

Scale-Aware Diffusion Conditioning for Consistent Multi-Style Image Generation Across Diverse Visual Domains

Erkan Spiess

School of Information Technology, University of Cincinnati, Cincinnati, OH, USA.
erkanmail@uc.edu

Abstract

Diffusion-based generative systems have become a dominant infrastructure for text-to-image synthesis, yet their extension to consistent multi-style generation across diverse visual domains remains a difficult systems problem. Existing approaches often rely on separate style encoders, per-style fine-tuning, or prompt engineering that does not generalize across photographic, illustrative, architectural, medical, and industrial visual domains. This paper examines scale-aware diffusion conditioning as an architectural and governance strategy for producing stylistically coherent outputs without sacrificing domain flexibility. The central argument is that style consistency should not be treated as a single global attribute but as a multi-scale property involving local texture, mid-level composition, and global semantic structure. When conditioning is separated across these scales, diffusion models can preserve content identity, reduce style drift, and support multiple styles within a shared deployment infrastructure. The paper analyzes structural trade-offs involving adapter routing, latent caching, memory bandwidth, training data provenance, evaluation metrics, and multi-tenant serving. It further considers robustness under distribution shift, fairness across cultural and domain-specific visual languages, sustainability implications of repeated fine-tuning, and policy questions concerning transparency, attribution, and content provenance. The discussion concludes that scale-aware conditioning is not merely a model modification but a socio-technical layer that connects model architecture, operational monitoring, and institutional accountability. Consistent multi-style generation therefore requires coordinated advances in conditioning design, infrastructure governance, and domain-sensitive evaluation.

Keywords

diffusion models, scale-aware conditioning, multi-style generation, visual domain adaptation, generative infrastructure, robustness, fairness, AI governance.

1. Introduction

Generative diffusion models have rapidly moved from research prototypes to production infrastructure for image synthesis. Their iterative denoising process enables high-fidelity generation and flexible conditioning, but it also introduces system-level challenges when models must serve many styles, domains, and user expectations simultaneously [1], [2], [3]. In early deployments, a single model could produce plausible images from text prompts, yet stylistic control often depended on prompt wording, auxiliary encoders, or costly fine-tuning. Such approaches become fragile when the same system must generate photographic realism, editorial illustration, technical drawing, medical visualization, and culturally specific visual idioms within one operational environment. The problem is not only model quality but consistency, maintainability, and governance across heterogeneous visual domains [4], [5].

Scale-aware diffusion conditioning addresses this problem by treating style as a layered signal rather than a monolithic embedding. Local scale governs texture, brushwork, grain, and material appearance; mid-level scale governs composition, shape language, and recurring motifs; global scale governs semantic atmosphere, color grammar, and domain conventions. If these scales are entangled in a single conditioning vector, a change in texture may unintentionally alter composition, and a change in global style may erase domain-specific content. Scale separation allows the system to adjust one layer while stabilizing others, which is essential for multi-style consistency across diverse visual domains [6]. This paper develops a system-level account of that idea, focusing on architecture, deployment, evaluation, sustainability, fairness, and policy rather than narrow model benchmarking.

The argument is that scale-aware conditioning should be understood as an infrastructural commitment. It shapes how training data is curated, how adapters are routed, how latents are cached, how quality is monitored, and how accountability is assigned when generated content causes harm. Diffusion systems are often discussed as if they were isolated models, but production systems are assemblages of data pipelines, accelerators, storage tiers, orchestration services, safety filters, and human review processes [23]. Multi-style generation intensifies these dependencies because each style may carry different cultural meanings, legal risks, and domain expectations. A scale-aware approach offers a way to make those dependencies explicit and governable, but only if it is paired with robust evaluation and transparent operational policy.

2. Background and Related Work

Diffusion probabilistic models learn to reverse a gradual noise process, producing samples through iterative refinement [1]. Denoising diffusion implicit models reduced sampling cost and improved practical deployability, while latent diffusion moved computation into a compressed representation and enabled high-resolution synthesis with manageable hardware demands [2], [3]. These advances made diffusion competitive with generative adversarial networks and, in many settings, superior for text-conditioned image generation [4]. The introduction of large-scale vision-language supervision further connected natural language prompts to visual concepts, allowing users to specify content and style through text [5]. More recent text-to-image systems have demonstrated photorealistic generation with deep language understanding, but they also reveal the difficulty of maintaining fine-grained stylistic control across domains [15].

