# DesignCon 2021

## Bridging the Organization Gap for EDA Machine Learning Data

Kerim Kalafala, IBM
kalafala@us.ibm.com

Hui Fu, Intel
hui.fu@intel.com

W. Rhett Davis, North Carolina State University
rhett_davis@ncsu.edu

Karthik Aadithya, Sandia National Laboratory
kvaadit@sandia.gov

Leigh Anne Clevenger, Silicon Integration Initiative
leighanne.clevenger@si2.org

## Abstract

Researchers at chip design companies, foundries, EDA tool suppliers, academic institutions, and national laboratories cannot investigate novel Machine Learning Algorithms without organized volumes of trusted, suitably labeled, non-proprietary design data for training. Access to quality data is fundamental to advancements in machine learning.

Undoubtedly, progress is being made in applying ML to EDA; however, big-payoff concepts require access to big data sets. The companies best positioned to make gains in this area are chipmakers with internal EDA and semiconductor development capabilities, with a reservoir of archived IC design and manufacturing data. Even then, processed data—data describing the human and tool actions that transpired at each process step, which are key to enable the next level of intelligent optimization—have often been left uncaptured, or, at least, uncaptured in a format usable to empower ML.

A comprehensive solution to the data access and formatting challenges that collectively supports competitive, university, and company research and development, is sorely needed.

This work presents motivation, requirements, and essential components for a collaborative, secure learning API for EDA, based on recent work by the Silicon Integration Initiative AI/ML in EDA Special Interest Group. Stakeholder authors, including chip design companies, foundries, academic institutions, and national laboratories, present their problem statements for a collaborative secure learning API through use case scenarios specifically targeting the ML training data gap.

## Author Biographies

**Kerim Kalafala** is a member of the IBM Academy of Technology, a Senior Technical Staff Member in the IBM Systems Group, and an IBM Master Inventor. He currently serves as lead architect of EDA analytic and static timing software tools used to design and verify the world's fastest microprocessors. He received his graduate degree in Computer and Systems Engineering from Rensselaer Polytechnic Institute.

**Hui Fu** is a senior principal technologist in the Design Enablement of Technology development group at Intel. His focus for design enablement is to leverage automation and AI/ML capabilities to accelerate design technology co-optimization, to optimize design rules for new technology nodes, and to improve product design kit quality. Prior to his current role, he worked as the design director of wireless baseband development and led Intel's last four generations of modem SoC development. He has held multiple R&D leadership positions with Intel, Infineon and Siemens in Germany, Singapore, China, and the U.S.

**W. Rhett Davis** is a Professor of Electrical and Computer Engineering at North Carolina State University and a founding faculty member of the Center for Advanced Electronics through Machine Learning (CAEML). He received M.S. and Ph.D. degrees in Electrical Engineering from the University of California, Berkeley. His research centers on electronic design automation for integrated systems in emerging technologies. He is best known for his efforts in design enablement, 3D-IC design, thermal analysis, circuit simulation, and power modeling for systems-on-chip and chip multi-processors.

**Karthik Aadithya** is a Senior Member of Technical Staff in the Electrical Models & Simulation Department at Sandia National Laboratories. He obtained his Ph.D. in Electrical Engineering from the University of California, Berkeley. His research focuses on developing algorithms and computational techniques for accurately modeling, analyzing, simulating, verifying, and debugging current- and next-generation electronic and biological systems.

**Leigh Anne Clevenger** is the Director of OpenStandards at Silicon Integration Initiative, Inc. (Si2), where she focuses on accelerating the Si2 OpenStandards Coalition collaboration efforts in machine learning and system level power modeling. She earned her doctorate in Software Engineering and Machine Learning at Pace University and has extensive experience in semiconductor design automation and semiconductor processing technology.

# 1. Introduction

Recent surveys, such as the 2020 Si2 AI/ML in EDA Survey, highlight that the current *status quo,* involving highly siloed development with proprietary formats, creates significant obstacles to collaboration in the field of EDA [1]. These difficulties largely arise from proprietary hurdles. Key issues include the lack of a standard means for partners to exchange processed data, the duplication of efforts needed to avoid prohibited exchanges of data, and access to metadata from an EDA toolset. The term *processed data* refers to any data derived from proprietary data, e.g., path delays for a circuit that have been generated using gate delays from a proprietary timing library. In a recent Silicon Integration Initiative (Si2) whitepaper, experts representing chip design companies, foundries, academic institutions, and national laboratories provide perspectives on the need for a holistic API to present processed data pertaining to IC design. We refer to this as a collaborative secure learning API.

A recent *Semiconductor Engineering* article presents the challenges for ML in EDA, and questions whether EDA is an appropriate area for ML [2]. Availability of sufficient high-quality data and the difficulty in sharing data from EDA tools are chief among the author's concerns. Obstacles to data collaboration include unease about data security and privacy, the potential for back-tracing original data, even after obfuscation, the cost of API and security implementation versus return on investment, and a desire to keep ML data and solutions proprietary.

