Compliance Cards: Automated EU AI Act Compliance Analyses amidst a Complex AI Supply Chain

Bill Marino [1,2,*], Yaqub Chaudhary [2], Yulu Pi [2,4], Rui-Jie Yew [5]
Preslav Aleksandrov [1],
Carwyn Rahman [3], William F. Shen [1]
Isaac Robinson [6], Nicholas D. Lane [1]

[1] University of Cambridge, UK [2] Leverhulme Center for the Future of Intelligence, UK
[3] Protocol Consulting, Ltd., UK [4] University of Warwick, UK
[5] Brown University, US
[6] Oxford University, UK

Abstract

As the AI supply chain grows more complex, AI systems and models are increasingly likely to incorporate multiple internally- or externally-sourced components such as datasets and (pre-trained) models. In such cases, determining whether or not the aggregate AI system or model complies with the EU AI Act ("AIA") requires a multi-step process in which compliance-related information about both the AI system or model *and* all its component parts is: (1) gathered, potentially from multiple arms-length sources; (2) harmonized, if necessary; (3) inputted into an analysis that looks across all of it to render a compliance prediction. This complex and time-consuming process threatens to strain the limited compliance resources of the AI providers (i.e., developers) who bear much of the responsibility for complying with the AIA. It also renders rapid or real-time compliance analyses infeasible in many AI development scenarios where they would be beneficial. To address these shortcomings, we introduce a complete system for automating provider-side AIA compliance analyses amidst a complex AI supply chain. This system has two key elements. First is an interlocking set of computational, multi-stakeholder transparency artifacts that capture AIA-specific metadata about both: (1) the provider's overall AI system or model; and (2) the datasets and pre-trained models it incorporates as components. Second is an algorithm that operates across all those artifacts to render a real-time prediction about whether or not the aggregate AI system or model complies with the AIA. All told, this system promises to dramatically accelerate and democratize provider-side AIA compliance analyses (and, by extension, foster provider-side AIA compliance).

1 Introduction

The EU AI Act ("AIA"), which recently entered into force (56), sets forth harmonized requirements for AI systems and models (which we collectively call "AI projects") (28, Art. 1(2)).

Importantly, these requirements not only regulate aspects of the AI projects themselves, but also aspects of the datasets and (pre-trained) models that those AI projects incorporate as components. The practical impact of this, amidst an AI landscape where both AI projects (88) and the supply chain

[*]Corresponding Author: `wlm27@cam.ac.uk`

2nd Workshop on Regulatable ML at NeurIPS 2024.

behind them (14; 24) are growing more complex, is that determining whether or not an AI project complies with the AIA often requires the following multi-step procedure (Figure 1):

1. Gathering information about the AI project *and all its component datasets and models*, potentially from multiple internal or external (and sometimes arms-length) sources;
2. Optionally harmonizing that information (if its sources used different formats to capture it);
3. Waging an analysis that looks across all of this collected/harmonized information in order to render a prediction about the compliance level of the aggregate AI project.

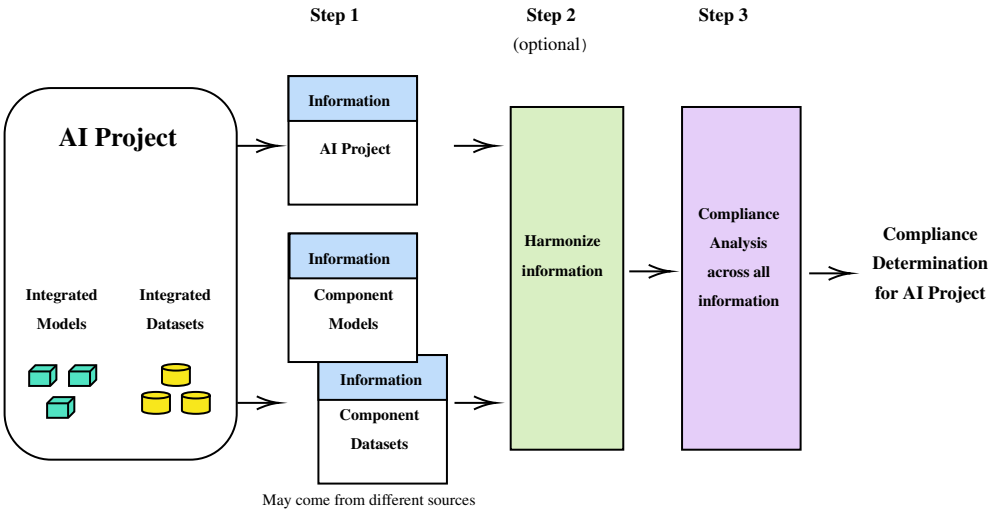


Figure 1: ***Today's AIA Compliance Analysis Procedure****: In Step 1, information about the overall AI Project as well as its component Models and Datasets is gathered from internal and/or external sources. In Step 2, this information is optionally harmonized. In Step 3, an analysis looks across all of the gathered/harmonized information to render a compliance analysis for the overall AI Project.*

This process may consume a lot of time and resources [2]. It therefore threatens to strain the limited compliance resources of the AI providers (i.e., developers) [3] who "bear the largest share of obligations under the [AIA]" (17). Importantly, it also renders real-time compliance assessments impossible in the many scenarios (listed in Section 2.1.2) where they would be highly beneficial to these providers.

To help solve these problems, we introduce the Compliance Cards system for automated provider-side AIA compliance analyses. It is built around two innovative technical elements:

- **Compliance Cards**: an interlocking set of open-source transparency artifacts (akin to Model or Data Cards (62; 68)) that capture, *in a computational format conductive to algorithmic manipulation*, compliance-related metadata about both: (1) a provider's AI project; and (2) any individual datasets and models it integrates;
- **Compliance Cards Algorithm**: an algorithm that manipulates the metadata in a set of Compliance Cards to render a run-time prediction about whether the provider's AI project complies with the AIA.

The vision is that the Compliance Cards artifacts for a given AI project (and its component datasets and models) will be populated, in advance and asynchronously, by the parties best-equipped to do so. Then, the artifacts can be assembled and inputted into the Compliance Cards Algorithm by

[2]For example, the AIA's Impact Assessment estimated that the process of "review[ing] documentation" to verify compliance for some AI systems would take up to "two and a half days" and cost up to €7500 (22, 68-69).

[3]The AIA defines several AI "operator" roles, with different obligations for each (28, Art. 3(8)). Among these, the "provider", which the AIA defines as anyone that develops an AI project (or that has an AI project developed) and places it on the market or puts it into service under their own name or trademark (28, Art. 3(3)), is awarded the highest level of responsibility for satisfying the AIA's requirements (28; 27, Arts. 2(1), 3(3)).

anyone who wishes to analyze whether the provider's AI project satisfies the AIA's requirements. Because the artifacts are pre-harmonized, our process eliminates the harmonization step in (Figure 1). Because they can be operated on by the accompanying algorithm, we also automate the analysis step in (Figure 1).[4] The net result is the streamlined AIA compliance analysis process depicted in (Figure 2), which promises to accelerate and democratize compliance assessments for providers, their third party assessors, regulators, or anyone else (see Section 4.3.1).

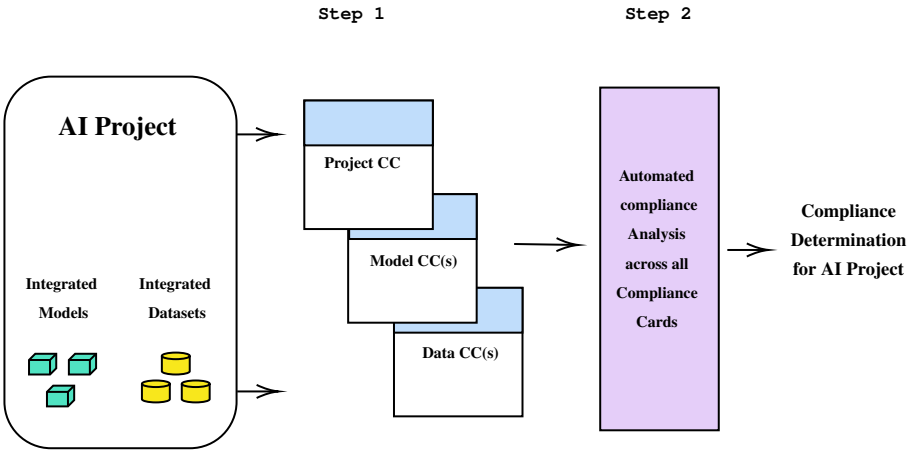


Figure 2: ***Compliance Cards AIA Compliance Analysis Procedure****: In Step 1, the Compliance Cards are used to collect, from internal and/or external sources, standardized metadata about the AI Project as well as its component Models and Datasets. In Step 2, an automated analysis looks across all of the gathered/harmonized metadata to render a compliance analysis for the overall AI Project.*

2 The Why and How of Compliance Cards

The Compliance Cards system is shaped by three key features of the modern AI landscape. First, the Compiance Cards' exact metadata is informed by AI's emerging regulatory dimension–namely, the AIA's dual focus on both project- and component-level features of an AI project. Second, the Compliance Cards system's specialized, separate artifacts are a response to the growing complexity of the AI supply chain faced by AI providers. Lastly, the growing need for real-time compliance analyses in various provider workflows inspired the Compliance Cards Algorithm and the complimentary, computational format of the artifacts.