Style transfer research offers an older lineage for understanding multi-style generation. Early neural style transfer separated content and style through feature statistics, enabling artistic rendering but often lacking semantic control [9]. Adaptive instance normalization improved arbitrary style transfer by modulating feature statistics in real time, yet it remained sensitive to content-style conflicts [8]. Later work used contrastive language-image objectives to adapt generators toward new domains, showing that style can be shifted through semantic guidance rather than explicit style labels [7]. Personalization methods such as DreamBooth and textual inversion demonstrated that subject or style concepts can be injected into diffusion models with limited examples, but they also increase model management complexity and risk overfitting [10], [13]. Low-rank adaptation provided a parameter-efficient alternative for specialization, while control adapters enabled spatial and structural conditioning [11], [12], [14]. These methods are valuable, but they often treat style as a separate module or a new checkpoint rather than a scale-structured signal within a shared backbone.

Scale-separated conditioning has emerged as a promising response. By conditioning different denoising stages or feature resolutions with distinct style signals, a model can preserve global coherence while allowing local variation, and it can reduce dependence on dedicated style encoders [6]. This is important for multi-style deployment because a single service may need to switch among styles without loading entirely separate models. However, the literature still lacks a unified account of how such conditioning interacts with infrastructure constraints, fairness concerns, and policy obligations. The remainder of this paper addresses that gap by connecting model-level conditioning to system-level design and governance.

3. Scale-Aware Conditioning as a Systems Problem

Scale-aware conditioning is often described as a modeling technique, but in operational terms it is a coordination mechanism. It determines which parts of a latent representation are allowed to vary when a user changes style, which parts remain anchored to content, and which parts are governed by domain-specific constraints. In a multi-style system, the conditioning stack must support style selection, intensity control, content preservation, and domain adaptation without forcing every combination into a separate model. This creates a combinatorial problem. If there are ten styles and ten domains, a naive deployment might require one hundred fine-tuned variants. A scale-aware system instead seeks a smaller set of shared representations and routing rules that can compose style and content across scales [6], [11].

The primary trade-off is between consistency and flexibility. Strong global conditioning improves stylistic coherence but may suppress domain-specific detail. Weak global conditioning preserves flexibility but allows style drift across images or batches. Local conditioning can add texture and material realism, yet it may produce seams, flicker, or inconsistent edges when applied independently per region. Mid-level conditioning can stabilize composition, but it may also impose unwanted templates that reduce creative variation. A well-designed system does not eliminate these tensions; it allocates them across scales and makes the allocation visible through controls, defaults, and audit logs. This is an architectural decision with direct user experience consequences, because creators and enterprises need predictable behavior when they switch styles or combine domains.

A second systems issue is compute and memory placement. Diffusion inference is iterative and memory-intensive, especially at high resolution. If scale-aware conditioning requires multiple adapters, cross-attention modules, or separate conditioning branches, the system may incur additional latency and accelerator memory pressure. Latent caching, adapter caching, and batched routing can mitigate these costs, but they introduce cache invalidation and versioning problems [3], [22]. A style update may invalidate cached latents for some domains but not others, and an adapter change may affect only one scale while leaving others stable. These partial dependencies make deployment more complex than monolithic model updates. They also create opportunities for fine-grained rollback, because a faulty local texture adapter can be disabled without discarding the entire style model.

4. Architecture and Training Infrastructure

A production architecture for scale-aware multi-style diffusion conditioning typically separates data ingestion, conditioning representation, denoising backbone, adapter routing, safety filtering, and serving. The data layer must track domain provenance, licensing, and style attribution because stylistic imitation can raise copyright and cultural appropriation concerns. Datasheets for datasets and model cards provide minimal documentation practices

that become more important when a system mixes many domains [20], [21]. The conditioning layer must encode style at multiple scales, possibly through text prompts, reference images, or learned tokens. The backbone performs denoising, while adapters or routing modules modulate specific feature levels. The serving layer manages batching, caching, failover, and tenant isolation. Each component introduces hidden technical debt if interfaces are not versioned or if evaluation is disconnected from deployment [23].