Section 2 describes field publications providing motivation and building blocks for a collaborative secure learning API. In Section 3, stakeholder authors from chip design companies, foundries, academic institutions, and national laboratories present their problem statements for a collaborative secure learning API through use case scenarios. Section 4 summarizes common problems in the semiconductor industry that can be addressed through secure data exchange, and common requirements for a collaborative secure learning API. This is followed by Section 5's discussion of example paths to address these challenges and requirements through the collaborative secure learning API, as well as advantages and disadvantages of certain resources which may address stakeholder challenges in the absence of an API. Section 6 presents the authors' conclusions.

# 2. Background

The motivation for a standard, common methodology for classification and structure of machine learning training and inference data for interoperability has been described in "A Collaborative Data Model for AI/ML in EDA," a white paper published by the Si2 AI/ML in EDA Special Interest Group with authors from eleven companies [3]. With examples of both analog and digital use cases, this paper highlights ways in which processed data from a common API can be optimized for machine learning training,

including the processes of determining data relevance, data cleaning, and data transformation.

Building blocks for a standard processed data API currently exist in open-source EDA endeavors such as the OpenROAD project. The OpenROAD "Safe Names" Conventions for the RTL-to-GDS space outline a format to define names for API input and output, without using proprietary or trademarked EDA tool names, by using verb+object+modifier [4]. OpenROAD defines metric-naming conventions based on flow stage names, run-level and tool metric-level nouns, and modifiers which can be recorded using Python and JSON [5]. These metrics can be used for ML training data. The IEEE CEDA Robust Design Flow (RDF) calibration data focuses on open-source interfaces to Logic Synthesis, Static Timing Analysis, and Detailed Routing intersections where tools can be freely swapped [6].

Security and privacy methods for EDA processed data are critical to enabling sufficient data-sharing to power ML methods such as neural networks. State-of-the-art cybersecurity applied to EDA processed data can assuage the concerns of potential collaborators. These methods include differential privacy that can prevent recreating input data from trained models [7], deterministic privacy that constrains database access at different security levels [8], and federated learning that enables secure cross-functional collaboration [9].

# 3. Use Case Problem Statements

A collaborative secure learning platform for EDA would contribute significantly to solving many pressing industry problems. Here, stakeholder authors present desired use-cases for such a collaborative ML platform as a set of tasks to be performed. The tasks are organized by stakeholder category: chip design companies, foundries, academic institutions, and national laboratories.

## Chip Design

When assembling a design for fabrication, a commercial chip design organization must execute a wide range of EDA design flows, from synthesis to physical design and verification, using multiple software tools and logical and physical IP obtained from many partners. These flows can be complex, and are often unpredictable. Machine learning is already used in many tools and flows but could bring great efficiency gains to many more. When developing these flows, a chip design organization may wish to do the following:

- *Collaborate with an external partner to develop and deploy an ML model training flow.* Developing a shared model training flow requires the sharing of sample data in a partially anonymized fashion. This data is used for benchmarking, flow

tuning, and training of ML models, and it must be both sent and received in a secure manner.

- *Validate that an external partner's EDA methodology is working as expected.* Such validation would involve running a portion of a partner's EDA methodology, perhaps on a public cloud, in a way that protects both design data and processed results. This may include both training and inferencing.
- *Validate that IP procured from a provider is being analyzed properly within a design-flow.* Validating IP would involve obtaining sample processed data (*e.g,* timing, power, or noise results) in order to validate that the IP in question is being analyzed properly within a particular design methodology.

Unfortunately, it is difficult for commercial chip design companies to fully collaborate as described while protecting proprietary design content. A collaborative secure learning API would provide a means of identifying and protecting sensitive information while enabling straightforward access to non-sensitive data.

## Foundry

Thanks to the success of ML in computer vision-related applications, foundries are among the early adopters of machine learning techniques to improve lithography and mask synthesis. ML-based lithography hot spot prediction is currently used to improve yield and design for manufacturability (DFM), while ML-based optical proximity correction and sub-resolution assist feature insertion are used to optimize mask synthesis efficiency. Foundries can claim some early victories in the lithography and mask synthesis areas, but machine learning is far from reaching its potential in aiding process development. Key collaborative ML-based tasks to unleash foundry potential include the following:

- *Simplify and accelerate Design Technology Co-Optimization (DTCO)*. Foundries are relying more and more on DTCO to drive the Performance, Power, Area, and Cost/Yield (PPAC) gains from one process generation to the next. DTCO now accounts for more than half of those gains for advanced process nodes. A fast loop from design rule definition to lithography DFM impact and design PPAC impact is required for modern foundry process development. Pure-play foundries, in particular, must exchange benchmark designs with customers in a secure manner in order to optimize the different types of designs (high performance, low power, cost sensitive, logic centric or memory centric). All this requires a secure way to access design data and establish a common "design dataset" for benchmarking. Similar situations arise for machine learning benchmarks between different AI hardware accelerators. Hardware vendors had their own benchmark models and published results based on non-standard benchmarking models until this was addressed by the MLperf benchmark set, which provided a common reference against which to compare the offerings from different machine learning hardware acceleration solutions. For DTCO, establishing a "public benchmark design dataset" to calibrate and benchmark all foundry PPAC claims/entitlements

is required. A secure and collaborative learning mechanism to build this dataset will help to accelerate this work.