2.1 The Dual Focus of the AIA's Requirements

A key feature of the AIA (and perhaps other upcoming AI regulation[5]) is that some of its requirements target an AI project at-large, while others target the datasets or models that the project incorporates as components. As an example, the AIA's record-keeping requirements target a project at-large (28, Art. 9, 12). But the AIA's data governance requirement mostly target the datasets that a project incorporates as components (28, Art. 10).

Notably, these sets of requirements are heavily interconnected. For example, the exact requirements that the dataset and model components must satisfy are a function of several high-level characteristics of the AI project[6], including — but not limited to — whether the AI project consists of an AI system or a general-purpose AI ("GPAI") model (28, Art. 1(2), 2, 6). Conversely, features of component

[4]The AIA itself champions the notion of an automated compliance analyses, albeit in relation to another type of AI operator: "The AI Office shall develop a template for a questionnaire, including through an automated tool, to facilitate deployers in complying with their obligations...in a simplified manner" (28, Art. 27(5)).

[5]The White House's Blueprint for an AI Bill of Rights, for example, also champions rules that are aimed at both AI systems and models at-large (80, 19), as well as their component datasets or models (80, 38).

[6]We call this special tranche of high-level characteristics, which play a key role in our Compliance Cards Algorithm, "dispositive characteristics."

models or datasets — such as the interpretation tools built into models — can shape what is required of the AI project at-large (e.g., the human oversight measures it must implement) (28, Art. 14).

The net effect of this duality — and this interconnectedness — is that providers cannot render a robust compliance analysis for an AI project without first gathering certain information about both the project at-large *and* its component datasets and models. To facilitate this, we introduce novel transparency artifacts — Compliance Cards, an evolution of earlier transparency artifacts such as Model Cards (62) — that are specifically designed to capture both kinds of essential information.

2.1.1 The growing complexity of the AI supply chain

Where an AI project is housed entirely inside one team, gathering all of the information required to render a compliance analysis — covering both the AI project at-large and any incorporated datasets or models — may be trivial. But, as the AIA itself notes, "[a]long the AI value chain multiple parties often supply ...components or processes that are incorporated" into AI projects (28, Rec. 88). These components, often sourced externally (5; 78; 19; 70), may include multiple[7] datasets and models obtained from open-source platforms like Hugging Face (64; 48) or third party models accessed via API (1). Notably, the provider integrating these components may lack insight into certain aspects of them (e.g., a model's training data) (53) and may never have direct contact with the supplier (10).

In such scenarios, it may prove difficult to have all of the different component suppliers (along with the provider) collectively fill out a single artifact capturing all the information required for a compliance analysis. What may be more feasible is to have each of the parties fill out an artifact strictly focused on their own contribution, in a decentralized fashion, then for these separate artifacts to be united in order to render the compliance analysis. The advantages of this approach include:

- Each party documents what they know best (i.e., their particular contribution to the aggregation), reaping the benefits of "division of knowledge" (37, 49);
- The artifacts can be filled-out (or updated) asynchronously, at each party's leisure;
- These artifacts can be efficiently shared in a one-to-many fashion (e.g., via Hugging Face);
- The parties do not need to be in contact in order for a compliance analysis to occur.[8]

To reap these benefits, the Compliance Cards system was designed with three separate (yet compatible) types of artifacts — Project, Data, and Model CCs — that can be filled-out asynchronously by the parties best equipped to do so, then combined to provide a holistic view of an AI project when it is time to run a compliance analysis.

2.1.2 The need for real-time AI regulation compliance analyses in various AI workflows

In a growing number of AI development scenarios, the ability to make real-time (or, at least, rapid) AI regulation compliance determinations would be highly advantageous for providers. These include:

Today's rapidly iterative AI development workflows Contemporary AI development is characterized by highly iterative experimentation that involves trying many different approaches — including with different models and datasets — to find the combination that performs best for the use case.(16; 86; 36; 66). During this phase of "continuous experimentation"(59), it would be beneficial for a provider to be able to check, in real-time, whether a candidate combination is regulation-compliant and, therefore, worth exploring.

Search and acquisition of components on AI communities and marketplaces Finding the right model or dataset to integrate into an AI project may involve searching through many candidates on ML communities (e.g., Hugging Face), model zoos, and marketplaces (18; 63). Such searches often involve multiple queries that are refined in real-time based on incoming search results (51). Real-time compliance analyses would allow providers, while searching for components, to instantly understand the compliance level of various candidate components as integrated into their overall AI project.

[7] For example, Apple's OpenELM LLM was trained on four datasets (61), including The Pile, which is itself an aggregation of 22 datasets (32).

[8] In this manner, Compliance Cards are "MultiActor Responsible AI (RAI) Artifacts" that can support direct or indirect communication between different stakeholders (50).

Federated learning Federated learning (FL) is a privacy-preserving ML technique in which models learn from data without directly accessing it (60). FL involves updating a global model locally, on private data, then sharing those updates (rather than the data) back to a central server for aggregation into the global model. In this scenario, the central server, so as not to slow model convergence, can ideally make rapid decisions about whether or not integrating a given batch of client-side data (via model updates) is AIA-compliant. Real-time compliance analyses would support this.

Continuous integration / continuous delivery There is great interest in automating the retraining of in-production AI models using production data, a process known as continuous integration and continuous delivery (33; 49). In such scenarios, it would be advantageous for a provider to know, in real-time, whether the integration of an incoming batch of new data (and/or the model retrained on that data), when combined with the overall AI project, are regulation-compliant.

The need for rapid or real-time compliance analyses in these AI development situations and others led to the design requirement that Compliance Cards must be readily manipulable by some accompanying algorithm in order to render low-latency compliance analyses. This, in turn, inspired the computational format of Compliance Cards — as well as the complimentary Compliance Cards Algorithm.

3 The Compliance Cards System

The Compliance Cards system has two key elements: (1) the Compliance Cards transparency artifacts; and (2) the Compliance Cards Algorithm. In this section, we take a closer look at each.

3.1 Compliance Cards

The Compliance Cards are an interlocking set of computational transparency artifacts that capture compliance-related metadata about both an AI project as well as its constituent datasets and models. Let's examine their overall format, then describe each type of card in more detail:

3.1.1 Compliance Cards format

A defining quality of Compliance Cards is that their metadata can be operated on by a Compliance Cards Algorithm to produce a compliance analysis. While there are different ways to embody this quality, they all follow the same overarching structure: a set of attribute-value pairs. These can inhabit many formats, including via human-readable data serialization language (e.g., YAML or JSON) or serialized python objects like dictionaries. In each of these, the attributes will capture specific characteristics of the subject (whether it is a project, model, or dataset) that relate to AIA compliance, while the values will quantify whether or how much the subject possesses that attribute. In the form we deem optimal, the latter are quantized — i.e., constrained to a finite set consisting of integers (e.g., confidence scores), boolean values, or other. Because it is the simplest format for computational manipulation, this is what we use in the initial, YAML-based Compliance Card templates[9] that are available at: `https://github.com/camlsys/compliancecards`.

3.1.2 Compliance Cards types

There are three types of Compliance Card ("CC") artifacts:

Project CC This type of Compliance Card captures AIA-relevant attributes of an AI project, including the "dispositive attributes" that heavily inform the compliance analysis. Table 2 presents an overview of the attributes this artifact captures. For a full list, see the `project_cc.yaml` template in the GitHub repository for this project: `https://github.com/camlsys/compliancecards`.

