusses the role of the Commons in defining standards and policies, and establishing governance that emphasizes accessibility and reproducibility of ML EDA research. Section 5 outlines a roadmap for proposed action items toward achieving this vision.

2 A Vision of an ML EDA Commons

Commons. The concept of "Commons" refers to shared resources where all stakeholders have an equal interest [23]. Tumeo [24] describes the *Open Hardware Technology Commons* (OHTC) as an "open and extensible portfolio of composable and interoperable hardware, software, design automation, and architecture design tools" that facilitates rapid prototyping. Kahng [25] provides the vision of an *EDA Commons* and highlights the complementarity of open-source and proprietary EDA technologies as an integral part of the OHTC. An *ML EDA Commons*, a critical piece of the EDA Commons, focuses on advancing ML EDA applications by providing shared resources (datasets, tools, flows, benchmarks, metrics). This forms a part of ML EDA infrastructure prioritizing the principles of standardization, accessibility, and reproducibility (Fig. 1).

Components of an ML EDA Commons. An ML EDA Commons for the RTL-to-GDSII flow consists of the following (Fig. 2).

(1) ML EDA infrastructure is composed of the essential tools, resources, frameworks, and standards required to enable ML EDA. It encompasses datasets, benchmarks, APIs, open-source tools, and proxies to facilitate model development, training, evaluation, and deployment across the RTL-to-GDSII design flow. These are highlighted in Fig. 2(b). Today, this infrastructure is nascent and requires significant effort to form a mature pool of resources.

(2) Standardization refers to the establishment of common protocols, formats, benchmarks, and best practices to ensure consistency, interoperability, and collaboration across tools, datasets, and research workflows. These standards are set by key stakeholders, and involve defining uniform data structures, evaluation metrics, and APIs to streamline the integration of ML techniques into EDA flows.

(3) Accessibility refers to the ability of researchers, developers, and practitioners to access and utilize resources, tools, and data for ML EDA research. This includes open and equitable access to datasets, benchmarks, models, and tools without prohibitive barriers (cost, intellectual property restrictions, technical complexity). Accessibility also involves creating user-friendly interfaces, APIs, and documentation to lower the entry barrier for newcomers and foster collaboration across academia, industry, and open-source communities [26].

(4) Reproducibility refers to the ability to consistently replicate experiments, results, and workflows across different environments, researchers, and institutions using shared resources. It ensures that models, algorithms, and research findings in ML EDA can be independently verified and validated by others. This involves providing well-documented datasets, standard benchmarks and leaderboards, publicly available code, and clear instructions for reproducing experiments. Reproducibility fosters trust, transparency, and collaboration in the ML EDA community, enabling researchers to build upon each other's work while reliably and credibly advancing the field.

The pillars of the ML EDA Commons, standards, accessibility, and reproducibility, align closely with the principles of FAIR (findable, accessible, interoperable, and reproducible). The ML EDA Commons

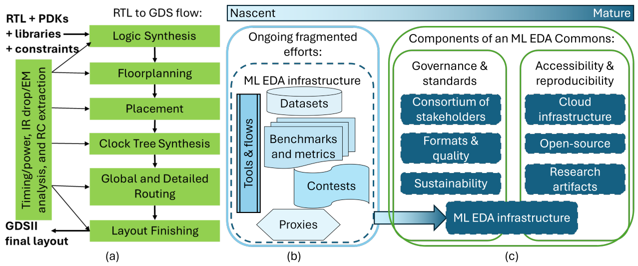

Figure 2: (a) The RTL-to-GDSII flow. (b) Existing individual efforts in creating ML EDA infrastructure at different levels of maturity. (c) Envisioned components of an ML EDA Commons.

must ensure each component developed, released and contributed to the community as a part of the ML EDA infrastructure aligns with FAIR principles. Stable repositories or platforms for hosting the infrastructure make it findable and accessible. Accessibility emphasizes open and transparent data sharing, supported by proper licensing and clear documentation. Standards ensure that resources contributed to the infrastructure are consistently formatted, annotated, and validated, make them interoperable. across diverse ML EDA applications and tools. Documentation, continuous integration/continuous delivery (CI/CD) pipelines, and version control, enables others to reproduce or extend previous works using the same datasets and methodologies.

Scope of the ML EDA Commons. The Commons, while focused on enabling ML EDA research, can stir a sense of excitement in students and can contribute shareable, modular training materials to engage and train the next generation of ML EDA researchers and engineers. However, it is important to acknowledge that an ML EDA Commons cannot realistically aim to address every stakeholder's needs comprehensively. For instance, access to tape-out shuttles for "lab-to-fab" prototyping, chip design workforce development resources such coursework and laboratories, IP access, licensing of EDA tools, etc. are perhaps better served by other initiatives such as Microelectronics Commons Hubs [27], NSF Chipshub [28, 29], Natcast WFPA [30], and NIST NAPMP efforts [31]. To be effective, the ML EDA Commons must focus on identifying the most pressing problems and prioritizing solutions that offer the highest impact for ML EDA research.

3 ML EDA infrastructure initiatives

Developing a robust ML EDA research infrastructure for the RTL-to-GDSII flow requires key components (Fig. 2(b)) such as datasets, benchmarks, metrics, contests, and proxies. Existing efforts to create this infrastructure remain nascent and individually-driven, and do not form a cohesive ecosystem. Maturity levels of each component vary, as shown in Fig. 2(b). Below, we outline example initiatives toward building each component, along with limitations.

3.1 Datasets

Datasets from domains where ML has traditionally excelled – text, images, code, audio, and video – differ significantly from those encountered in the RTL-to-GDSII design flow [32]. IC data comes in specialized formats such as hardware description languages (e.g., SystemVerilog), reports, and EDA file formats (e.g., LEF/DEF), which are less intuitive or widely understood than natural language or images. To make this data suitable for ML models, it is typically transformed into structured formats that ML excels at processing.

Examples. Numerous efforts have sought to gather and release datasets across different flow stages.

