Digital Weight: Towards an Ecological Evaluation Layer for Generative AI in Computational Creativity

Jekaterina Kaizere
Independent Artist-Researcher
Idar-Oberstein, Germany
katerina@kaizere.de

Abstract

Generative AI systems are increasingly positioned as tools of computational creativity, yet the material conditions enabling them rarely enter the evidential record of computational creativity evaluation. Existing evaluation discourse typically asks what systems generate (novelty, value, quality), how they generate it (process characteristics, autonomy, interaction), and how outputs are received. Under contemporary generative practice, compute regime, generation count, inference cost, emissions, and water use condition how such claims can be compared, justified, and audited. Infrastructural Situatedness is proposed here as a cross-cutting conditioning layer within computational creativity evaluation. Methodologically, when systems make comparable claims about novelty, value, autonomy, or creative search while relying on materially different infrastructures, those differences should be visible in the evaluation record. Two precedents matter here: computational creativity evaluation already debates what evidence counts for creativity claims and how systems can be compared systematically; broader AI evaluation argues that efficiency and the "price tags" of model development and use should be reported (Strubell et al. 2019; Schwartz et al. 2020). A case demonstration, Digital Weight, operationalises per-generation accounting by translating published per-inference energy coefficients into time-based perceptual structure. The case makes infrastructural constraints part of generative logic and available to evaluation. The contribution is a compact evaluation record for infrastructural conditions in computational creativity.

Introduction

Computational creativity has developed a mature evaluation culture (Jordanous 2012; Colton 2008): debates about machine creativity are typically grounded in explicit criteria, defensible evidence, and domain-relative standards, moving beyond informal "it seems creative" claims. Evaluation frameworks and tutorials formalise how creativity claims can be substantiated across artefact properties, generative processes, and audience reception (Lamb, Brown and Clarke 2018). For example, creativity is often operationalised through novelty, surprise, and value, following Margaret A. Boden's characterisation of creative ideas or artefacts as "new, surprising, and valuable" (Boden 2004). In CC evaluation, these dimensions have also been treated as tractable criteria for assessing design creativity (Grace et al. 2015).

At the same time, many contemporary generative systems used in creative practice rely on large-scale computational infrastructures. Unlike earlier eras of symbolic creativity software, the resource intensity of a creative system can now be estimated at the level of repeated use: deployment cost can be benchmarked in energy and carbon per 1,000 inferences, and task modality matters strongly. In a large comparative study of inference costs, Luccioni, Jernite and Strubell (2024) report that multimodal tasks (including image captioning and image generation) sit at the high end of the measured spectrum (0.06–2.9 kWh per 1,000 inferences), and their task-level summary table reports image generation with a mean of 2.907 kWh per 1,000 queries – orders of magnitude higher than many text classification settings.

This paper argues that infrastructural cost is rarely treated as part of computational creativity evaluation. Existing frameworks can accommodate process-level considerations, yet they do not typically require the material conditions of producing creative artefacts to be disclosed, contextualised, or compared. This omission was historically understandable when computation could be treated, for many evaluative purposes, as a relatively neutral substrate. Contemporary generative systems make this assumption less stable. Two systems may produce outputs with similar claims to novelty and value, or describe similar processes of exploration, autonomy, and interaction, while operating through radically different compute regimes. Without infrastructural disclosure, the evaluation record cannot distinguish constrained search from resource-amplified sampling, nor can it show when a creativity claim depends on extensive repeated generation, discarded candidates, or large hosted models.

This paper proposes Infrastructural Situatedness as a cross-cutting conditioning layer within computational creativity evaluation. It leaves creativity judgements to existing novelty, value, process, autonomy, and reception-based criteria. Methodologically, it affects how these criteria are read when systems rely on materially different computational infrastructures. At minimum, the layer records compute regime, estimation method, and, where available, electricity, emissions, and water estimates, so that artefact-, process-, and reception-oriented evaluations can be read under known infrastructural assumptions. The aim is auditability under uncertainty. As a case demonstration, Digital Weight translates published per-inference coefficients into perceptible temporal structure: cadence, duration, and accumulation become part of the work's generative constraints. The paper contributes a compact evaluation record for making creativity claims more comparable, justified, and auditable under contemporary generative AI conditions.