Data CC This type of Compliance Card captures AIA-relevant attributes of a component dataset (whether training, evaluation, or test set). Table 3 presents an overview of the attributes it captures.[10] For a full list, see the `data_cc.yaml` template in the GitHub repository for this project.

[9] Our templates are composed in YAML so that they can readily be added to the existing YAML metadata sections of the (popular) Hugging Face Model Cards, lowering the friction of adoption (40).

[10] Note that an AI project integrating multiple models or datasets will have a one-to-many relationship between its Project CC and/or Data and Model CCs.

Model CC This type of Compliance Card captures AIA-relevant attributes of a component model (including general-purpose AI models[11]). Table 4 presents an overview of the attributes it captures. For a full list, see the `model_cc.yaml` template in the GitHub repository for this project.

3.2 Compliance Cards Algorithm

The Compliance Cards Algorithm looks across all of the metadata in the Compliance Cards to render a real-time prediction about whether the aggregate AI project complies with the AIA (Figure 3.2).

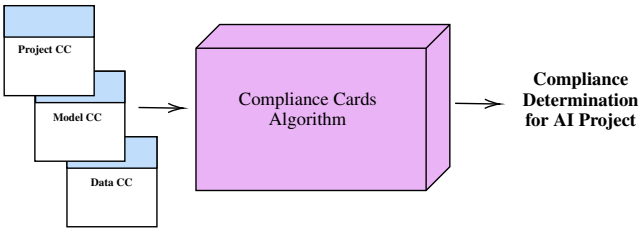


Figure 3: ***The Compliance Cards Algorithm****: The Compliance Cards Algorithm is any algorithm that accepts the Compliance Cards as its input and outputs a compliance analysis.*

While the Compliance Cards Algorithm can take on many different embodiments, the initial embodiment that we put forward in conjunction with this paper is rules-based and uses a series of if-then statements to represent and encode the requirements of the the AIA. A Python implementation of this algorithm is available at the Compliance Cards GitHub repository. Importantly, other embodiments are not only possible but may scale better to unseen AI projects. In Section 6, for example, we express our intent to produce an LLM-assisted embodiment of the Compliance Cards Algorithm.

4 How to Use the Compliance Cards System

How should the Compliance Cards system actually be used in practice? Here are some ideas:

4.1 Who should populate the Compliance Cards?

In light of the information asymmetry that exists along the AI supply chain (10), Compliance Cards are ideally populated by whichever entity has the best information about the subject of the respective card. As a rule of thumb, we recommend the entities in Table 1 populate each artifact.[12]

Table 1: Who should populate the Compliance Card?

Artifact	Recommended populator
Project CC	Provider
Data CC	Dataset Supplier(s)
Model CC	Model Supplier(s)

4.2 How should the Compliance Cards be populated?

While parties can populate their Compliance Cards any way they like, here are some methods that we expect to be popular:

- **Manual population via user interface (UI)**: Compliance Cards can be populated manually through a questionnaire-style UI akin to those used to populate existing artifacts (44). In the

[11]The AIA highlights that GPAI models "can be integrated as a component of a downstream AI system covered by the AIA, but also...can be used 'as such' as an AI system falling under the scope of the [AIA]" (30).

[12]Note that the AIA actively encourages AI component makers to provide documentation to integrators of those components — something we argue is achieved by sharing filled-out Compliance Cards (28, Rec. 88, 89).

case of Compliance Cards, a UI like this might present each individual attribute to the user, asking them to enter the appropriate value for that attribute.

- **Programmatic population via library**: Like some existing transparency artifacts, Compliance Cards could hypothetically be populated programmatically, via a library that is referenced from within model training, evaluation, and deployment code (43; 69; 29). In Section 6, we express our intent to release a similar library for Compliance Cards.
- **Population via large language model (LLM)**: Text documents, including but not limited to training and evaluation code, data governance policies, or other artifacts such as Model and Data Cards, could be inputted into a specialized LLM that outputs one or more populated Compliance Cards. In Section 6, we express our intent to release a toolkit that allows users to leverage existing LLMs to populate Compliance Cards in this manner.

4.3 What happens to Compliance Cards after they are populated?

Our hope is that Data and Model CCs, like existing transparency artifacts, will be uploaded to dataset and model repositories on ML communities and marketplaces (e.g., Hugging Face). There, they can be readily accessed by providers or other parties (discussed in Section 4.3.1) who may wish to combine them with a Project CC and/or other Data and Model CCs, then render a compliance analysis for a potential or existing AI project. Alternatively, the suppliers of datasets and models may wish to generate these artifacts and keep them on-hand for private transmission to such parties.

Because ML communities and marketplaces are not yet, as far as we know, instrumented to host Project CCs in the same way they can host Data and Model CCs, we expect that the Project CC for a given AI project is more likely to be generated by an AI provider and then kept locally until that provider (or any of the parties listed in Section 4.3.1) wishes to pair it with any relevant Data and Model CCs and render a compliance analysis for the project.

4.3.1 Who will run the Compliance Cards Algorithm on the Compliance Cards?

In short, anyone with the full set of Compliance Cards associated with an AI project can run the Compliance Cards Algorithm on them and output a compliance analysis. That said, here are some users and use cases we expect to be commonplace:

- **Conformity assessors**: The AIA requires AI providers, before putting certain AI projects on the EU market, to undergo a conformity assessment — performed internally or by a third party — verifying the project's compliance (28, Art. 16, 43). We believe the Compliance Cards system can accelerate these assessments, regardless of who performs them.
- **Market surveillance entities**: Under the AIA, market surveillance authorities have the power to investigate AIA non-compliance (28, Art. 75). We believe the Compliance Cards system can accelerate such investigations as regards providers.[13]
- **Providers**: Because they bear the most responsibility for complying with the AIA (28; 27, Arts. 2(1), 3(3)), and also own the development workflows (Section 2.1.2) that benefit the most from accelerated compliance analyses, we foresee AI providers making our system a key part of the internal assessments that are "the core of [AIA] compliance" (57; 81).
- **Dataset and model suppliers**: We envision a scenario in which model and dataset suppliers use the Compliance Cards system to evaluate the compliance level of their wares when paired with various providers' Project CCs, using the output to inform everything from their future development plans to their marketing strategy.

4.3.2 How will they run the Compliance Cards Algorithm on the Compliance Cards?

Our Python implementation of the Compliance Cards Algorithm is open source and can be used by anyone interested in rendering a compliance analysis for an AI project. Our hope is that the algorithm will be integrated into various applications or platforms, made available to their users to democratize its use. In the meantime, we also plan to host the algorithm in a web application that anyone can use to run analyses across a set of Compliance Cards they generate and/or upload.

[13] AI regulators may suffer from a comparative lack of capacity, resources, and expertise (7). By helping overcome this, the Compliance Cards may help boost AIA enforcement.

5 Related Work

Researchers have developed various transparency artifacts to facilitate the sharing of metadata about models, datasets, and other facets of AI systems. Some of these artifacts have been popularized by industry. Most recently, there have been efforts to produce transparency artifacts that are machine-readable as well as efforts to produce transparency artifacts specialized to AI regulation, including the AIA. In parallel, there has been work on algorithms that encode the requirements of the AIA. What is still missing — and what Compliance Cards provides — is a holistic system that brings together all the necessary building blocks for real-time compliance analyses by pairing computational, AIA-specialized transparency artifacts with a connected, AIA-encoding algorithm.

5.1 Human- and machine-readable transparency artifacts

Multiple efforts have produced human-readable transparency artifacts that facilitate the sharing of metadata about AI components. This includes artifacts for datasets (11; 34; 62; 38; 21; 46; 73), models (62; 29; 79; 74; 75; 2; 47), services (8), and multi-model systems (67). Recognizing the need for transparency artifacts that were more readily analyzable by computers, researchers have recently captured various transparency artifacts in machine-readable formats like the OWL 2 Web Ontology Language (6), linked data and knowledge graphs (23), YAML (12), and JSON (72). Perhaps most notable due to their widespread use[14], Hugging Face Model Cards are captured as machine-readable Markdown files with optional YAML sections (41; 40), a format that enabled a large-scale analysis of these Model Cards' contents by Liang et al. (52).