Training such a system requires more than a large image-text corpus. It requires balanced domain sampling, style disentanglement objectives, and validation across scales. If one domain dominates, the model may interpret its conventions as universal, marginalizing other visual languages. If style labels are coarse, the system may conflate subject matter with style, causing content leakage. If scale separation is not supervised or regularized, the model may learn entangled representations that appear effective on average but fail under composition. Parameter-efficient adaptation methods can reduce training cost, but they also create a proliferation of adapters that must be tested, documented, and retired [14]. The infrastructure therefore needs a registry that maps styles, domains, scales, metrics, and safety status to specific model artifacts.

Infrastructure governance is also a training concern. Reproducibility depends on deterministic data splits, pinned library versions, hardware profiles, and evaluation harnesses. Without these, a style improvement in one domain may be impossible to attribute or reproduce. Multi-tenant systems intensify the problem because different customers may supply private reference images or fine-tuned adapters. Isolation, access control, and deletion guarantees become part of the model architecture in practice, even if they are not expressed in the denoising equations. Scale-aware conditioning can support tenant isolation by keeping customer-specific style signals in separate adapter spaces, but only if the routing layer enforces boundaries and logs access.

5. Consistency Across Multi-Style Generation

Consistency in multi-style generation has several meanings. Within a single image, it means that local texture, mid-level forms, and global semantics do not contradict one another. Across a batch, it means that the same style produces similar visual grammar despite variations in content. Across time, it means that model updates do not abruptly change the style users have incorporated into their workflows. Across domains, it means that a style can be adapted to photography, illustration, or technical visualization without collapsing into a single average look. Scale-aware conditioning helps because it provides separate control points for these forms of consistency, but it does not automatically resolve conflicts. A local texture may be consistent while global composition drifts, or a global style may remain stable while domain-specific details are lost [6], [7].

Style drift often arises from distribution shift. If a style is learned primarily from one domain, applying it to another domain may force the model to exaggerate recognizable features. For example, an illustrative style trained on fantasy landscapes may impose dramatic lighting when applied to architectural diagrams, reducing legibility. Mid-level conditioning can preserve structural conventions, while global conditioning can maintain a coherent atmosphere. But the system must know which conventions are essential. This requires domain-sensitive evaluation, not only aggregate similarity scores. It also requires user controls that allow intensity adjustment and region-specific application. Personalization and control adapters offer mechanisms for such control, but they must be integrated with scale routing rather than bolted on as separate pipelines [10], [11], [12].

Consistency also has an institutional dimension. Enterprises that adopt generative systems need predictable outputs for brand assets, product catalogs, and training materials. If a model update changes style behavior, downstream content may become inconsistent with prior materials. A scale-aware architecture can support versioned conditioning profiles, where global style remains fixed while local or mid-level adapters are updated. This resembles semantic versioning for visual identity. However, versioning increases storage and serving complexity, and it requires clear deprecation policies. Without governance, users may unknowingly mix incompatible style versions, producing outputs that appear broken or unprofessional.

6. Deployment, Robustness, and Domain Diversity

Deployment of scale-aware diffusion conditioning spans cloud inference, on-premise clusters, and edge or mobile scenarios. Cloud deployment offers elastic accelerators and centralized updates, but it raises latency, cost, and data sovereignty concerns. Edge deployment reduces network dependency and may improve privacy, but it constrains model size and adapter count, making efficient architectures important [22]. A practical system may use a hybrid approach: global style conditioning and safety filtering in the cloud, with lightweight local adapters for texture and preview generation on device. Such a split introduces synchronization problems, because local and remote scales must remain compatible. If an edge adapter is updated without a corresponding cloud change, the resulting style may drift or fail.