Background on evaluation in computational creativity

Computational creativity has developed a mature evaluation culture in which creativity claims are expected to be supported by explicit criteria and defensible evidence. Carolyn Lamb, Daniel G. Brown, and Charles L. A. Clarke provide an interdisciplinary evaluation tutorial that surveys diverse evaluation methods and emphasises the need for explicit evaluation choices and assumptions because there is no single consensus method for evaluating a creative system (Lamb, Brown and Clarke 2018).

Three evaluation anchors position Infrastructural Situatedness as contextual evidence within existing evaluation methodology.

First, Anna Jordanous proposes the SPECS procedure for evaluating creative systems: researchers specify what creativity means for a system, derive evaluation criteria, and apply domain-sensitive tests (Jordanous 2012). Jordanous also emphasises that comparative and systematic evaluation is essential for progress in computational creativity research (Jordanous 2012).

Second, Graeme Ritchie develops empirical criteria for attributing creativity to computer programs based on observable relationships between inputs and outputs. Importantly, he notes that if two systems produce similarly rated outputs, the system generating them in a relatively simpler manner might be considered more creative, while also recognising the difficulty of incorporating such "effort" information into evaluation (Ritchie 2007).

Third, Simon Colton introduces the "creative tripod", describing creative systems in terms of skill, appreciation, and imagination, and distinguishes creativity from the perception of creativity by observers (Colton 2008). For this argument, Colton's distinction matters because it supports the comparison and justification of creativity claims without asserting a single universal metric.

Together, these accounts make the proposed layer methodologically continuous with existing CC evaluation practice: if CC evaluation culture already requires explicit evidence and has precedent for considering "simplicity/effort" as part of comparative judgement, then resource intensity can be treated as contextual evidence of infrastructural effort and search scale in contemporary generative systems.

FACE/IDEA and related descriptive models make clear that CC evaluation already extends beyond final artefacts. Colton, Charnley and Pease (2011) describe creative acts through tuples of generative acts involving expressions, concepts, aesthetic measures, and framing information; IDEA adds impact and audience reception. Lamb, Brown and Clarke (2018) similarly organise evaluation across product, process, person/producer, and press. FACE/IDEA help locate Infrastructural Situatedness inside the evaluative record: what was generated, by what process, under what constraints, with what framing, and within what reception conditions.

Wiggins' Creative Systems Framework, developed through a formalisation of Boden's account, further supports comparison between creative systems in terms of conceptual spaces, rules, search procedures, and constraints (Wiggins 2006). Ventura's discussion of Mere Generation clarifies why generation alone is insufficient for creativity claims: novelty, value, and intentionality require an evaluative context, especially when outputs can be produced by very different processes (Ventura 2016). Mondol and Brown's Algorithmic Information Theory account is relevant here: it connects creativity evaluation with computational history, effort, complexity, novelty, and value through concepts such as logical depth and sophistication (Mondol and Brown 2021a; Mondol and Brown 2021b). Infrastructural Situatedness gives this concern with process and effort a practical evaluation record: it makes compute regime, generation count, resource intensity, and transparency available as contextual evidence for evaluating and comparing creative search. Energy, carbon, and water use are not treated here as creativity criteria. They become part of the evidential context through which materially different generative systems can be compared, justified, and audited.

Why infrastructure matters now

The infrastructural case is grounded in measurable changes in scale, deployment, and resource demand.

At the macro level, International Energy Agency estimates that electricity consumption from data centres is around 415 TWh, or about 1.5% of global electricity consumption in 2024 (IEA 2025), with substantial uncertainty and scenario-based projections for growth. For computational creativity, the issue is comparative: when creative systems rely on data centres and accelerated servers, infrastructure is no longer a neutral "implementation detail," especially in aggregate and at scale.