5.2 Transparency artifacts for AI regulation (including AIA) compliance

Although the transparency artifacts above capture some of the metadata needed for an AIA compliance determination (15), they do not capture all of it. To address this gap, recent efforts have introduced transparency artifacts specialized to the AIA compliance (or, at least, AI regulation compliance) use case. In particular, Brajovic et al. (13) introduced human-readable transparency artifacts to help align AI systems with AI regulations during development while Hupont et al. (45) introduced human-readable Use Case Cards to capture metadata related to AIA's risk levels. Lohachab and Urovi (55) rendered model artifacts as blockchain-based non-fungible tokens to, among other things, advance regulatory compliance. Most recently, Golpayegani et al. (35) produced a transparency artifact that captures metadata related to the corner of the AIA addressing AI risk management technical documentation. Importantly, none of these efforts include a comprehensive (and certainly not automated) method for looking across their metadata to produce a compliance analysis.

5.3 AIA-encoding Algorithms

RegCheckAI is a web application that renders an AIA compliance prediction based solely on the metadata captured in a Hugging Face Model Card — an approach we argue is incomplete (39). Meanwhile, Future of Life Institute (31) offers a web application that predicts the AIA requirements that may apply to AI project based on its inputted details; importantly, it does not analyze whether the project satisfies those requirements. A handful of startups also offer web applications advertising similar functionality, sometimes behind registration walls (82; 3). While the algorithms undergirding these web applications are not public, we infer that they include some type of encoding of the AIA's rules (or, at least, a portion of them).

5.4 The void that remains (and Compliance Cards fills)

What has been sorely missing from existing research is a holistic system that interweaves all of the the necessary building blocks for real-time AIA compliance analyses: (1) Computation-amenable artifacts that captures all of the metadata necessary for a robust AIA compliance analysis; (2) A companion algorithm that operates on those artifacts to output a 360 AIA compliance analysis. The Compliance Cards system is meant to fill this void.

[14] Over 32,000 Model Cards and 7000 Dataset Cards have been filled-out and shared on Hugging Face (52; 87).

6 Future Work

There are many opportunities to extend this line of research, with the Compliance Cards system as a foundation:

6.1 Evaluation

It would be useful to thoroughly evaluate the accuracy, speed, and expense of the Compliance Cards system as compared to the status quo (the manual compliance analyses depicted in Figure 1).

6.2 Automate the process of Compliance Card population

Automating the population of Compliance Cards will further streamline the process of AIA compliance analyses. There are multiple paths towards this goal:

- Values could be programmatically deposited into Compliance Cards during the training and evaluation processes. This could include the outputs of existing responsible AI benchmarking frameworks (54; 65). This process could be managed by a library that bears similarity to the existing artifact-populating libraries like the Model Cards Toolkit of (29; 79; 4).
- Training, evaluation, and product code, product specifications, existing transparency artifacts, and other documentation can be inputted into an LLM that, perhaps via contextual learning, outputs populated Compliance Cards.

6.3 Provide tools to verify Compliance Cards metadata

Sometimes, component models or datasets will not be accessible to their integrators or would-be integrators; this may happen wherever these components are closed by design (9), before their exchange during arms-length business transactions (84), or in federated learning (85). In such scenarios, it would nonetheless be valuable for model and dataset integrators to be able to verify that the metadata shared in Model and Data Compliance Cards is accurate (58; 71). There has been promising work using zk-SNARK zero-knowledge proofs to verify ML model accuracy and fairness without direct model access (77), though it may come with high latency. Differently, Hugging Face has unveiled a feature that lets model makers run evaluations on Hugging Face-controlled servers and receive metrics, digitally signed by Hugging Face, to add to their Model Cards (42). It would be valuable to explore how these or other closed-box methods can be used to verify the metadata captured in Compliance Cards (in a cost-effective and rapid manner).

6.4 Adapt Compliance Cards to AIA technical standards and other AI regulation

As is typical in EU lawmaking, multiple EU standardization bodies have been asked to create harmonized standards ("common specifications") to accompany the AIA (26; 25; 83). AI projects that conform to these "more concrete" (76) requirements will be presumed to be AIA-compliant (28, 41(3)). Once issued, there could be value in replacing some or all of the existing Compliance Card attributes with ones that reflect these more granular specifications. In addition, other AI regulations are being gestated worldwide (20). The Compliance Cards framework could be extended beyond the AIA and to these other regulations as they materialize.

7 Conclusion

The Compliance Card system is designed to reduce the burden of AIA compliance analyses that firms face admidst an increasingly complex AI supply chain. Our hope (and hunch) is that lowering the capacity required to run AIA compliance analyses will give AI firms (especially resource-starved small- and medium-sized businesses) more bandwidth to cure any of the deficiencies these analyses uncover or to investigate nuanced issues that an automated analysis may not fully capture. Accordingly, we see Compliance Cards as a pathway to increased AIA compliance and, since the AIA encodes so many core responsible AI principles, a step forward for responsible AI in general.

References

[1] Ada Lovelace Institute. What is a foundation model? https://www.adalovelaceinstitute.org/resource/foundation-models-explainer/, 2023. [Accessed 23-08-2024].

[2] David Adkins, Bilal Alsallakh, Adeel Cheema, Narine Kokhlikyan, Emily McReynolds, Pushkar Mishra, Chavez Procope, Jeremy Sawruk, Erin Wang, and Polina Zvyagina. Prescriptive and descriptive approaches to machine-learning transparency. *CoRR*, abs/2204.13582, 2022. doi: 10.48550/ARXIV.2204.13582. URL https://doi.org/10.48550/arXiv.2204.13582.

[3] Holistic AI. EU AI Act risk calculator. https://www.holisticai.com/eu-ai-act-risk-calculator, 2024. [Accessed 23-08-2024].

[4] Amazon SageMaker. Model card FAQs. https://docs.aws.amazon.com/sagemaker/latest/dg/model-cards-faqs.html, 2024. [Accessed 22-08-2024].

[5] Saleema Amershi, Andrew Begel, Christian Bird, Robert DeLine, Harald C. Gall, Ece Kamar, Nachiappan Nagappan, Besmira Nushi, and Thomas Zimmermann. Software engineering for machine learning: A case study. In Helen Sharp and Mike Whalen, editors, *Proceedings of the 41st International Conference on Software Engineering: Software Engineering in Practice, ICSE (SEIP) 2019, Montreal, QC, Canada, May 25-31, 2019*, pages 291–300. IEEE / ACM, 2019. doi: 10.1109/ICSE-SEIP.2019.00042. URL https://doi.org/10.1109/ICSE-SEIP.2019.00042.

[6] Muhammad Tuan Amith, Licong Cui, Degui Zhi, Kirk Roberts, Xiaoqian Jiang, Fang Li, Evan Yu, and Cui Tao. Toward a standard formal semantic representation of the model card report, 2022. URL https://doi.org/10.1186/s12859-022-04797-6.

[7] Usman Anwar, Abulhair Saparov, Javier Rando, Daniel Paleka, Miles Turpin, Peter Hase, Ekdeep Singh Lubana, Erik Jenner, Stephen Casper, Oliver Sourbut, Benjamin L. Edelman, Zhaowei Zhang, Mario Günther, Anton Korinek, José Hernández-Orallo, Lewis Hammond, Eric J. Bigelow, Alexander Pan, Lauro Langosco, Tomasz Korbak, Heidi Zhang, Ruiqi Zhong, Seán Ó hÉigeartaigh, Gabriel Recchia, Giulio Corsi, Alan Chan, Markus Anderljung, Lilian Edwards, Yoshua Bengio, Danqi Chen, Samuel Albanie, Tegan Maharaj, Jakob N. Foerster, Florian Tramèr, He He, Atoosa Kasirzadeh, Yejin Choi, and David Krueger. Foundational challenges in assuring alignment and safety of large language models. *CoRR*, abs/2404.09932, 2024. doi: 10.48550/ARXIV.2404.09932. URL https://doi.org/10.48550/arXiv.2404.09932.

[8] Matthew Arnold, Rachel K. E. Bellamy, Michael Hind, Stephanie Houde, Sameep Mehta, Aleksandra Mojsilovic, Ravi Nair, Karthikeyan Natesan Ramamurthy, Alexandra Olteanu, David Piorkowski, Darrell Reimer, John T. Richards, Jason Tsay, and Kush R. Varshney. Factsheets: Increasing trust in AI services through supplier's declarations of conformity. *IBM J. Res. Dev.*, 63(4/5):6:1–6:13, 2019. doi: 10.1147/JRD.2019.2942288. URL https://doi.org/10.1147/JRD.2019.2942288.