- *Enable early-phase process development for advanced technologies.* In the early phases of development, foundries use synthetic patterns (e.g., a 6x6 squish pattern) to explore the range of possible patterns to fabricate and drive yield learning; however, patterning for advanced nodes will require a significantly larger matrix (e.g., a 16x16 squish matrix pattern) to capture the essential transition windows. A 6x6 squish matrix represents a pattern space topology greater than 68 billion patterns and 16x16 matrix makes the overall pattern space intractable. Fig. 1 illustrates the difficulty of covering the pattern space using the random-synthetic-pattern approach. Red dots in the figure represent the layout patterns extracted from the design database that are not covered, and the synthetic patterns generated are potentially areas of future yield surprises. To solve this problem, foundries must obtain early sample designs to identify high-probability patterns from the sea of possible patterns.
- *Reduce the cost of EDA tool-chain integration.* In order for ML to resolve the first two issues above, foundries must overcome the formidable tasks of exporting the design database from the EDA tool chain, performing inference on a custom-trained ML model, fixing any reported issues within and without the EDA tool chain, and importing the results back into the EDA tool chain to complete the flow. Fig. 2 shows an example of an ML-based hotspot prediction and fixing training and inference process. During the inference process, EDA tool processed data is exported to an external ML model input data feature set. After the ML inference and hotspot fixing, data is imported into the EDA toolset to complete the process. The overhead cost of database export/import is likely to nullify much of the efficiency gains from ML without an efficient API.

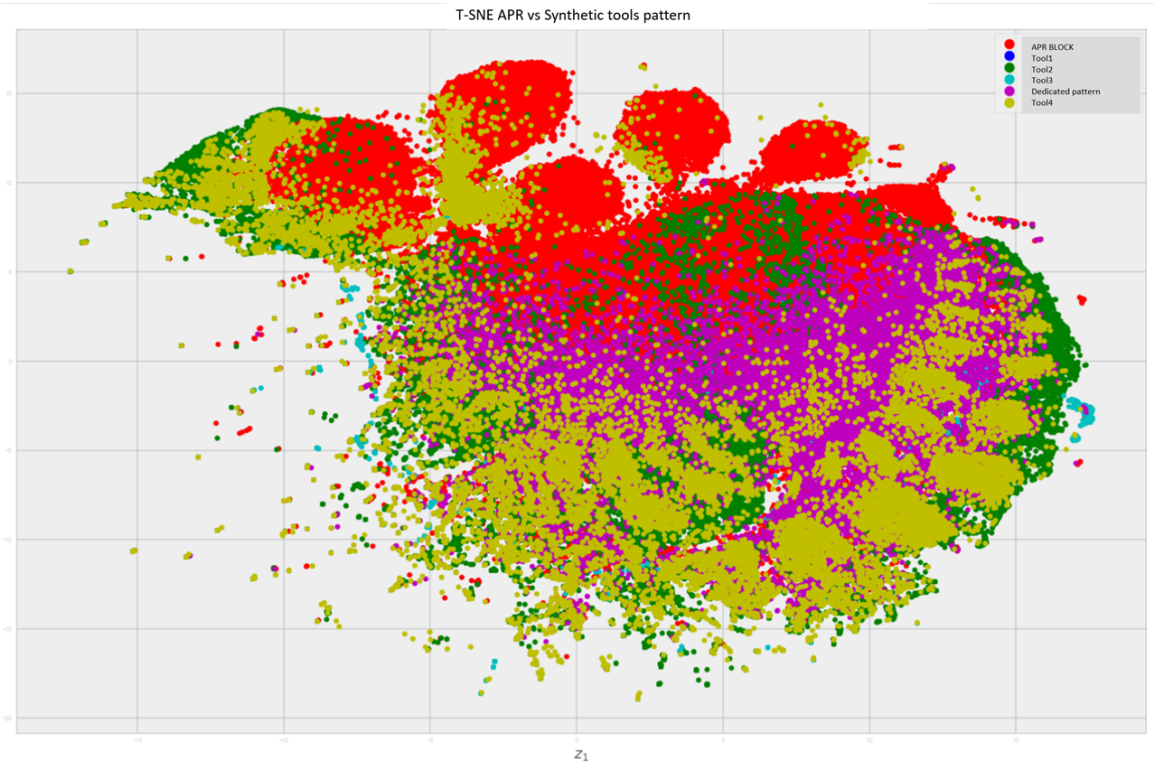


Fig. 1 – Comparison of the pattern spaces of an APR layout
to various synthetic pattern spaces [10]