RTL and logic synthesis. The creation of RTL datasets has seen various individual efforts, such as VeriGen [33] and RTLCoder [18]. VeriGen [33] has gathered Verilog datasets compiled from various GitHub repositories and Verilog textbooks. RTLCoder [18] created an automated training dataset generation flow, producing over 27,000 training samples. Each sample includes a design description instruction and the corresponding reference RTL code. Another effort was the community-driven LLM4HW [22] contest held at ICCAD 2024, aiming to gather an open-source, large-scale, high-quality dataset to fine-tune LLMs for generating Verilog code. For logic synthesis, OpenABC-D [19] is a labeled dataset with 29 hardware IP designs synthesized with 1500 yosys-abc synthesis flows, with 870,000 and-inverter graphs (AIGs) in PyTorch format.

Physical design. The high cost of creating place-and-route datasets results in limited availability of such benchmarks. Examples include FloorSet [34], CircuitNet [21], BeGAN [35], and EDA Corpus [36]. FloorSet [34] contains two benchmark datasets, FloorSet-Prime and FloorSet-Lite, each with 1M training samples and 100 test samples of synthetic fixed-outline floorplan layouts, reflecting real SoC distributions and modern design flow constraints. CircuitNet [21] features samples from 10,242 layouts generated through synthesis and physical design of 28nm RISC-V designs.

BeGAN [35] is a synthetic dataset for power delivery networks (PDNs), created using generative adversarial networks (GANs) and transfer learning from urban satellite images. It contains realistic PDN benchmarks with 1000 data points per technology node for three open-source technologies, preserving data privacy of the underlying IP used for training. EDA Corpus [36] is an open-source dataset for the OpenROAD EDA toolchain, with 1000 data points that include pairwise sets of question prompts with prose answers and code prompts with corresponding OpenROAD scripts, facilitating LLM-focused research on user productivity for EDA flows.

Maturity level. The released ML EDA datasets are typically small, lack realism, and are inadequate for training today's large-scale ML models with billions of parameters (e.g., LLMs). These datasets are often developed in isolation, lack standardization, are not significantly diverse, and fail to cater to the diverse applications required for ML EDA research across the entire RTL-to-GDSII flow. Researchers have created datasets that are publicly available, but these are still in their infancy and require significant advancements in quality, quantity, and standards to make a meaningful impact. Datasets must be continuously updated to remain relevant, which becomes challenging when they are created by students for specific research papers; maintaining and updating these datasets after students graduate is difficult due to a lack of incentives and continuity.

3.2 Contests, benchmarks, and metrics

There have been many contests in EDA over the past several decades, including ones that create benchmarks and metrics for ML EDA research. These contests have been crucial to the advancement of EDA research. Similar to the ML community, these contests have the potential to transform ML EDA research. Below, we highlight examples of the contests, describe their benchmarks and metrics, and then comment on their level of maturity in the context of an ML EDA Commons.

Examples. Contests in ML EDA have recently gained prominence for their role in setting benchmarks and driving innovation. ICCAD 2023 featured a problem on static IR drop estimation [37], and ICCAD 2024 presented a problem on gate sizing [38]. The ISPD 2024 and 2025 contests focused on global routing [39]. Each of these contests has released benchmarks tailored to their respective problems.

While contests are a good source of benchmarks, there have also been significant efforts outside of contests to create useful benchmarks for EDA research. VerilogEval [40] serves as a benchmark for evaluating LLM performance in Verilog code generation, using a dataset of 156 problems from HDLBits. It tests a range of tasks from simple circuits to complex finite-state machines, providing a standardized way to assess the functional correctness of generated Verilog code. Efforts to confirm reproducibility have also led to new benchmarks such as those in the MacroPlacement [41]. This repository consists of open-source designs such as Ariane, MemPool, and NVDLA, in enablements such as NanGate45, ASAP7, and SKY130HD; these serve as benchmarks for the macro placement problem.

Maturity level. Many current contests are based on oversimplified problems and unrealistic assumptions. The research that uses these benchmarks may have limited practical impact, due to these unrealistic assumptions. The benchmarks they provide fail to address the complexities of real-world challenges. Moreover, each contest introduces its own benchmarks and nonstandard metrics, leading to further fragmentation. The benchmarks almost always fall victim to Goodhart's law [42] ("when a measure becomes a target, it ceases to be a good measure") and need to be regularly updated.

3.3 Open-source ML EDA models

The ML community has benefited from large open-source foundation models such as Llama and, more recently, DeepSeek [43]. These models are trained on large amounts of data and generalize to different modalities. The ML EDA community has extensively leveraged these large models, e.g., fine-tuning them for domain-specific applications.

Examples. Open domain-specific models include [18] for RTL generation and [5] for physical design (answering questions, scripting).

Maturity level. The existing open-source models in ML EDA are usually task-specific and unsuitable for developing broader or general-purpose solutions. These models are still at a very early stage of development and are often released as an artifact of a research paper. Typically, they require peripheral harnesses to enable reuse, and are very far from open-source foundation models for EDA that can be generalized across different modalities and applications.

3.4 Open-source EDA tools and flows

The open-source EDA community serves as a cornerstone of the ML EDA infrastructure by enabling scalable data generation without IP restrictions, facilitating the integration of ML into design flows, and embedding ML capabilities within EDA tools.

Examples. Open-source EDA tools such as OpenROAD and Yosys, along with flows such as OpenROAD-flow-scripts [44], have been available for some time. These tools and flows are actively updated and maintained, have garnered a large user base, and have been successfully used to tape out multiple chips.

Maturity level. The open-source EDA ecosystem [14, 44, 45] has reached sufficient maturity to play a vital role in advancing ML EDA research by making it accessible and reducing entry barriers. While some prior efforts have enabled an ML EDA research playground [46], further development is needed to fully meet community requirements, including support for ML-native data structures within tools, Python APIs that provide ML-friendly outputs (e.g., NumPy arrays), and other features tailored to ML researchers.