Robustness under diverse visual domains requires more than clean benchmark accuracy. The system must handle unusual aspect ratios, low-light images, medical or industrial content, cultural symbols, and adversarial prompts. Safety filters can reduce harmful outputs, but they may also over-filter legitimate artistic or educational content, creating fairness concerns [16], [17]. Concrete problems in AI safety highlight specification, robustness, and unintended behavior as persistent challenges, and generative systems add new forms of ambiguity because style and content are entangled [24]. A scale-aware approach can help by separating safety-critical content constraints from stylistic variation, but only if the governance layer defines which scales are subject to which policies. For example, global semantics may require strict content moderation, while local texture may allow broader creative expression.

Domain diversity also raises fairness issues. Visual domains are not neutral; they encode cultural histories, professional norms, and power relations [25]. If a model trains primarily on Western art, commercial photography, and English-language captions, it may treat those conventions as default. Scale-aware conditioning can support multiple visual languages, but it cannot correct representational imbalance by itself. Fairness requires diverse data, participatory design, and evaluation that includes affected communities. It also requires transparency about which domains are supported and which are approximated. Model cards and datasheets are useful starting points, but they must be updated as styles and domains are added [20], [21]. Without such documentation, users may overtrust the system in domains where it has little validated competence.

7. Sustainability, Fairness, and Governance

The sustainability of multi-style diffusion systems depends on how often models are retrained, how many adapters are stored and served, and how efficiently inference is scheduled. Large-scale training has significant energy and policy implications, and repeated fine-tuning for every style and domain multiplies those costs [18]. Green AI argues for reporting efficiency alongside accuracy, not treating compute as an invisible externality [19]. Scale-aware

conditioning can improve sustainability by sharing a backbone across many styles and reducing the need for full model duplication. However, it can also increase inference complexity if multiple adapters are activated per request. The net environmental effect depends on routing efficiency, batching, hardware utilization, and the frequency of updates. Sustainable design therefore requires measurement at the system level, including training, fine-tuning, storage, and serving.

Fairness in this context is both representational and procedural. Representational fairness concerns whose visual cultures, bodies, and environments are depicted and how they are stylized. Procedural fairness concerns who can access the system, who can define style categories, and who can contest harmful outputs. A scale-aware conditioning layer may make it easier to isolate culturally specific style signals, but it may also commodify them if governance is weak. If styles are treated as interchangeable settings, the system may obscure their origins and reproduce them without consent. Fairness therefore requires provenance tracking, licensing review, and mechanisms for opt-out or benefit sharing. These are not purely technical features; they are policy commitments embedded in data pipelines and interfaces.

Governance must also address accountability across the model lifecycle. When a generated image causes harm, responsibility may be distributed among data providers, model developers, platform operators, and users. Scale-aware conditioning can support accountability by logging which scales, adapters, and style profiles contributed to an output. Such logs can enable audits and incident response, but they also raise privacy concerns. Governance frameworks must balance transparency with confidentiality, especially for enterprise or personal reference images. The atlas of AI perspective reminds us that generative systems are embedded in material and political economies, not just technical pipelines [25]. Scale-aware conditioning should therefore be designed as a governable infrastructure, with clear ownership, auditability, and contestability.

8. Evaluation and Operational Metrics

Evaluation of scale-aware multi-style generation cannot rely on a single similarity score. It requires metrics for prompt alignment, content preservation, style consistency, domain appropriateness, robustness, fairness, and efficiency. Prompt alignment can draw on vision-language representations, but these representations may not capture fine-grained style or cultural nuance [5]. Content preservation may require identity or structural metrics, while style consistency may require both perceptual and human judgments. Domain appropriateness is especially difficult because it depends on professional and cultural standards that vary across contexts. A medical illustration and a fashion editorial may both value clarity, but they define it differently. Evaluation must therefore be stratified by domain and scale rather than aggregated into a single number.