[9] Simone Balloccu, Patrícia Schmidtová, Mateusz Lango, and Ondrej Dusek. Leak, cheat, repeat: Data contamination and evaluation malpractices in closed-source LLMs. In Yvette Graham and Matthew Purver, editors, *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics, EACL 2024 - Volume 1: Long Papers, St. Julian's, Malta, March 17-22, 2024*, pages 67–93. Association for Computational Linguistics, 2024. URL https://aclanthology.org/2024.eacl-long.5.

[10] Iain Barclay, Alun D. Preece, Ian J. Taylor, and Dinesh C. Verma. Quantifying transparency of machine learning systems through analysis of contributions. *CoRR*, abs/1907.03483, 2019. URL http://arxiv.org/abs/1907.03483.

[11] Emily M. Bender and Batya Friedman. Data statements for natural language processing: Toward mitigating system bias and enabling better science. *Trans. Assoc. Comput. Linguistics*, 6:587–604, 2018. doi: 10.1162/TACL_A_00041. URL https://doi.org/10.1162/tacl_a_00041.

[12] Thomas M Booth and Sudipto Ghosh. Machine learning model cards toward model-based system engineering analysis of resource-limited systems. In *Signal Processing, Sensor/Information Fusion, and Target Recognition XXXII*, volume 12547, pages 188–203. SPIE, 2023.

[13] Danilo Brajovic, Niclas Renner, Vincent Philipp Goebels, Philipp Wagner, Benjamin Frész, Martin Biller, Mara Klaeb, Janika Kutz, Jens Neuhüttler, and Marco F. Huber. Model reporting for certifiable AI: A proposal from merging EU regulation into AI development. *CoRR*, abs/2307.11525, 2023. doi: 10.48550/ARXIV.2307.11525. URL https://doi.org/10.48550/arXiv.2307.11525.

[14] Ian Brown. Allocating accountability in AI supply chains. https://www.adalovelaceinstitute.org/resource/ai-supply-chains/, 2023. [Accessed 23-08-2024].

[15] Diego Calanzone, Andrea Coppari, Riccardo Tedoldi, Giulia Olivato, and Carlo Casonato. An open source perspective on AI and alignment with the EU AI act. In Gabriel Pedroza, Xiaowei Huang, Xin Cynthia Chen, Andreas Theodorou, Huáscar Espinoza, Nikolaos Matragkas, José Hernández-Orallo, Mauricio Castillo-Effen, Richard Mallah, John A. McDermid, David M. Bossens, Bettina Könighofer, Sebastian Tschiatschek, and Anqi Liu, editors, *Proceedings of the IJCAI-23 Joint Workshop on Artificial Intelligence Safety and Safe Reinforcement Learning (AISafety-SafeRL 2023) co-located with the 32nd International Joint Conference on Artificial Intelligence(IJCAI2023), Macau, China, August 21-22, 2023*, volume 3505 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2023. URL https://ceur-ws.org/Vol-3505/paper_10.pdf.

[16] Nicholas Carlini. Rapid iteration in machine learning research. https://nicholas.carlini.com/writing/2022/rapid-iteration-machine-learning-research.html, 2022. [Accessed 23-08-2024].

[17] Center for Democracy & Technology. Public authorities: What role in the AI Act? https://cdt.org/insights/public-authorities-what-role-in-the-ai-act, 2024. [Accessed 23-08-2024].

[18] Adriane Chapman, Elena Simperl, Laura Koesten, George Konstantinidis, Luis-Daniel Ibáñez, Emilia Kacprzak, and Paul Groth. Dataset search: A survey. *VLDB J.*, 29(1):251–272, 2020. doi: 10.1007/S00778-019-00564-X. URL https://doi.org/10.1007/s00778-019-00564-x.

[19] Shamik Chaudhuri, Kingshuk Dasgupta, Michael Le Isaac Hepworth, Mark Lodato, Mihai Maruseac, Sarah Meiklejohn, Tehila Minkus, and Kara Olive. Securing the AI software supply chain. https://research.google/pubs/securing-the-ai-software-supply-chain/, 2024. [Accessed 23-08-2024].

[20] Foo Yun Chee and Tassilo Hummel. Europe sets benchmark for rest of the world with landmark AI laws. https://www.reuters.com/world/europe/eu-countries-back-landmark-artificial-intelligence-rules-2024-05-21/, 2024. [Accessed 23-08-2024].

[21] Kasia S. Chmielinski, Sarah Newman, Matt Taylor, Josh Joseph, Kemi Thomas, Jessica Yurkofsky, and Yue Chelsea Qiu. The Dataset Nutrition Label (2nd gen): Leveraging context to mitigate harms in artificial intelligence. *CoRR*, abs/2201.03954, 2022. URL https://arxiv.org/abs/2201.03954.

[22] EUROPEAN COMMISSION. Commission staff working document impact assessment accompanying the proposal for a regulation of the european parliament and of the council laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts. https://ec.europa.eu/newsroom/dae/redirection/document/75792, 2021. [Accessed 28-08-2024].

[23] Andy Donald, Edward Curry Apostolos Galanopoulos, Emir Mu noz, Ihsan Ullah, M.A. Waskow, Maciej Dabrowski, and Manan Kalra. Towards a semantic approach for linked dataspace, model and data cards. In *Companion Proceedings of the ACM Web Conference 2023*, WWW '23 Companion, page 1468–1473, New York, NY, USA, 2023. Association

for Computing Machinery. ISBN 9781450394192. doi: 10.1145/3543873.3587659. URL https://doi.org/10.1145/3543873.3587659.

[24] Alex Engler and Andrea Renda. Reconciling the AI value chain with the EU's Artificial Intelligence Act. https://www.ceps.eu/ceps-publications/reconciling-the-ai-value-chain-with-the-eus-artificial-intelligence-act/, 2022. [Accessed 23-08-2024].

[25] European Commission. Draft standardisation request to the European Standardisation Organisations in support of safe and trustworthy artificial intelligence. https://ec.europa.eu/docsroom/documents/52376, 2022. [Accessed 23-08-2024].

[26] European Committee for Electrotechnical Standardization. Artificial intelligence. https://www.cencenelec.eu/areas-of-work/cen-cenelec-topics/artificial-intelligence/, 2024. [Accessed 23-08-2024].

[27] European Digital Rights. Introduce obligations on users of high-risk ai systems. https://edri.org/wp-content/uploads/2022/05/Obligations-on-users-AIA-Amendments-17022022.pdf, 2022. [Accessed 23-08-2024].

[28] European Union. Regulation (eu) 2024/1689 of the european parliament and of the council of 13 june 2024 laying down harmonised rules on artificial intelligence and amending regulations (ec) no 300/2008, (eu) no 167/2013, (eu) no 168/2013, (eu) 2018/858, (eu) 2018/1139 and (eu) 2019/2144 and directives 2014/90/eu, (eu) 2016/797 and (eu) 2020/1828 (artificial intelligence act)text with eea relevance. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L_202401689, 2024. [Accessed 23-08-2024].

[29] Huanming Fang and Hui Miao. Introducing the Model Card Toolkit for easier model transparency reporting. https://blog.research.google/2020/07/introducing-model-card-toolkit-for.html, 2020. [Accessed 23-08-2024].

[30] David Fernández-Llorca, Emilia Gómez, Ignacio Sánchez, and Gabriele Mazzini. An interdisciplinary account of the terminological choices by eu policymakers ahead of the final agreement on the ai act: Ai system, general purpose ai system, foundation model, and generative ai. *Artificial Intelligence and Law*, Aug 2024. ISSN 1572-8382. doi: 10.1007/s10506-024-09412-y. URL https://doi.org/10.1007/s10506-024-09412-y.

[31] Future of Life Institute. EU AI Act Compliance Checker. https://artificialintelligenceact.eu/assessment/eu-ai-act-compliance-checker/, 2023. [Accessed 21-08-2024].

[32] Leo Gao, Stella Biderman, Sid Black, Laurence Golding, Travis Hoppe, Charles Foster, Jason Phang, Horace He, Anish Thite, Noa Nabeshima, Shawn Presser, and Connor Leahy. The pile: An 800gb dataset of diverse text for language modeling, 2020. URL https://arxiv.org/abs/2101.00027.