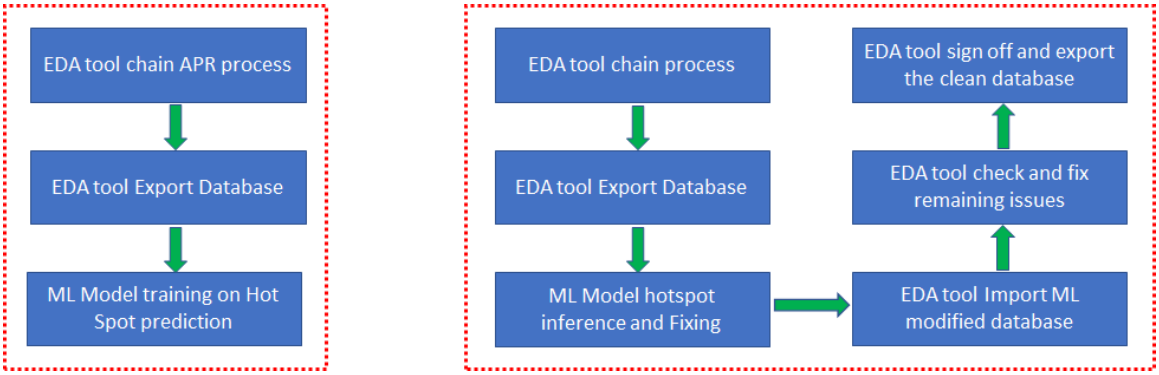


Fig. 2 – ML training Process, ML inference Process [11]

## Academia

Because academic researchers often work to solve the problems of both chip design organizations and foundries, they are assumed to have similar use-cases to the ones already presented. Additionally, academic researchers are often free to explore new algorithms or techniques that are considered too risky for commercial entities. ML algorithms hold great promise and present many opportunities for academic research;

however, academics are also required to publish their work and present convincing evidence of their results to sponsors and publication venues. In the process of performing these tasks, an academic researcher may wish to do the following:

- *Experiment with new algorithms using realistic data.* Academic researchers often demonstrate new models or algorithms with synthetic data, which can be easily dismissed by industry researchers as unrealistic. Industry partners may attempt to obtain access to proprietary data for their academic partners but find themselves unable to satisfy their company's data sharing safeguards. As a result, the technology transfer from academia to industry is often severely limited.
- *Publish work without violating license agreements.* In the process of their investigations, academic researchers often need to evaluate the quality of one model or training algorithm relative to another, duplicating the results of another researcher to ensure proper understanding of the technology. When publishing their work, they must ensure that they do not unintentionally leak confidential information; this can be especially difficult if the new model or algorithm create data derived from a tool or library with licensing that prohibits benchmarking. Researchers inherently want to release their subject data so that the results of their work can be duplicated and expanded. If the results can be used to reverse-engineer a tool's performance, then the EDA company could revoke the researcher's license for the tool, and potentially sue the university.

## National Laboratories

National laboratories often design circuits and systems that have national security implications. These systems have to perform to their specifications in harsh, sometimes radiation-heavy environments; therefore, national laboratories construct their own, internal models for electrical component behavior in such environments, as well as unique design flows for analyzing, simulating, verifying, and validating their systems under extreme conditions. Such internal models and design flows typically cannot be shared outside of the laboratory (including with EDA vendors, academia, or other chip design houses) due to security concerns. In the process of designing these circuits, national laboratories may wish to do the following:

- *Use the latest ML-enabled EDA tools without violating government security protocols.* National laboratories stand to benefit from the performance improvements that accompany each new technology node, as well as from the latest EDA techniques and algorithms developed by researchers in academia or at chip design companies, which leverage ML to solve EDA problems commonly encountered in the newest technologies (e.g. timing analysis, parasitic extraction and layout, characterization of noise margins and variability, etc.). Under the *status quo*, security concerns often prevent national laboratories from taking advantage of these improvements.
- *Share non-sensitive ML research findings with the EDA community.* Not all securely developed models and design flows are security sensitive. As a taxpayer-funded entity, national laboratories wish to offer their non-sensitive ML/EDA

research findings to the rest of the EDA community; however, extracting these new algorithms, models, tools, and techniques from security-sensitive data is often a difficult task.