3.5 Proxies for designs and PDKs

Proxies serve as substitutes for commercial or proprietary resources when direct access or sharing is restricted. These include proxy PDKs that emulate proprietary process design kits, proxy designs that replicate the characteristics of commercial designs, and proxy tools, such as open-source alternatives, that provide functional stand-ins for proprietary EDA tools. Although proxies may not achieve the exact performance, power, and area metrics of their commercial equivalents, they are sufficiently accurate to facilitate research, development, and benchmarking in resource-constrained scenarios.

Examples. Efforts include proxies for PDKs [47, 48], RTL [49], and netlists [50]. To bridge the gap between open-source and commercial PDKs, [48] calibrated proxy design enablements which autotuned scaling factors for standard-cell timing and power models, setup and hold timing, pin capacitance, and BEOL resistance and capacitance. This work has narrowed the gap between the open-source ASAP7 enablement and an unnamed leading-edge commercial 7nm technology. Other proxies include [49] and [50] which respectively create synthetic RTL and gate-level netlists for ML applications.

Maturity level. Initial efforts to create proxies for commercial designs and PDKs are primarily driven by very few individual initiatives. These proxies still need adoption and recognition; efforts to drive awareness are still needed to gain traction. Their value is yet to be recognized in the community (including reviewers, conferences, and for ML EDA) and they must further be developed to make meaningful impacts in a data- and IP-constrained world.

4 Components of the ML EDA Commons

We now outline the components of the ML EDA Commons, including a mature ML EDA infrastructure and a governance framework to establish standardization for accessibility and reproducibility.

4.1 Maturing the ML EDA infrastructure

The several individual efforts described in Section 3 establish various different components of an ML EDA infrastructure. These efforts must be matured, maintained, and unified into a shared platform to have an impact. The acute need for a shared open collaborative ML EDA infrastructure was echoed in a 2023 NSF workshop [12]. Such an ML EDA infrastructure will be an important component of the ML EDA Commons as a pool of resources.

SLICE. The 2023 NSF workshop on Shared Infrastructure for Machine Learning EDA [12] culminated in SLICE [51], a Shared machine Learning Infrastructure for the Community of EDA that will potentially play a crucial role in the Commons. Inspired by the open shared infrastructure available in the ML community, SLICE aims to create an analogous infrastructure for ML EDA as shown in Fig. 3, with all the components described in Section 3.

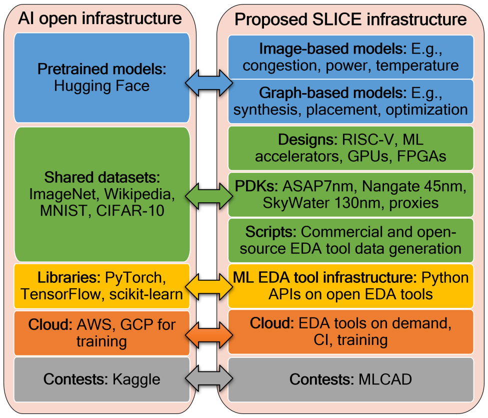

Figure 3: Inspired by open AI infrastructure, SLICE [51] aims to develop an ML EDA infrastructure.

Aligned with the broader goal of the Commons to unify ML EDA infrastructure efforts, SLICE aims to serve as a one-stop shop for data, models, scripts, and software interfaces with EDA tools for easy access by the EDA community [51]. SLICE was conceived based on inputs from a large community of researchers spanning academia, industry, and government with a focus on digital design and verification. SLICE seeks to advance existing ML EDA infrastructure through the following contributions.

- *Datasets and benchmarks.* SLICE aims to create datasets in two ways. The first way curates EDA tool runs across different flow stages with different parameter settings; the second way leverages synthetic data generation techniques [35] to create large datasets for training large models. SLICE will develop proxies for designs and PDKs, while also updating and maintaining benchmarks for a variety of ML EDA applications across the RTL-to-GDSII flow. Beyond the datasets from SLICE and other sources, *a mature infrastructure as a part of the Commons will gather datasets into a single platform. The Commons will enforce standards for quality, documentation, and diversity.*
- *Pretrained models.* SLICE will develop and open pre-trained models for a variety of applications including automatic HDL code generation, testbench generation, RTL code bug detection, EDA tool script generation [5], etc. These will be released along with the datasets used to train them, enabling reduced training costs for users and development of foundation models that can be fine-tuned for specific ML EDA applications. The ML EDA Commons must enforce that models are *supported by collaterals to ensure usability, including scripts for processing data that feeds into the models, and APIs for model training and fine-tuning.*
- *Software infrastructure.* SLICE will enhance existing software infrastructure [46] to support ML inside open-source EDA tools, and outside EDA tools within open EDA flows. SLICE will develop Python APIs that interact with EDA tools to directly support ML-native data return types in PyTorch or NumPy arrays.

4.2 Governance in ML EDA Commons

All of the above-mentioned infrastructure pieces will serve as a pool of resources managed by the ML EDA Commons. *Commons infrastructure and its development also require governance to establish quality standards, equitable access, and prioritization of reproducibility.*

Table 1: Roles of stakeholders in activities of the Commons.