[33] Satvik Garg, Pradyumn Pundir, Geetanjali Rathee, P. K. Gupta, Somya Garg, and Saransh Ahlawat. On continuous integration / continuous delivery for automated deployment of machine learning models using MLOps. *CoRR*, abs/2202.03541, 2022. URL https://arxiv.org/abs/2202.03541.

[34] Timnit Gebru, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé III, and Kate Crawford. Datasheets for datasets. *Commun. ACM*, 64(12): 86–92, nov 2021. ISSN 0001-0782. doi: 10.1145/3458723. URL https://doi.org/10.1145/3458723.

[35] Delaram Golpayegani, Isabelle Hupont, Cecilia Panigutti, Harshvardhan J. Pandit, Sven Schade, Declan O'Sullivan, and Dave Lewis. AI Cards: Towards an applied framework for machine-readable AI and risk documentation inspired by the EU AI Act, 2024. URL https://arxiv.org/abs/2406.18211.

[36] Grace Guo, Dustin Arendt, and Alex Endert. Explainability in JupyterLab and beyond: Interactive XAI systems for integrated and collaborative workflows. *CoRR*, abs/2404.02081, 2024. doi: 10.48550/ARXIV.2404.02081. URL https://doi.org/10.48550/arXiv.2404.02081.

[37] Friedrich August Hayek. Economics and knowledge. *Economica*, 4(13):33–54, 1937. ISSN 00130427, 14680335. URL http://www.jstor.org/stable/2548786.

[38] Sarah Holland, Ahmed Hosny, Sarah Newman, Joshua Joseph, and Kasia Chmielinski. The Dataset Nutrition Label: A framework to drive higher data quality standards, 2018. URL https://api.semanticscholar.org/CorpusID:13665181.

[39] Hugging Face. Model Card Regulatory Check. https://huggingface.co/spaces/society-ethics/model-card-regulatory-check, 2023. [Accessed 23-08-2024].

[40] Hugging Face. Example metadata to be added to a model card. https://github.com/huggingface/hub-docs/blob/main/modelcard.md, 2024. [Accessed 23-07-2024].

[41] Hugging Face. Model Card Template. https://github.com/huggingface/huggingface_hub/blob/main/src/huggingface_hub/templates/modelcard_template.md, 2024. [Accessed 23-08-2024].

[42] Hugging Face. facebook/bart-large-cnn · Add verifyToken field to verify evaluation results are produced by Hugging Face's automatic model evaluator — huggingface.co. https://huggingface.co/facebook/bart-large-cnn/discussions/23, 2024. [Accessed 23-08-2024].

[43] Hugging Face. Model Cards. https://huggingface.co/docs/hub/en/model-cards, 2024. [Accessed 23-08-2024].

[44] Hugging Face. Modelcard Creator. https://huggingface.co/spaces/huggingface/Model_Cards_Writing_Tool, 2024. [Accessed 23-08-2024].

[45] Isabelle Hupont, David Fernández Llorca, Sandra Baldassarri, and Emilia Gómez. Use case cards: a use case reporting framework inspired by the European AI act. *Ethics Inf. Technol.*, 26(2):19, 2024. doi: 10.1007/S10676-024-09757-7. URL https://doi.org/10.1007/s10676-024-09757-7.

[46] Ben Hutchinson, Andrew Smart, A. Hanna, Emily L. Denton, Christina Greer, Oddur Kjartansson, Parker Barnes, and Margaret Mitchell. Towards accountability for machine learning datasets: Practices from software engineering and infrastructure, 2020. URL https://api.semanticscholar.org/CorpusID:225067460.

[47] IBM. Using AI Factsheets for AI Governance | IBM Cloud Pak for Data as a Service. https://dataplatform.cloud.ibm.com/docs/content/wsj/analyze-data/factsheets-model-inventory.html, 2024. [Accessed 23-08-2024].

[48] Jason Jones, Wenxin Jiang, Nicholas Synovic, George K. Thiruvathukal, and James C. Davis. What do we know about Hugging Face? A systematic literature review and quantitative validation of qualitative claims. *CoRR*, abs/2406.08205, 2024. doi: 10.48550/ARXIV.2406.08205. URL https://doi.org/10.48550/arXiv.2406.08205.

[49] Bojan Karlas, Matteo Interlandi, Cédric Renggli, Wentao Wu, Ce Zhang, Deepak Mukunthu Iyappan Babu, Jordan Edwards, Chris Lauren, Andy Xu, and Markus Weimer. Building continuous integration services for machine learning. In Rajesh Gupta, Yan Liu, Jiliang Tang, and B. Aditya Prakash, editors, *KDD '20: The 26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Virtual Event, CA, USA, August 23-27, 2020*, pages 2407–2415. ACM, 2020. doi: 10.1145/3394486.3403290. URL https://doi.org/10.1145/3394486.3403290.

[50] Anna Kawakami, Daricia Wilkinson, and Alexandra Chouldechova. Do responsible ai artifacts advance stakeholder goals? four key barriers perceived by legal and civil stakeholders, 2024. URL https://arxiv.org/abs/2408.12047.

[51] Jin Young Kim, Mark Cramer, Jaime Teevan, and Dmitry Lagun. Understanding how people interact with web search results that change in real-time using implicit feedback. In Qi He, Arun Iyengar, Wolfgang Nejdl, Jian Pei, and Rajeev Rastogi, editors, *22nd ACM International Conference on Information and Knowledge Management, CIKM'13, San Francisco, CA, USA, October 27 - November 1, 2013*, pages 2321–2326. ACM, 2013. doi: 10.1145/2505515.2505663. URL https://doi.org/10.1145/2505515.2505663.

[52] Weixin Liang, Nazneen Rajani, Xinyu Yang, Ezinwanne Ozoani, Eric Wu, Yiqun Chen, Daniel Scott Smith, and James Zou. What's documented in AI? Systematic analysis of 32k AI model cards, 2024.

[53] Andreas Liesenfeld and Mark Dingemanse. Rethinking open source generative AI: open washing and the EU AI act. In *The 2024 ACM Conference on Fairness, Accountability, and Transparency, FAccT 2024, Rio de Janeiro, Brazil, June 3-6, 2024*, pages 1774–1787. ACM, 2024. doi: 10.1145/3630106.3659005. URL https://doi.org/10.1145/3630106.3659005.

[54] Stephanie Lin, Jacob Hilton, and Owain Evans. Truthfulqa: Measuring how models mimic human falsehoods. In Smaranda Muresan, Preslav Nakov, and Aline Villavicencio, editors, *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2022, Dublin, Ireland, May 22-27, 2022*, pages 3214–3252. Association for Computational Linguistics, 2022. doi: 10.18653/V1/2022.ACL-LONG.229. URL https://doi.org/10.18653/v1/2022.acl-long.229.

[55] Ankur Lohachab and Visara Urovi. A blockchain-based approach for model card accountability and regulatory compliance. In João Paulo A. Almeida, Claudio Di Ciccio, and Christos Kalloniatis, editors, *Advanced Information Systems Engineering Workshops - CAiSE 2024 International Workshops, Limassol, Cyprus, June 3-7, 2024, Proceedings*, volume 521 of *Lecture Notes in Business Information Processing*, pages 37–48. Springer, 2024. doi: 10.1007/978-3-031-61003-5_4. URL https://doi.org/10.1007/978-3-031-61003-5_4.

[56] Natasha Lomas. EU's AI Act gets published in bloc's Official Journal, starting clock on legal deadlines | TechCrunch. https://techcrunch.com/2024/07/12/eus-ai-act-gets-published-in-blocs-official-journal-starting-clock-on-legal-deadlines, 2024. [Accessed 23-08-2024].

[57] Laura Lucaj, Patrick van der Smagt, and Djalel Benbouzid. AI regulation is (not) all you need. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency, FAccT 2023, Chicago, IL, USA, June 12-15, 2023*, pages 1267–1279. ACM, 2023. doi: 10.1145/3593013.3594079. URL https://doi.org/10.1145/3593013.3594079.

[58] Bill Marino. The EU AI Act's technical "tension areas". https://www.lcfi.ac.uk/news-events/blog/post/the-eu-ai-acts-technical-tension-areas, 2024. [Accessed 23-08-2024].