ML will drive the future of EDA, and, unlike traditional EDA approaches, ML-based tools and techniques tend to vastly improve with increased availability of both primary and processed data. It is therefore in the interest of the national labs to maximize collaboration on ML initiatives with academia and industry participants, while respecting their own security constraints and the confidentiality/licensing/trade-secret concerns of their would-be collaborators.

# 4. Summary of Common API Requirements

From a commercial chip-design standpoint, a collaborative secure learning platform would provide a standard means of sharing the processed data needed to address IP protection, including the various levels of security required by the myriad application domains. This IP protection component is vital to enabling collaboration.

Furthermore, given the highly integrated nature of commercial-grade EDA flows, linking processed data to underlying physical structures, as well as managing across multiple levels of hierarchy, is of critical importance, particularly for big data analytics and machine-learning. The digital implementation flow of a commercial chip design house, for example, often includes a mix of homegrown and commercial tools. Hierarchical design methodologies have evolved to group large designs into smaller sub-units, which require careful integration of the implementation with analysis results to ensure the system functions properly. As a result, a holistic framework for collaboration (i.e., one that supports hierarchy, as well as linking multiple domains of processed data back to the appropriate design content) is needed to address the industry demand to operate efficiently with multiple partners in a variety of environments.

Additionally, there are multiple opportunities to analyze the physical design of digital ICs during iterations of the design process. These include not only evaluation of the results, but also identifying the particular inputs that lead to the results and, ultimately, driving improvements [12]. A viable analysis system will allow queries at the intersections of logic, placement, wiring, timing, power, noise, etc. A tool-agnostic, efficient API layer would also need offline query capability into a known data model, as well as capacity for in-process queries.

Fig. 3 indicates that existing models, such as Si2's OpenAccess database, can be used to store design data. A processed (or derived) data model for the output of timing, noise, or power analysis can be linked to the design data through a unified persistent data model to enable ML training. This elevates the processed data to primary data objects, as opposed to properties which require connection to a design data object, as illustrated in Fig. 4.

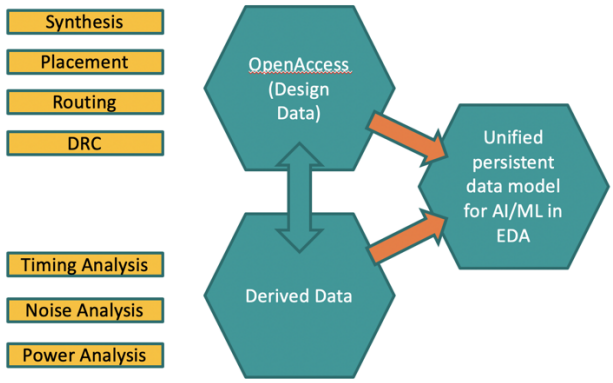


Fig. 3 – High-level representation of a Unified Persistent Data Model for enabling AI/ML in EDA [3]

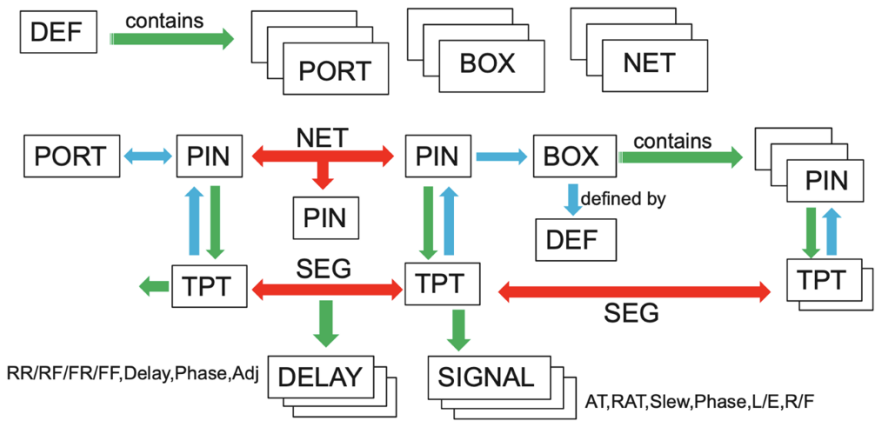


Fig. 4 – Processed data, such as timing, as a first-class citizen of a Unified Persistent Data Model [3]

Multiple points in the implementation flow can have API connections to the design and processed data, referred to as derived data (DD) in Fig. 5. This unified data model should be designed for minimal impact to implementation run time.