Key stakeholders	Define standards	Develop infrastructure	Provide incentives	Serve on advisory board	Ensure sustenance	Engage community
Academia	✓	✓		✓	✓	✓
Industry	✓	✓	✓	✓	✓	✓
Professional societies	✓				✓	✓
Government agencies	✓		✓	✓		
Open-source community	✓	✓			✓	

4.2.1 *Commons activities.*
Key Commons activities include:

- *Defining standards and policies.* Standards and policies should define clear and comprehensive guidelines for contributing datasets and benchmarks, developing formats, and developing guidelines for reproducibility and artifact evaluations (AE). These standards should incorporate quality assurance measures, establish protocols to mitigate data poisoning risks, and mandate the use of CI/CD pipelines for managing datasets. Additionally, they should include maintenance policies to ensure long-term usability of datasets and benchmarks. Furthermore, the guidelines should provide specific recommendations for conference organizers, reviewers, and journal editors to promote reproducibility and strongly encourage the availability of necessary artifacts.
- *Providing incentives and funding.* Incentives and funding mechanisms are essential to encourage contributions to ML EDA Commons infrastructure. The Commons can offer incentives to its contributors, e.g., certifications and support in the pursuit of funding opportunities. Incentives such as badges provided by the Commons can encourage reproducible ML EDA research.
- *Contributing to ML EDA infrastructure and resource allocation.* Contributing to ML EDA infrastructure includes the community-driven engineering effort in releasing datasets, and developing supporting libraries and tools that follow guidelines set by defined policies and standards. Resource allocation and management activities in the Commons must ensure availability of its resources (cloud and infrastructure) to all users, while managing access to IP, tools, and PDKs.
- *Providing direction through an advisory board.* An advisory board to the ML EDA Commons must guide the development and periodic updates of standards to keep pace with the rapid advancements in ML research. The board should also identify key areas requiring investment based on the progress and evolving needs of ML EDA research. The board also serves as a channel through which community needs are communicated and considered.
- *Sustaining the effort over time.* These efforts involve developing robust frameworks to maintain standards and infrastructure within the Commons. This includes regularly updating benchmarks to mitigate the effects of Goodhart's Law, ensuring that they reflect real-world complexities and do not encourage narrow optimization. Additionally, these efforts must create strategies to ensure the long-term viability of the ML EDA Commons, such as fostering community engagement, incentivizing contributions, and securing funding to support ongoing development and innovation.
- *Promoting community engagement.* Community engagement activities span the roles of stakeholders in conference/workshop organization, mechanisms for conflict resolution within the community to prevent fragmentation, and strategies to ensure that the ML EDA Commons can sustain itself over the long term.

4.2.2 Role of stakeholders. The stakeholders in the ML EDA Commons will form an essential part of the governance. Their roles, shown in Table 1, are outlined below:

- *Academic members.* Academia contributes by developing prototype frameworks and publishing research that can inform the creation of ML EDA standards. Academics also implement and contribute datasets, and advise on standards for the next generation of ML EDA algorithms. They play a key role in steering conferences and workshops, which indirectly drive research, and can advise on areas for investment within the Commons.
- *Industry members.* Members from industry play a role in all activities, as shown in Table 1. They guide development of standards based on their required ML EDA applications (an example includes the Si2 efforts [52]); contribute toward developing datasets [34], benchmarks [40], designs and open-source PDKs, and provide compute resources to the community. They also can incentivize contributions to the Commons via funding opportunities. Industry ensures sustainability by integrating ML EDA outcomes into commercial workflows and supporting community projects.
- *Professional societies.* Professional organizations, including IEEE and ACM, play a vital role as they sponsor conferences, oversee publications, and promote reproducibility through badges and awards, such as best artifact and best paper recognitions. As newer conferences and workshops emerge to focus on ML EDA [53, 54], these can actively prioritize reproducible research, encouraging the use of open-source tools, datasets, benchmarks, and metrics by establishing practices such as AE as incentives.
- *Government agencies.* Government agencies shape policies to align ML EDA advancements with national priorities such as technological leadership. They provide funding, guide roadmaps, and identify investment areas through advisory roles, exemplified by workshops on ML EDA infrastructure [12, 55].
- *Open-source communities.* The community contributes to creating open standards, ensuring interoperability and accessibility of ML EDA tools. Open-source developers build and maintain tools, frameworks, and libraries essential for implementation of ML EDA systems, and ensure continuous improvement of tools and resources to ensure the long-term viability of the Commons.

4.3 ML EDA data standardization

The ML EDA Commons must establish standards for research, including data formats and quality standards for benchmarks, datasets, and metrics. These standards can be established if the ML EDA Commons defines them through responsible governance by key stakeholders.

4.3.1 Standards for data formats. Users today face significant challenges in exploring ML with proprietary EDA tools due to several barriers: (i) the lack of standardized reporting formats and tool metrics; (ii) the unique and proprietary parameters of each tool, which complicate comparisons and adjustments; and (iii) the extensive fragmentation of naming conventions and formats across toolchains and vendors, creating a "Tower of Babel" that impedes the sharing of ML models and expertise. Moreover, proprietary tool command sets and reports are often copyrighted and confidential, restricting the deployment of ML EDA. These challenges are well recognized, and efforts to address these obstacles and overcome the "Tower of Babel" date back to the late 1990s. The work [56] proposed METRICS1.0, which was then realized in [57]. With the rise of AI/ML, renewed efforts have emerged to address these challenges.

(1) *METRICS2.1.* A collaborative effort between the OpenROAD project and IEEE CEDA DATC, METRICS2.1[58] introduces an open-source standard for metrics naming conventions, a metrics dictionary, and a reference JSON-based implementation within the OpenROAD/OpenROAD-flow-scripts platforms. Thousands of RTL-to-GDSII datasets, with configuration files for reproducibility, are available in DATC's GitHub [59].
(2) *CircuitOps.* A data representation format developed for integrating ML in EDA, CircuitOps [60] minimizes preprocessing overhead for ML EDA, lowering barriers to entry for researchers in applying ML models for tasks such as timing optimization, power estimation, and design space exploration. The design data is represented as labeled property graphs (LPGs) backed by intermediate representation tables (IR Tables), simplifying the process of custom dataset generation for ML applications.
(3) *EDA Schema.* A recent effort toward creating a schema to drive ML EDA applications is EDA schema [61]. It uses a property graph that represents the physical attributes and quality-of-results (QoR) metrics of a circuit across various stages of the physical design flow. The property graph encapsulates information about timing paths, interconnects, and parasitics.