[59] Silverio Martínez-Fernández, Justus Bogner, Xavier Franch, Marc Oriol, Julien Siebert, Adam Trendowicz, Anna Maria Vollmer, and Stefan Wagner. Software engineering for ai-based systems: A survey. *ACM Trans. Softw. Eng. Methodol.*, 31(2):37e:1–37e:59, 2022. doi: 10.1145/3487043. URL https://doi.org/10.1145/3487043.

[60] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Agüera y Arcas. Communication-efficient learning of deep networks from decentralized data. In Aarti Singh and Xiaojin (Jerry) Zhu, editors, *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, AISTATS 2017, 20-22 April 2017, Fort Lauderdale, FL, USA*, volume 54 of *Proceedings of Machine Learning Research*, pages 1273–1282. PMLR, 2017. URL http://proceedings.mlr.press/v54/mcmahan17a.html.

[61] Sachin Mehta, Mohammad Hossein Sekhavat, Qingqing Cao, Maxwell Horton, Yanzi Jin, Chenfan Sun, Seyed-Iman Mirzadeh, Mahyar Najibi, Dmitry Belenko, Peter Zatloukal, and Mohammad Rastegari. Openelm: An efficient language model family with open training and inference framework. *CoRR*, abs/2404.14619, 2024. doi: 10.48550/ARXIV.2404.14619. URL https://doi.org/10.48550/arXiv.2404.14619.

[62] Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timnit Gebru. Model cards for model reporting, 2019. URL https://doi.org/10.1145/3287560.3287596.

[63] Zachary Mueller. Supercharged searching on the Hugging Face Hub. https://huggingface.co/blog/searching-the-hub, 2022. [Accessed 23-08-2024].

[64] Cailean Osborne, Jennifer Ding, and Hannah Rose Kirk. The AI community building the future? A quantitative analysis of development activity on Hugging Face Hub. *CoRR*, abs/2405.13058, 2024. doi: 10.48550/ARXIV.2405.13058. URL https://doi.org/10.48550/arXiv.2405.13058.

[65] Alicia Parrish, Angelica Chen, Nikita Nangia, Vishakh Padmakumar, Jason Phang, Jana Thompson, Phu Mon Htut, and Samuel R. Bowman. BBQ: A hand-built bias benchmark for question answering. In Smaranda Muresan, Preslav Nakov, and Aline Villavicencio, editors, *Findings of the Association for Computational Linguistics: ACL 2022, Dublin, Ireland, May 22-27, 2022*, pages 2086–2105. Association for Computational Linguistics, 2022. doi: 10.18653/V1/2022.FINDINGS-ACL.165. URL https://doi.org/10.18653/v1/2022.findings-acl.165.

[66] David Piorkowski, John T. Richards, and Michael Hind. Evaluating a methodology for increasing AI transparency: A case study. *CoRR*, abs/2201.13224, 2022. URL https://arxiv.org/abs/2201.13224.

[67] Chavez Procope, Adeel Cheema, David Adkins, Bilal Alsallakh, Nekesha Green, Emily McReynolds, Grace Pehl, Erin Wang, and Polina Zvyagina. System-level transparency of machine Learning. https://ai.meta.com/research/publications/system-level-transparency-of-machine-learning/, 2022. [Accessed 23-8-2024].

[68] Mahima Pushkarna, Andrew Zaldivar, and Oddur Kjartansson. Data cards: Purposeful and transparent dataset documentation for responsible ai, 2022. URL https://doi.org/10.1145/3531146.3533231.

[69] PyPI. Model Card Toolkit. https://pypi.org/project/model-card-toolkit/, 2024. [Accessed 23-08-2024].

[70] Elizabeth M. Renieris, David Kiron, and Steven Mills. Building robust RAI programs as third-party AI tools proliferate. *MIT Sloan Manage. Rev*, 2023. URL https://sloanreview.mit.edu/projects/building-robust-rai-programs-as-third-party-ai-tools-proliferate/.

[71] Anka Reuel, Ben Bucknall, Stephen Casper, Tim Fist, Lisa Soder, Onni Aarne, Lewis Hammond, Lujain Ibrahim, Alan Chan, Peter Wills, Markus Anderljung, Ben Garfinkel, Lennart Heim, Andrew Trask, Gabriel Mukobi, Rylan Schaeffer, Mauricio Baker, Sara Hooker, Irene Solaiman, Alexandra Sasha Luccioni, Nitarshan Rajkumar, Nicolas Moës, Jeffrey Ladish, Neel Guha, Jessica Newman, Yoshua Bengio, Tobin South, Alex Pentland, Sanmi Koyejo, Mykel J. Kochenderfer, and Robert Trager. Open problems in technical ai governance, 2024. URL https://arxiv.org/abs/2407.14981.

[72] Anthony Cintron Roman, Jennifer Wortman Vaughan, Valerie See, Steph Ballard, Nicolas Schifano, Jehu Torres Vega, Caleb Robinson, and Juan M. Lavista Ferres. Open Datasheets: Machine-readable documentation for open datasets and responsible AI assessments. *CoRR*, abs/2312.06153, 2023. doi: 10.48550/ARXIV.2312.06153. URL https://doi.org/10.48550/arXiv.2312.06153.

[73] Negar Rostamzadeh, Diana Mincu, Subhrajit Roy, Andrew Smart, Lauren Wilcox, Mahima Pushkarna, Jessica Schrouff, Razvan Amironesei, Nyalleng Moorosi, and Katherine Heller. Healthsheet: Development of a transparency artifact for health datasets, 2022. URL https://doi.org/10.1145/3531146.3533239.

[74] Christin Seifert, Stefanie Scherzinger, and Lena Wiese. Towards generating consumer labels for machine learning models. In *2019 IEEE First International Conference on Cognitive*

Machine Intelligence (CogMI), Los Angeles, CA, USA, December 12-14, 2019, pages 173–179. IEEE, 2019. doi: 10.1109/COGMI48466.2019.00033. URL https://doi.org/10.1109/CogMI48466.2019.00033.

[75] Hong Shen, Wesley H. Deng, Aditi Chattopadhyay, Zhiwei Steven Wu, Xu Wang, and Haiyi Zhu. Value Cards: An educational toolkit for teaching social impacts of machine learning through deliberation, 2021. URL https://doi.org/10.1145/3442188.3445971.

[76] Charlotte Siegmann and Markus Anderljung. The Brussels effect and artificial intelligence: How EU regulation will impact the global AI market. *CoRR*, abs/2208.12645, 2022. doi: 10.48550/ARXIV.2208.12645. URL https://doi.org/10.48550/arXiv.2208.12645.

[77] Tobin South, Alexander Camuto, Shrey Jain, Shayla Nguyen, Robert Mahari, Christian Paquin, Jason Morton, and Alex 'Sandy' Pentland. Verifiable evaluations of machine learning models using zksnarks. *CoRR*, abs/2402.02675, 2024. doi: 10.48550/ARXIV.2402.02675. URL https://doi.org/10.48550/arXiv.2402.02675.

[78] Marius Take, Sascha Alpers, Christoph Becker, Clemens Schreiber, and Andreas Oberweis. Software design patterns for AI-systems. In Agnes Koschmider and Judith Michael, editors, *11th International Workshop on Enterprise Modeling and Information Systems Architectures, Kiel, Germany, May 21-22, 2021*, volume 2867 of *CEUR Workshop Proceedings*, pages 30–35. CEUR-WS.org, 2021. URL https://ceur-ws.org/Vol-2867/paper5.pdf.

[79] TensorFlow. Model Card Toolkit | Responsible AI Toolkit. https://www.tensorflow.org/responsible_ai/model_card_toolkit/guide, 2023. [Accessed 23-08-2024].

[80] The White House. Blueprint for an AI Bill of Rights. https://www.whitehouse.gov/ostp/ai-bill-of-rights/, October 2022. [Accessed 23-08-2024].

[81] Eva Thelisson and Himanshu Verma. Conformity assessment under the EU AI Act general approach. *AI Ethics*, 4(1):113–121, 2024. doi: 10.1007/S43681-023-00402-5. URL https://doi.org/10.1007/s43681-023-00402-5.

[82] trail. EU AI Act Compliance Checker. https://www.trail-ml.com/eu-ai-act-compliance-checker, 2024. [Accessed 22-08-2024].