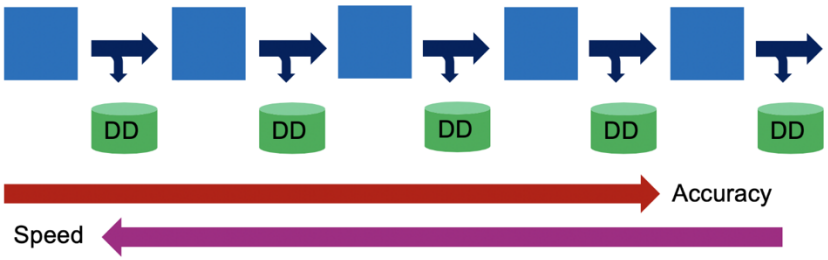


Fig. 5 – Multiple API connections to implementation flow [3]

An industry-standard secure API for processed output data would provide important infrastructure to enable AI and ML training and inference from data produced by a variety of EDA tools. This API could replace separate data translation and tagging programs to reduce the turnaround time of providing structured and properly labeled datasets, enabling innovation through secure collaboration, as shown in Fig. 6.

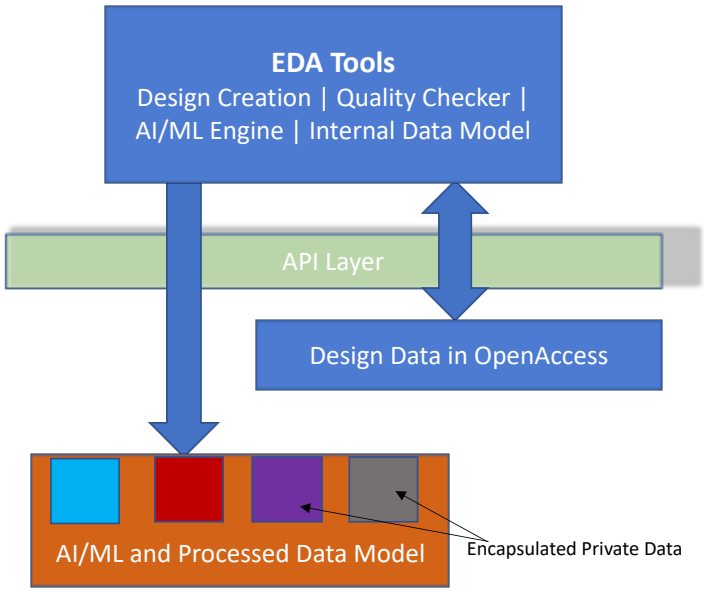


Fig. 6 – Secure Processed Data API

# 5. Discussion

Given the stakeholder use cases and the requirements for a collaborative secure learning API, this section describes possible scenarios for an API and the different levels of data security needed to address these challenges. Until a collaborative secure learning API is complete, design teams will continue utilizing other approaches to obtaining ML training data. Potential training applications of open-source datasets and synthetic data generation are described here.

## EDA Built-in API for Customer ML Model

Design companies can customize ML model and EDA tool chains to interact with each other in a variety of ways. One option is ML model integration in the EDA toolset, which would provide a native API to access the processed data within the EDA toolset. Without the native API to the EDA tool chain, the user exports the data from the EDA tool chain, trains the model, performs inference, then imports the data back to the EDA tool chain and completes the remaining steps.

There are established native APIs for designers to access, but these are often limited to specific application areas. One next step would be to expand the native API access to other design application and signoff databases, and to allow users to run their ML model in conjunction with the database residing in the EDA tool chain.

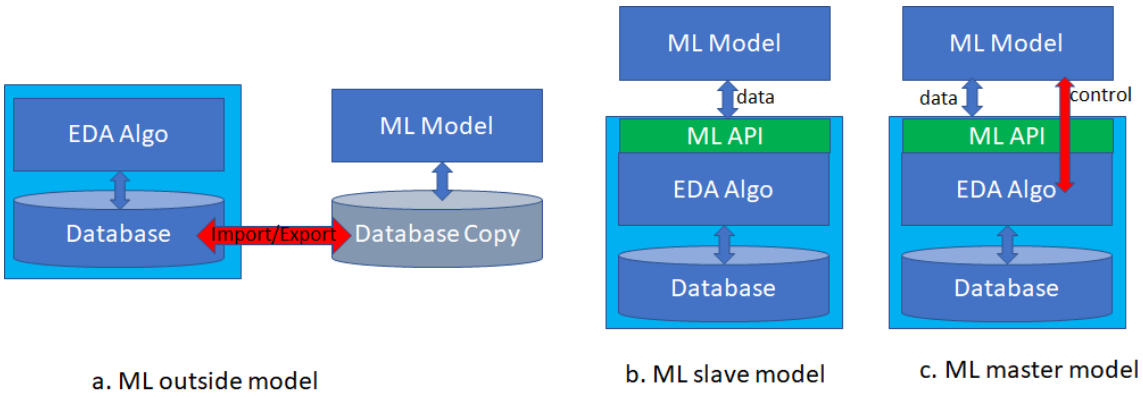


Fig. 7a-c – Three possible API implementations

Fig. 7a represents a current machine learning interaction model with an EDA tool chain. This model requires the database to be exported from the current EDA toolset and the machine learning inference to be conducted outside of the EDA tool chain.