Operational metrics should connect evaluation to deployment. Latency percentiles, adapter cache hit rates, failure rates, safety filter pass rates, and rollback frequency are as important as image quality scores. If a scale-aware system improves quality but doubles inference cost, it may be unsustainable for high-volume use. If it reduces style drift but increases false positives in safety filtering, it may harm legitimate creators. A balanced scorecard can make these trade-offs visible to operators and policymakers. Model cards and datasheets provide structured reporting templates, but they must be adapted to multi-scale conditioning so that style, domain, and scale-specific limitations are explicit [20], [21]. Continuous monitoring is necessary because visual domains and user expectations change over time.

Evaluation also needs adversarial and participatory methods. Red-teaming can reveal prompt injection, style imitation, and unsafe combinations that average metrics miss [24]. Participatory evaluation can reveal how communities interpret style, appropriation, and representation. These methods are slower and more qualitative than automated benchmarks, but they are essential for fairness and legitimacy. A system that performs well on standard datasets but fails in underrepresented domains is not robust in a meaningful sense. Scale-aware conditioning offers a framework for organizing such evaluations, because it separates local, mid-level, and global effects that can be tested independently and in combination.

9. Policy Implications and Forward-Looking Perspectives

Policy for multi-style generative systems must address transparency, intellectual property, safety, and competition. Transparency requires disclosure of training data sources, style provenance, and known limitations. Intellectual property debates are complicated by style, which is not easily reduced to discrete protected works. A scale-aware system could support attribution by tracking which style profiles contributed to an output, but attribution is not the same as legal clearance. Policymakers and industry standards bodies will need to clarify how style signals should be licensed, documented, and contested. Without such clarity, creators may face imitation without consent, while developers may face legal uncertainty that stifles legitimate innovation.

Safety policy must also account for combinatorial risk. A style that is harmless in one domain may become harmful in another, and scale-aware conditioning can create many combinations that were not explicitly tested. Regulatory approaches based on static model approval may struggle with continuously updated adapters and style profiles. Risk-based governance, mandatory incident reporting, and post-market monitoring may be more appropriate. However, these mechanisms require technical hooks for logging, auditing, and rollback. Scale-aware conditioning can provide such hooks if governance requirements are designed into the architecture rather than added after deployment. This aligns with broader arguments that foundation models require anticipatory governance and interdisciplinary oversight [17], [23].

Forward-looking research should integrate technical and social perspectives. On the technical side, promising directions include disentangled scale representations, efficient adapter routing, uncertainty estimation for style transfer, and domain-aware safety controls. On the social side, research is needed on consent, cultural preservation, labor impacts on artists, and equitable access. Policy experiments could include style registries, provenance standards, and procurement rules that require fairness and sustainability reporting. The goal is not to freeze generative systems but to make their behavior governable across diverse visual domains. Scale-aware conditioning is one component of that larger effort, and its value depends on how well it is embedded in institutions that can monitor, contest, and revise its use.

10. Conclusion

Scale-aware diffusion conditioning offers a principled way to manage consistent multi-style image generation across diverse visual domains. By separating local, mid-level, and global style signals, it reduces the need for monolithic style encoders and per-style model duplication while preserving content identity and stylistic coherence. Yet the approach is not solely a modeling innovation. It reshapes training infrastructure, adapter governance, serving economics, evaluation practices, and policy obligations. The trade-offs between consistency and flexibility, efficiency and control, and creativity and safety are unavoidable. They must be

managed through architecture, monitoring, and institutional accountability rather than hidden in default settings or unexamined metrics.

The paper has argued that scale-aware conditioning should be treated as a socio-technical layer. Its success depends on data provenance, domain-sensitive evaluation, fairness auditing, sustainability measurement, and transparent governance. Deployed well, it can support a more pluralistic generative infrastructure that respects diverse visual languages and enables predictable enterprise workflows. Deployed poorly, it can accelerate stylistic appropriation, concentrate power, and obscure responsibility. The path forward requires interdisciplinary collaboration among model architects, systems engineers, domain experts, policymakers, and affected communities. Consistent multi-style generation is therefore not only a technical benchmark but a test of whether generative systems can be built as accountable public and private infrastructure.