Each of the above data formats has its own limitations and often fails to cater to a broad spectrum of ML use cases of the industry. To address this, the Silicon Integration Initiative (Si2) – a non-profit organization that creates standards for the semiconductor industry – is collaborating with its members, Drexel University, and Arizona State University to develop a unified ML EDA schema [52]. This initiative aims to serve the ML EDA data format needs of member companies and their ML EDA applications. Recent advancements in this effort include enhancing CircuitOps and transitioning it from a data format to a formal schema. These enhancements include documented CircuitOps APIs that call PyTorch, Pandas, or Graph Tool library APIs under the hood, as well as example CircuitOps use cases and scripts [62] designed to interact with OpenROAD via Python APIs [46]. *Such efforts in developing standard formats must be governed and overseen by the ML EDA Commons.* ML EDA Commons must ensure standards for the proposed schema, including but not limited to the following.

(1) *Modularity and extensibility.* The Commons should, for example, standardize a schema supporting multiple stages of the EDA flow (RTL, synthesis, place-and-route, etc.), extensible to allow adding new data fields or components for evolving ML EDA applications.
(2) *Interoperability.* Data formats should be compatible with existing EDA tools (proprietary and open-source) and ML frameworks.
(3) *ML-friendly representation.* Data should be stored in forms easily consumable by ML algorithms, e.g., tabular data for metrics, graph-based representations for netlists and circuit connectivity, and image-like formats for congestion maps or layouts. It must allow for ML-specific annotations, such as labels for supervised learning or features for reinforcement learning.
(4) *Scalability.* ML EDA methods must handle large datasets efficiently, including hierarchical designs and full-chip layouts.
(5) *Human and machine readability.* The chosen data formats should ensure a balance between readability (for debugging and inspection) and machine-optimized formats (for speed and scalability).

4.3.2 *Standards for data quality.*

An ML EDA Commons must establish standards for dataset quality that include robust safeguards against data poisoning through automated validation pipelines to ensure data security and integrity. These standards should mandate comprehensive documentation for all datasets, specifying details of data collection methods, tools, and parameters used. Datasets must also exhibit diversity, encompassing a broad range of PDKs and design classes such as RISC-V processors or accelerators. High-quality benchmarks should require datasets, such as RTL collections, to include synthesized and verified designs, with complete, open enablement extending to GDSII generation. Synthetic or proxy datasets should meet strict quality criteria to ensure their utility and accuracy for ML training. Additionally, the integration of CI/CD pipelines for dataset updates and maintenance will help maintain consistency and reliability, and support the growing needs of the ML EDA community.

4.3.3 *Leaderboards for standard benchmarks.*

Standard benchmarks enable leaderboards that provide a consistent framework for evaluating new methodologies. The ML EDA Commons must implement live, CI/CD-based leaderboards that are dynamically and automatically updated to assess submissions with the latest builds of ML EDA models and tools. An example template can be based on nightly builds of OpenROAD-flow-scripts [44]. By tracking live metrics such as runtime and power-performance-area (PPA), these leaderboards promote frictionless reproducibility and research innovation.

4.4 Accessibility and reproducibility

Accessibility refers to the ease with which researchers can obtain and utilize essential EDA tools, flows, models, PDKs, compute resources, and datasets. Open-source resources are pivotal as they are ready to use, lowering barriers to entry for newcomers. Addressing challenges related to computational resources is essential to making ML EDA broadly accessible, and not limited to groups that can afford the computing resources. ML EDA Commons efforts can also promote reproducibility of research through open sharing of data and code.

4.4.1 *Role of open source in accessibility.*

Open-source tools [14], flows [44, 45], PDKs [15–17], and models are vital for accessibility.
Tools, flows, and PDKs. Open-source EDA tools play a pivotal role in enabling scalable data generation for ML EDA research, offering the freedom to run multiple tool instances without license restrictions, thereby supporting parallelized data generation. Moreover, these tools provide a unique opportunity to integrate ML directly within EDA workflows – an area traditionally inaccessible to anyone outside of EDA tool vendors. This capability has the potential to drive significant advancements in EDA optimizations. For example, OpenROAD [14] has been extensively used in prior ML EDA research [46, 63] for scalable data generation and has recently been leveraged for integrating ML inside the tool using Python APIs [64] for timing optimization. Establishing a Commons that incorporates standards and open resources can further incentivize the adoption of these tools.

Open-source workflows [44, 45] allow researchers to replicate studies and validate results. These workflows are particularly beneficial for newcomers, as they provide comprehensive flows that would be challenging to develop without extensive chip design expertise. They also enable cross-stage ML EDA research, such as in [63], where OpenROAD-flow-scripts was used to predict post-detailed routing timing during the global routing timing optimization stage, demonstrating how ML can enable flow-specific improvements.

Furthermore, open-source flows and tools facilitate benchmarking, which is often limited in commercial counterparts. With CI/CD pipelines, flows such as OpenROAD-flow-scripts and the RDF repository [45] enable tracking the best achievable PPA for specific benchmarks and flows, providing a transparent foundation for research.
Open-source ML EDA models. Open-source foundation models for ML EDA will be crucial in a data-scarce domain, much as open-source foundation models have propelled research in other ML fields. An ML EDA Commons can play a pivotal role in guiding the development of such models by establishing standardized practices for hosting them in a shared environment. It can also provide clear guidelines for accessing and utilizing these models, enabling users to easily download, fine-tune, or run inference through simple, well-documented APIs. Such an approach will ensure accessibility and encourage adoption.

4.4.2 *Cloud compute resources.*

The ML EDA Commons should provide accessible computing resources for researchers. When large AI models require extensive training times on state-of-the-art GPUs, ML EDA research can be out of reach for smaller groups with limited budgets. The Commons must establish partnerships with industry cloud providers, such as NVIDIA, Google Cloud, AWS, and Azure. These resources will be available for Commons users, as well as for hosting and maintaining Commons infrastructure.

By leveraging academic program initiatives from leading providers, the Commons can offer credits and resources to democratize computing. Additionally, these resources can be allocated to the community for hosting AE and reproducibility challenges [65].