Fig. 7b shows an API implementation that supports external ML models, in which the ML model could access the data residing in the EDA toolset and perform inference through the EDA ML backend. EDA toolsets can then use the ML inference results to perform the necessary actions to correct and/or modify the database. One typical implementation would be ML-based hotspot prediction (for EM/RV, DFM hotspot, congestion hotspot, etc.). EDA controlled software would query the ML model based on the current database or implementation and perform the necessary corrections. The external ML model would then be called through the API and respond to requests for action. Under this model, the API could act as a typical ML API backend (e.g., TensorFlow, PyTorch).

The structure outlined in Fig. 7c would allow the ML model to control the EDA software, based on the inference result, to perform necessary corrections. The ML model would also control the overall flow of the EDA tool setting to obtain the best performance. This reinforcement learning-based ML model would then query the implementation results

based on this setting, and output optimal tool settings for the next run based on the given constraints.

## Levels of Data Security

Distributed learning and database security for semiconductor design are areas of growing interest and concern. Many of the models envisioned for ML in EDA require training data samples and features that are scattered among different organizations. Recent research in distributed learning techniques could allow adversarial companies to train an ML model collaboratively, while protecting trade secrets that competitors may attempt to infer from the training data. Differential privacy can guarantee that secret information cannot be probabilistically inferred from the shared information, and is considered the "gold standard" for privacy protection [13]. The federated learning algorithm allows collaborative ML model training without moving the training data across organizational boundaries [9]; however, models trained by FL and other learning paradigms can leak private training information, leading to possibilities of training data recovery [14].

Differentially private ML approaches rely on the ability to "hide" the private information within the shared training gradients, which may be too risky for some semiconductor companies to consider; these techniques may not always be able to prevent parties from leaking trade secrets to their collaborators. An alternative approach is to publish training data in a secure database, with policies that restrict data accesses and apply inference control methods to determine whether information leakage is possible [15]. As the amount of sensitive information grows, mechanisms enabling private access to secure data must become more fine-grained and complex. These elaborate security policies pose a different kind of information leakage risk, as an adversarial user with incorrectly configured access could deterministically infer sensitive information. Whether privacy is guaranteed differentially, through inference control methods, or by some other means, it is paramount that the common API quantifies information leakage in order to enable collaborative learning.

Differential privacy can be applied to the method of federated learning, which "aims at training a machine learning algorithm, for instance deep neural networks, on multiple local datasets contained in local nodes without explicitly exchanging data samples. The general principle consists of training local models on local data samples and exchanging parameters (e.g., the weights and biases of a deep neural network) between these local nodes at some frequency to generate a global model shared by all nodes" [16].

With federated learning, a design company need not disclose their own design data, and shares only the trained parameters with participating members. Likewise, an EDA company need not reveal their core algorithm, and could merely integrate all distributed parameters. To allow for this kind of collaboration, a federated learning infrastructure, a proper learning community membership structure, and data protection mechanisms must be developed and agreed-upon by the industry.

### Open-Source Datasets

Academia and industry alike are beginning to build open-source datasets for IC design, including open-source PDKs, APR benchmark datasets, hotspot detection public datasets, and digital and analog design datasets. Unfortunately, these available datasets are often:

- Old, outdated, and lacking state-of-the-art design characteristics
- Very small, particularly compared to successful computer vision public image datasets, for example
- Not well-maintained, with no quality assurance on label accuracy and consistency (which leads to reduced training accuracy)

Given these limitations, industry members should seek to build a high-quality public dataset for PDKs, digital design (logic centric design, interconnect centric design, etc.), analog design, and signoff datasets. Performance, power, area (PPA) benchmarks, in particular, would benefit from a common reference dataset. Beyond 28nm, process node naming has lost its meaning, since companies use their own marketing imagination to define the process node, causing confusion in the industry. There is no common terminology for PPA benchmarks. An open benchmark dataset would eliminate the confusion caused by the process node naming and allow for appropriate process comparisons against the same reference.

### Synthetic Data Generation

Another way to produce data for machine learning is through generating synthetic data. Analog layout generators, such as the Berkeley Analog Generator, automatically generate varied layouts for the same function for different topologies, device sizes and routing width-spacing patterns [17]. Foundries often lack the tools for data generation, data augmentation and data management, and there is no standard native API to integrate data from foundries and design companies in a collaborative manner. Synthetic data generation can bridge the gap without requiring the design companies to risk leaking proprietary design data.

## 6. Conclusions

A common secure learning API will enable a smooth collaboration between fabless design companies and foundries to accelerate early phase technology development. In particular, a cross-EDA platform API that enables foundries to deliver their machine learning model-based design rule definition will minimize the gaps between design rule development, lithography recipe development, yield learning and design tool deployment. Other approaches, such as open design data sets, synthetic data generation tool development and federated learning, have great potential to further hasten machine learning adoption by foundries.