References

1. Ho, J., Jain, A., & Abbeel, P. (2020). Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems*, 33, 6840–6851.
2. Song, J., Meng, C., & Ermon, S. (2021). Denoising diffusion implicit models. *International Conference on Learning Representations*.
3. Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 10684–10695.
4. Dhariwal, P., & Nichol, A. (2021). Diffusion models beat GANs on image synthesis. *Advances in Neural Information Processing Systems*, 34, 8780–8794.
5. Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., & Sutskever, I. (2021). Learning transferable visual models from natural language supervision. *Proceedings of the 38th International Conference on Machine Learning*, 8748–8763.
6. Zhang, Jingtao, et al. "Scale-separated conditioning for style-encoder-free diffusion stylization." *arXiv preprint arXiv:2608.19719* (2026).
7. Gal, R., Patashnik, O., Maron, H., Bermano, A. H., Chechik, G., & Cohen-Or, D. (2022). StyleGAN-NADA: CLIP-guided domain adaptation of image generators. *ACM Transactions on Graphics*, 41(4), 1–13.
8. Huang, X., & Belongie, S. (2017). Arbitrary style transfer in real-time with adaptive instance normalization. *Proceedings of the IEEE International Conference on Computer Vision*, 1501–1510.
9. Gatys, L. A., Ecker, A. S., & Bethge, M. (2016). Image style transfer using convolutional neural networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2414–2423.
10. Ruiz, N., Li, Y., Jampani, V., Pritch, Y., Rubinstein, M., & Aberman, K. (2023). DreamBooth: Fine tuning text-to-image diffusion models for subject-driven generation. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 22500–22510.

11. Zhang, L., Rao, A., & Agrawala, M. (2023). Adding conditional control to text-to-image diffusion models. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 3836–3847.
12. Mou, C., Wang, X., Xie, L., Wu, Y., Zhang, J., Qi, Z., & Shan, Y. (2024). T2I-Adapter: Learning adapters to dig out more controllable ability for text-to-image diffusion models. *Proceedings of the AAAI Conference on Artificial Intelligence*, 38(5), 4296–4304.
13. Gal, R., Alaluf, Y., Atzmon, Y., Patashnik, O., Bermano, A. H., Chechik, G., & Cohen-Or, D. (2023). An image is worth one word: Personalizing text-to-image generation using textual inversion. *International Conference on Learning Representations*.
14. Hu, E. J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L., & Chen, W. (2022). LoRA: Low-rank adaptation of large language models. *International Conference on Learning Representations*.
15. Saharia, C., Chan, W., Saxena, S., Li, L., Whang, J., Denton, E., Ghasemipour, S. K. S., Ayan, B. K., Mahdavi, S. S., Lopes, R. G., Salimans, T., Ho, J., Fleet, D. J., & Norouzi, M. (2022). Photorealistic text-to-image diffusion models with deep language understanding. *Advances in Neural Information Processing Systems*, 35, 36479–36494.
16. Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623.
17. Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N., Chen, A., Creel, K., Davis, J. Q., Demszky, D., ... Liang, P. (2021). On the opportunities and risks of foundation models. *arXiv preprint arXiv:2108.07258*.
18. Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and policy considerations for deep learning in NLP. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 3645–3650.
19. Schwartz, R., Dodge, J., Smith, N. A., & Etzioni, O. (2020). Green AI. *Communications of the ACM*, 63(12), 54–63.
20. Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé III, H., & Crawford, K. (2021). Datasheets for datasets. *Communications of the ACM*, 64(12), 86–92.
21. Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., & Gebru, T. (2019). Model cards for model reporting. *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 220–229.
22. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). MobileNetV2: Inverted residuals and linear bottlenecks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 4510–4520.
23. Sculley, D., Holt, G., Golovin, D., Davydov, E., Phillips, T., Ebner, D., Chaudhary, V., Young, M., Crespo, J. F., & Dennison, D. (2015). Hidden technical debt in machine learning systems. *Advances in Neural Information Processing Systems*, 28, 2503–2511.
24. Amodei, D., Olah, C., Steinhardt, J., Christiano, P., Schulman, J., & Mané, D. (2016). Concrete problems in AI safety. *arXiv preprint arXiv:1606.06565*.

25. Crawford, K. (2021). *The atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press.