4.4.3 *Artifact evaluation and reproducibility contests.*

Artifact evaluation is a critical initiative to promote open, reproducible research. The process allows authors to submit the codes, datasets, training scripts, and inference scripts used to produce the key results of their accepted papers. These submitted elements, known as *artifacts*, undergo peer evaluation to check if they meet standards of availability (*available* badge), functionality (*functional* or *reviewed* badge), and reproducibility (*reproducible* badge) defined by ACM/IEEE. Papers that meet these standards are awarded badges. In 2024, the MLCAD symposium [54] adapted ACM badges and developed AE standards based on MLCommons [66] to create criteria tailored to the ML EDA community [67]. However, *the above taxonomy itself merits further discussion. For instance, the "available" badge alone may hold limited value without confirmation of "functional" or "reproducible" artifacts.* The ML EDA Commons must refine these standards.

In addition to AE, peer-reviewed reproducibility challenges such as [65] can serve as examples of efforts that prioritize reproducibility. The ML EDA Commons must partner with conference organizers to support and organize such efforts to elevate reproducibility to a first-order priority in research publishing.

5 Roadmap to an ML EDA Commons

As adjacent ecosystems such as [27, 28, 30] – along with many others worldwide – continue to emerge, now is an opportune time to establish an ML EDA Commons and capitalize on the existing momentum in the semiconductor and design domains. We now outline roadmaps and proposed metrics of success for development of the ML EDA

Commons components highlighted in Section 4. These roadmaps and metrics are designed to guide development, establish priorities, and ensure sustained efforts over time.

(1) Maturing ML EDA infrastructure. The development of ML EDA infrastructure can occur incrementally and in parallel for each component (dataset, models, etc.). As an example, we propose a roadmap and metrics for the development of datasets, divided into three phases, with foci for each stage detailed as follows.

- *Early stage (Years 1 and 2).* Focus on building datasets that are large enough in *scale* to support the training of large models. Metrics for evaluation can include the number of updated and newly introduced benchmark suites, and the number that use agreed-upon evaluators; additionally, community engagement can include contributors, pull requests, papers, and contests that use the datasets.
- *Intermediate stage (Years 3 and 4).* Prioritize *high-quality* datasets, with emphasis on completeness and reproducibility. Metrics at this stage include the number of benchmarks equipped with golden checkers and evaluators; the number of benchmarks supported by live leaderboards; and metrics that capture dataset diversity, such as coverage of different design classes and PDKs.
- *Advanced stage (Years 4 and beyond).* Ensure that datasets *remain relevant* by addressing real-world challenges and aligning with industry advancements. Metrics include support of new (manufacturable or proxy) PDKs; the number of new or updated designs and IPs; and ongoing monitoring of community engagement and broader impacts of the datasets.

Some aspects of development in each stage, such as creating and updating datasets, can be achieved by individual research groups. On the other hand, infrastructure efforts such as maintaining live leaderboards with CI/CD require engineering support and adherence to standards. These are areas where the ML EDA Commons can invest resources to ensure robust and sustainable infrastructure.

(2) Efforts toward standardization and governance. Standardizing and establishing a governance structure for the ML EDA Commons can also be performed in three stages, as follows.

- *Early stage (Years 1 and 2).* Define *basic data formats*, benchmarks, and protocols to ensure interoperability across tools and workflows. Collaborative efforts with organizations such as Si2 can help formalize initial standards for formats and evaluation metrics. This phase should encourage contributions from individuals, research groups, and companies to promote consensus and up-front adoptability of these early standards.
- *Intermediate stage (Years 3 and 4). Expand governance activities*, including formation of an advisory board to oversee the maintenance of Commons-related standards. Policies for contributions of any infrastructure component or updates to standards should also be implemented. Centralized infrastructure, such as repositories or platforms to host datasets, benchmarks, etc., is essential at this stage.
- *Advanced stage (Years 4 and beyond).* Create a *self- sustaining and relevant ecosystem* supported by a formalized governance structure. Standards and policies should remain adaptable to accommodate emerging advancements in ML EDA. Ongoing collaboration between academia, industry, and the open-source community will be critical to ensuring relevance of these standards.

Success metrics include measures of standards adoption, e.g., number of papers and contests that use the format. Other metrics could include feedback from the community regarding usability and interoperability, and the number of users and contributors to the centralized infrastructure. Activities in standardization and governance must proactively and systematically solicit inputs from the community and stakeholders.

(3) Efforts toward accessibility and reproducibility. Activities toward accessibility and reproducibility can also proceed in stages. For example, AE efforts for research publications can be advanced according to the following roadmap.

- *Early stage (Years 1 and 2).* Early activities include establishing AE practices at conferences, guided by IEEE/ACM badging policies. Conferences can begin by encouraging authors to submit artifacts alongside their papers and awarding badges to recognize excellence in reproducibility. Hosting of reproducibility challenges, modeled after successful examples such as [65] can further promote community participation and collaboration.
- *Intermediate stage (Years 3 and 4).* Activities in this stage focus on updating the use of badge taxonomy to incentivize meaningful artifacts. Conference organizers can integrate AE more formally into their processes, e.g., requiring artifact submission and/or the use of standard benchmarks for specific tracks or research categories.
- *Advanced stage (Years 4 and beyond).* More advanced efforts would pursue culture changes for ML EDA research, such as changing the peer review process to evaluate submitted papers not only for scientific merit but also for the completeness and reproducibility of artifacts. Conferences should also provide long-term infrastructure to host artifacts, leaderboards for reproducibility challenges, and tools to track the impact of artifacts on the community.

Progress metrics include the number of conferences adopting AE processes, the number of papers receiving badges, and participation in the AE review process. Impact metrics, such as the reuse of artifacts in subsequent research and the number of reproducibility challenges hosted, will highlight the broader influence of these efforts.