Academic researchers would benefit from a collaborative secure processed data API through the opportunity to demonstrate new models and algorithms on data from industry partners, enhancing technology transfer from academia to industry. The API can be used to create ML databases with EDA processed data content, which can be used for common research projects and for academic and industry metrics alike.

National laboratories would drive the security parameters of the API, to the benefit of the EDA industry. Satisfying the security requirements of national laboratories to share data and non-security-sensitive ML research would enrich learning for the industry, and facilitate use of the latest ML-enabled EDA tools by national laboratories.

An industry-standard secure API for processed output data would provide essential infrastructure for AI and ML training and inference from data produced by a variety of EDA tools. This API could replace separate data translation and tagging programs to reduce turnaround time for providing structured and properly labeled data sets, enabling innovation through secure collaboration. Given the highly integrated nature of commercial-grade EDA flows, linking processed data to underlying physical structures, as well as managing across multiple levels of hierarchy, is of critical importance, particularly for big data analytics and machine-learning; hence, a holistic framework for collaboration (i.e., one that supports hierarchy, as well as linking multiple domains of processed data back to the appropriate design content) is needed to address the industry demand to operate efficiently with multiple partners in a variety of environments.

## References


[1] J. Das, A. Kumar, A. Dey, et al., ‘Gaps and Opportunities for AI/ML Techniques in the EDA Domain,’ white paper, 2020, https://si2.org/ai-ml-downloads.

[2] B. Bailey, ‘Roadblocks for ML in EDA,’ SemiconductorEngineering, 2021, https://semiengineering.com/roadblocks-for-ml-in-eda.

[3] K. Kalafala, V. Parthasarathy, N. Chang, et al., ‘A Collaborative Data Model for AI/ML in EDA,’ white paper, 2020, https://si2.org/ai-ml-downloads.

[4] OpenROAD project, ‘OpenROAD Safe Names Conventions v1.0,’ white paper, 2019, https://theopenroadproject.org/wp-content/uploads/2019/12/OpenROAD-Safe-Names-Conventions-v1.0.pdf.

[5] A. Kahng, T. Spyrou, ‘The OpenROAD Project: Unleashing Hardware Innovation,’, GOMACTech 2021, https://vlsicad.ucsd.edu/Publications/Conferences/383/c383.pdf.

[6] J. Chen, I. Jiang, J. Jung, et al., ‘DATC RDF-2019: Towards a Complete Academic Reference Design Flow,’ 2019 IEEE/ACM ICCAD, 2019.

[7] Y. Zhang, R. Jia, H. Pei, W. Wang, B. Li, and D. Song, ‘The secret revealer: Generative model-inversion attacks against deep neural networks’, In Proceedings

of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 253–261, 2020.

[8] R. Chirkova and T. Yu, 'Exact Detection of Information Leakage: Decidability and Complexity,' Trans. on Large-Scale Data- and Knowledge-Centered Systems, Springer, 2017.

[9] J. Konečný, H. B. McMahan, F. X. Yu, P. Richtárik, A. T. Suresh, and D. Bacon, 'Federated learning: Strategies for improving communication efficiency,' arXiv preprint arXiv:1610.05492, 2016.

[10] H. Fu, et al.,'Towards measuring layout pattern coverage: a Machine Learning Approach,' Design Automation Conference, 2021.

[11] H. Fu, et al.,'Machine learning Assisted Design Rule Debug and Rule Ranking Automation,' Design Automation Conference, 2021.

[12] L. Stok et al., 'Empowering the Designer Through Advanced Analytics and Machine Learning,' 56thDesign Automation Conf., Las Vegas, NV, USA, 2019.

[13] C. Dwork. 'Differential privacy: A survey of results', Intl. Conf. on Theory and Applications of Models of Computation, Springer, pp. 1–19, 2008. https://link.springer.com/chapter/10.1007/978-3-540-79228-4_1

[14] B. Hitaj, G. Ateniese, and F. Perez-Cruz, 'Deep Models under the GAN: Information Leakage from Collaborative Deep Learning,' ACM SIGSAC Conf. on Computer and Communications Security, pp.603–618, 2017. https://dl.acm.org/doi/abs/10.1145/3133956.3134012

[15] A. Brodsky, C. Farkas, and S. Jajodia, 'Secure Databases: Constraints, Inference Channels, and Monitoring Disclosures,' IEEE Trans. on Knowledge and Data Engineering, vol. 12, no. 6, Nov./Dec. 2000, pp. 900-919.

[16] Wikipedia, 'Federated Learning', accessed May 2021, https://en.wikipedia.org/wiki/Federated_learning

[17] E. Chang, J. Han, W. Bae, et al.,'BAG2: A process-portable framework for generator-based AMS circuit design,' IEEE Custom Integrated Circuits Conference (CICC), April 2018.