[83] Sandra Wachter. Limitations and loopholes in the EU AI Act and AI Liability Directives: what this means for the European Union, the United States, and beyond. *Yale Journal of Law & Technology*, 26(3), 2024. URL https://yjolt.org/limitations-and-loopholes-eu-ai-act-and-ai-liability-directives-what-means-european-union-united.

[84] Jiasi Weng, Jian Weng, Chengjun Cai, Hongwei Huang, and Cong Wang. Golden Grain: Building a secure and decentralized model marketplace for MLaaS. *IEEE Trans. Dependable Secur. Comput.*, 19(5):3149–3167, 2022. doi: 10.1109/TDSC.2021.3085988. URL https://doi.org/10.1109/TDSC.2021.3085988.

[85] Herbert Woisetschläger, Alexander Erben, Bill Marino, Shiqiang Wang, Nicholas D. Lane, Ruben Mayer, and Hans-Arno Jacobsen. Federated learning priorities under the European Union Artificial Intelligence Act, 2024. URL https://arxiv.org/abs/2402.05968.

[86] Doris Xin, Stephen Macke, Litian Ma, Jialin Liu, Shuchen Song, and Aditya Parameswaran. HELIX: holistic optimization for accelerating iterative machine learning. *Proc. VLDB Endow.*, 12(4):446–460, dec 2018. ISSN 2150-8097. doi: 10.14778/3297753.3297763. URL https://doi.org/10.14778/3297753.3297763.

[87] Xinyu Yang, Weixin Liang, and James Zou. Navigating dataset documentations in AI: A large-scale analysis of dataset cards on Hugging Face, 2024. URL https://doi.org/10.48550/arXiv.2401.13822.

[88] Matei Zaharia, Omar Khattab, Lingjiao Chen, Jared Quincy Davis, Heather Miller, Chris Potts, James Zou, Michael Carbin, Jonathan Frankle, Naveen Rao, and Ali Ghodsi. The shift from models to compound AI systems. https://bair.berkeley.edu/blog/2024/02/18/compound-ai-systems/, 2024. [Accessed 23-08-2024].

A Tables

Table 2: Project CC Attributes (Overview)

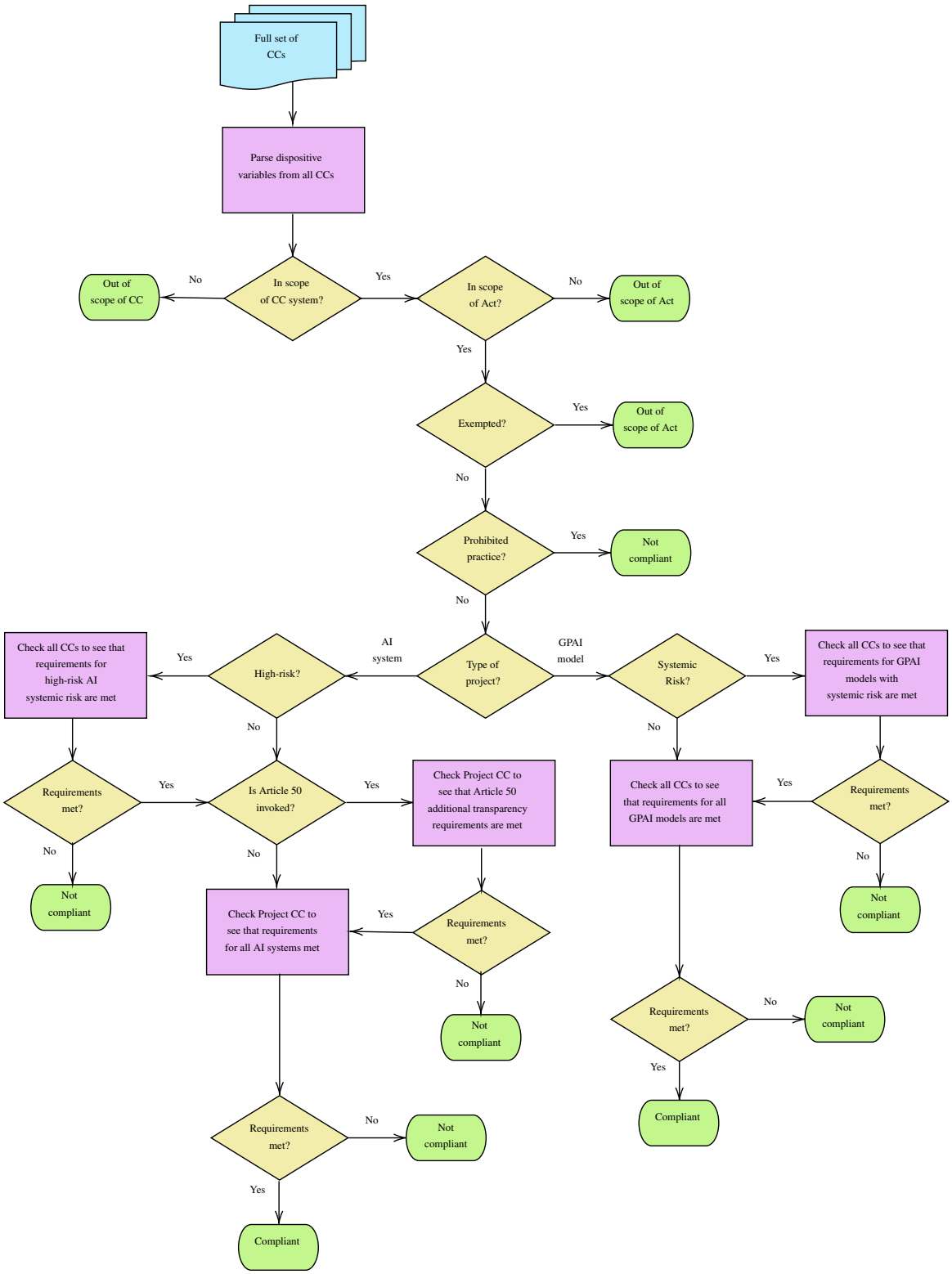
Attribute category (dispositive characteristics in italics)	Relevant Section(s) of AIA
Operator role (provider)	Art. 2
Intended purpose	Art. 6, 9, 10, 13, 14, 15
Whether on EU market	Art. 2
Whether AI System, High-risk AI System, GPAI model, GPAI model with Systemic Risk	Art. 3(1), 3(63), 6
Whether excepted from AIA	Art. 2(6)
Whether prohibited practices	Art. 5
Risk management (High-risk AI systems)	Art. 9
Data and data governance (High-risk AI systems)	Art. 10
Technical documentation (High-risk AI systems and GPAI models))	Art. 11, 53(1); Annex XI(2), XI(1)
Record-keeping (High-risk AI systems)	Art. 12
Transparency and provision of information to deployers (High-risk AI systems)	Art. 13
Human oversight (High-risk AI systems)	Art. 14
Accuracy, robustness, and cybersecurity (High-risk AI systems)	Art. 15
Registration, etc. (AI systems)	Art. 16
Fundamental Rights Assessment (High-risk AI systems)	Art. 27
Transparency (AI systems)	Art. 50

Table 3: Data CC Attributes (Overview)

Attribute category (dispositive characteristics in italics)	Relevant Section(s) of AIA
Intended purpose	Art. 6, 9, 10, 13, 14, 15
Data and data governance (High-risk AI systems)	Art. 10; Rec. 10, 67
Technical documentation (High-risk AI systems and GPAI models)	Art. 11, 13, 53(1); Annex IV, XI;
Accuracy, robustness and cybersecurity (High-risk AI systems)	Art. 15
Quality management system (High-risk AI systems)	Art. 17(1)(f)

Table 4: Model CC Attributes (Overview)

Attribute category (dispositive characteristics in italics)	Relevant Section(s) of AIA
Intended purpose	Art. 6, 9, 10, 13, 14, 15
Risk management (High-risk AI systems)	Art. 9(2), 9(6)
Data and data governance (High-risk AI systems)	Art. 10
Technical documentation (High-risk AI systems and GPAI models)	Art. 11, 53(1); Annex IV, XI;
Transparency and provision of information to deployers (High-risk AI systems, GPAI models, and GPAI models with systemic risk)	Art. 13
Human oversight (High-risk AI systems)	Art. 14(3)(a), 14(4)(d)
Accuracy, robustness and cybersecurity (High-risk AI systems)	Art. 15

B Diagrams

B.1 Compliance Cards Algorithm Flowchart