6 Conclusion

This paper emphasizes the urgent need for an ML EDA Commons to unify and advance ML EDA research for the RTL-to-GDSII flow. We highlight existing initiatives that can help to seed Commons infrastructure, as well as critical gaps in ML EDA research. The paper outlines the main components of an ML EDA Commons which focuses on maturing existing infrastructure and establishing governance, standardization, accessibility, and reproducibility as core principles. We also outline a roadmap toward establishing the ML EDA Commons. These efforts aim to bridge the divide between the open culture of ML research and the traditionally closed EDA domain, ultimately enabling the community to accelerate innovation in ML EDA.

7 Acknowledgments

We acknowledge the contribution of members of the SLICE team, including S. Garg, Y. Chen, C. Hao, A. Tyagi, and M. Quinn. We thank H.-R. Jiang and other ISPD 2025 organizers for the invitation to present this paper.

References

[1] Y.-H. Huang *et al.*, "Routability-driven macro placement with embedded CNN-based prediction model," in *Proc. DATE*, 2019.
[2] S. Liu *et al.*, "Global placement with deep learning-enabled explicit routability optimization," in *Proc. DATE*, 2021.
[3] E. C. Barboza *et al.*, "Machine learning-based pre-routing timing prediction with reduced pessimism," in *Proc. DAC*, 2019.
[4] Z. Guo *et al.*, "A timing engine inspired graph neural network model for pre-routing slack prediction," in *Proc. DAC*, 2022.
[5] U. Sharma *et al.*, "OpenROAD-Assistant: An open-source large language model for physical design tasks," in *Proc. MLCAD*, 2024.
[6] H. Wu *et al.*, "ChatEDA: A large language model powered autonomous agent for EDA," *IEEE T. Comput. Aid. D.*, vol. 43, no. 10, pp. 3184–3197, 2024.
[7] D. Donoho, "Data science at the singularity," *Harvard Data Sci. Rev.*, vol. 6, no. 1, Jan. 2024, https://hdsr.mitpress.mit.edu/pub/g9mau4m0.
[8] J. Deng *et al.*, "ImageNet: A large-scale hierarchical image database," in *Proc. CVPR*, 2009.
[9] "Kaggle." [Online]. Available: https://www.kaggle.com/
[10] "MLCommons," accessed: 2025-01-18. [Online]. Available: https://mlcommons.org/
[11] V. J. Reddi *et al.*, "MLPerf inference benchmark," in *Proc. ISCA*, 2020.
[12] "Report of NSF workshop on shared infrastructure for machine learning EDA," 2023. [Online]. Available: https://sites.google.com/view/ml4eda/home
[13] A. B. Kahng, "Machine learning applications in physical design: Recent results and directions," in *Proc. ISPD*, 2018.
[14] T. Ajayi *et al.*, "Invited: Toward an open-source digital flow: First learnings from the OpenROAD project," in *Proc. DAC*, 2019.
[15] "Skywater open source PDK," accessed: 2025-01-18. [Online]. Available: https://github.com/google/skywater-pdk
[16] "Nangate45," accessed: 2025-01-18. [Online]. Available: https://github.com/The-OpenROAD-Project/OpenROAD-flow-scripts/tree/master/flow/platforms/nangate45
[17] "ASAP7," accessed: 2025-01-18. [Online]. Available: https://github.com/The-OpenROAD-Project/asap7
[18] S. Liu *et al.*, "RTLCoder: Fully open-source and efficient LLM-assisted RTL code generation technique," *IEEE T. Comput. Aid. D.*, 2024.
[19] A. B. Chowdhury *et al.*, "OpenABC-D: A large-scale dataset for machine learning guided integrated circuit synthesis," *arXiv preprint arXiv:2110.11292*, 2021.
[20] Z. Wei *et al.*, "HLSDataset: Open-source dataset for ML-assisted FPGA design using high level synthesis," in *Proc. ASAP*, 2023.
[21] Z. Chai *et al.*, "CircuitNet: An open-source dataset for machine learning in VLSI CAD applications with improved domain-specific evaluation metric and learning strategies," *IEEE T. Comput. Aid. D.*, vol. 42, no. 12, pp. 5034–5047, 2023.
[22] "ICCAD Contest on LLM-Assisted Hardware Code Generation," 2024, accessed: 2025-01-18. [Online]. Available: https://nvlabs.github.io/LLM4HWDesign/
[23] Wikipedia, "Commons," accessed: 2025-01-18. [Online]. Available: https://en.wikipedia.org/wiki/Commons
[24] A. Tumeo, "Open hardware technology commons," presentation at Open Source EDA and Benchmarking Summit birds-of-a-feather meeting, ACM/IEEE Design Automation Conference, July 2022. [Online]. Available: https://open-source-eda-birds-of-a-feather.github.io/doc/slides/Antonino%20Tumeo_BOF.pdf
[25] A. B. Kahng, "A mixed open-source and proprietary EDA commons for education and prototyping," in *Proc. ICCAD*, 2022.
[26] "Pathways to enable open-source ecosystems (POSE)," 2024, accessed: 2025-01-18. [Online]. Available: https://new.nsf.gov/funding/opportunities/pose-pathways-enable-open-source-ecosystems
[27] "Southwest advanced prototyping (SWAP) hub," accessed: 2025-01-18. [Online]. Available: https://microelectronicscommons.org/connect/southwest-advanced-prototyping-swap-hub/
[28] "Enabling access to the semiconductor chip ecosystem for design, fabrication, and training (chip design hub)," 2023, accessed: 2025-01-18. [Online]. Available: https://new.nsf.gov/funding/opportunities/chip-design-hub-enabling-access-semiconductor-chip-ecosystem-design/nsf24-522/solicitation
[29] "The network coordination hub for the national network for microelectronics education program," 2024, accessed: 2025-01-18. [Online]. Available: https://new.nsf.gov/news/nsf-department-commerce-announce-30m-funding-opportunity
[30] "NSTC workforce partner alliance program," 2024, accessed: 2025-01-18. [Online]. Available: https://natcast.org/workforce/wfpa
[31] "NIST notice of funding opportunity: Chips NAPMP," 2024, accessed: 2025-01-18. [Online]. Available: https://www.nist.gov/chips/r%2526d-funding-opportunities/notice-funding-opportunity-chips-napmp
[32] A. Ghose *et al.*, "Use cases and deployment of ML in IC physical design," *Proc. ASP-DAC*, 2025.
[33] S. Thakur *et al.*, "VeriGen: A large language model for Verilog code generation," *ACM Trans. Des. Autom. Electron. Syst.*, vol. 29, no. 3, Apr. 2024.
[34] U. Mallappa *et al.*, "FloorSet-a VLSI floorplanning dataset with design constraints of real-world SoCs," *arXiv preprint arXiv:2405.05480*, 2024.
[35] V. A. Chhabria *et al.*, "BeGAN: Power grid benchmark generation using a process-portable GAN-based methodology," in *Proc. ICCAD*, 2021.
[36] B.-Y. Wu *et al.*, "EDA Corpus: A large language model dataset for enhanced interaction with OpenROAD," in *Proc. LLAD*, 2024.
[37] G. S. P. Kadagala and V. A. Chhabria, "Invited paper: 2023 ICCAD CAD contest Problem C: Static IR drop estimation using machine learning," in *Proc. ICCAD*, 2023.
[38] B.-Y. Wu *et al.*, "Invited paper: 2024 ICCAD CAD contest Problem C: Scalable logic gate sizing using ML techniques and GPU acceleration," in *Proc. ICCAD*, 2024.
[39] R. Liang *et al.*, "GPU/ML-enhanced large scale global routing contest," in *Proc. ISPD*, 2024.
[40] M. Liu *et al.*, "VerilogEval: Evaluating large language models for Verilog code generation," in *Proc. ICCAD*, 2023.
[41] "MacroPlacement," accessed: 2025-01-18. [Online]. Available: https://github.com/TILOS-AI-Institute/MacroPlacement
[42] C. A. E. Goodhart, *Problems of Monetary Management: The UK Experience*. London: Macmillan Education UK, 1984, pp. 91–121.
[43] A. Liu *et al.*, "DeepSeek-V3 technical report," *arXiv preprint arXiv:2412.19437*, 2024.
[44] "OpenROAD-flow-scripts," accessed: 2025-01-18. [Online]. Available: https://github.com/The-OpenROAD-Project/OpenROAD-flow-scripts
[45] "IEEE-CEDA-DATC Robust Design Flow," accessed: 2025-01-18. [Online]. Available: https://github.com/ieee-ceda-datc/Robust-Design-Flow
[46] V. A. Chhabria *et al.*, "OpenROAD and CircuitOps: Infrastructure for ML EDA research and education," in *Proc. VTS*, 2024.
[47] J. Jung *et al.*, "IEEE CEDA DATC: Expanding research foundations for IC physical design and ML-Enabled EDA," in *Proc. ICCAD*, 2022.
[48] ——, "Invited paper: IEEE CEDA DATC emerging foundations in IC physical design and MLCAD research," in *Proc. ICCAD*, 2023.
[49] S. Liu *et al.*, "Towards big data in ai for eda research: Generation of new pseudo circuits at RTL stage," *Proc. ASP-DAC*, 2025.
[50] D. Kim *et al.*, "Construction of realistic place-and-route benchmarks for machine learning applications," *IEEE T. Comput. Aid. D.*, vol. 42, no. 6, pp. 2030–2042, 2023.
[51] "SLICE: A Shared machine Learning Infrastructure for the EDA Community," accessed: 2025-01-18. [Online]. Available: https://slice-ml-eda.github.io/
[52] R. Aslett, "Q2 2024 CEO Message," Silicon Integration Initiative, July 2024, accessed: 2025-01-18. [Online]. Available: https://si2.org/q2-si2-newsletter/
[53] [Online]. Available: https://www.islad.org/
[54] [Online]. Available: https://mlcad.org/symposium/2025/
[55] "National Science Foundation Workshop: "ImageNets" for EDA," accessed: 2025-01-18. [Online]. Available: https://wp.nyu.edu/imagenets_eda/
[56] S. Fenstermaker *et al.*, "METRICS: A system architecture for design process optimization," in *Proc. DAC*, 2000.
[57] A. Kahng and S. Mantik, "A system for automatic recording and prediction of design quality metrics," in *Proc. ISQED*, 2001.
[58] J. Jung *et al.*, "METRICS2.1 and flow tuning in the IEEE CEDA robust design flow and OpenROAD ICCAD special session paper," in *Proc. ICCAD*, 2021.
[59] "DATC-RDF-Metrics4ML," accessed: 2025-01-18. [Online]. Available: https://github.com/ieee-ceda-datc/datc-rdf-Metrics4ML
[60] R. Liang *et al.*, "CircuitOps: An ML infrastructure enabling generative AI for VLSI circuit optimization," in *Proc. ICCAD*, 2023.
[61] P. Shrestha *et al.*, "EDA-schema: A graph datamodel schema and open dataset for digital design automation," in *Proceedings of the ACM Great Lakes Symposium on VLSI*, 2024.
[62] "CircuitOps," accessed: 2025-01-18. [Online]. Available: https://github.com/NVlabs/CircuitOps
[63] V. A. Chhabria *et al.*, "From global route to detailed route: ML for fast and accurate wire parasitics and timing prediction," in *Proc. MLCAD*, 2022.
[64] W. Jiang *et al.*, "IR-aware ECO timing optimization using reinforcement learning," in *Proc. MLCAD*, 2024.
[65] "ML reproducibility challenge," accessed: 2025-01-18. [Online]. Available: https://reproml.org/
[66] "MLCommons artifact evaluation," accessed: 2025-01-18. [Online]. Available: https://github.com/ctuning/artifact-evaluation/blob/master/docs/reviewing.md
[67] "MLCAD artifact evaluation," accessed: 2025-01-18. [Online]. Available: https://github.com/ml-eda/artifact-evaluation