At the micro level, deployment and inference cost can be reported per generation or per 1,000 inferences. Luccioni, Jernite and Strubell (2024) operationalise deployment cost as energy and carbon for 1,000 inferences, showing large differences by modality and task structure. They report that multimodal tasks such as image captioning and image generation are on the highest end of the spectrum (0.06–2.9 kWh per 1,000 inferences), and their task-level means indicate image generation at ~2.907 kWh per 1,000 queries. They also report extreme variation across tasks (a factor exceeding 1,000 in their analysis), underscoring that "what task you do" can dominate "how big the model is" in energy terms (Luccioni, Jernite and Strubell 2024).

Carbon captures only part of infrastructural impact. Data centre impacts include water consumption for cooling and electricity generation pathways, and these impacts vary strongly by geography, cooling practice, and energy mix. A Massachusetts Institute of Technology explainer cites an often-used order-of-magnitude estimate of about two litres of water per kWh of energy a data centre consumes for cooling (Zewe 2025), illustrating why water can be reported alongside carbon as a separate, assumption-bound metric.

These coefficients are context dependent. David Mytton reports that water intensity for electricity generation varies widely: consumption ranges from 0.00 to 4.4 L/kWh, while withdrawals can be much higher depending on cooling technology and generation mix (Mytton 2021).

Finally, emissions conversions require explicit assumptions about grid carbon intensity. A widely cited 2024 transition review reports global power-sector CO_2 intensity reaching approximately 480 gCO_2/kWh in 2023, offering an example emission factor that can be transparently parameterised (and replaced with regional factors where available) (Ember 2024).

Operationalising Infrastructural Situatedness within CC Evaluation

This section defines Infrastructural Situatedness and its evaluation record for use with output-, process-, and reception-oriented frameworks.

Definition

Infrastructural Situatedness is a cross-cutting conditioning layer that requires explicit disclosure and contextualisation of the infrastructural conditions enabling a generative process. At minimum, this includes the compute regime, the estimation method, and the conversion factors used to express electricity, emissions, and (optionally) water.

The proposal is motivated by two converging precedents:

- Within CC, Ritchie's discussion shows that "cost/effort" has already been considered as a possible differentiator when outputs are similar, even if the field lacks standard ways to weigh it (Ritchie 2007).
- Within broader AI evaluation culture, Green AI argues for making efficiency visible alongside accuracy and for reporting the "price tag" of developing, training, and running models (Schwartz et al. 2020).

How Infrastructural Situatedness Interacts with Creativity Evaluation

Infrastructural Situatedness specifies the evidential context in which creativity claims are compared and audited. The methodological issue is what changes when the evaluation record includes compute regime, generation count, resource intensity, and transparency. The layer matters most where systems make claims about novelty, value, autonomy, quality, or creative search while relying on materially different infrastructures.

Comparative interpretation. CC evaluation is often comparative, even when it avoids a single creativity score. Two systems may produce artefacts judged similarly novel, valuable, or high quality, while reaching them through very different search conditions. One may operate through constrained search, limited local inference, or a carefully specified generative procedure. Another may depend on large-scale hosted generation, repeated prompt variation, extensive sampling, and many discarded candidates. Artefact-level assessment alone misses this difference. Infrastructural Situatedness distinguishes compact or constrained search from resource-amplified sampling. This connects to Ritchie's observation that effort and simplicity may matter when systems produce similarly rated outputs (Ritchie 2007), while keeping resource use separate from creativity ranking.

Process interpretation. Claims about computational creativity often rest on process: autonomy, iteration, selection, evaluation, interaction, and the system's role in shaping its own outputs. For contemporary generative AI workflows, the process also includes generation count, discarded samples, batch size, model choice, runtime, prompt revision, local or hosted compute, and the degree to which these conditions can be inspected. These factors make process evidence more complete. A claim that a system explored a space has a different evidential profile if the exploration involved ten constrained generations, ten thousand unlogged generations, or a hosted model whose sampling and energy profile remain unavailable. Mondol and Brown's Algorithmic Information Theory account is relevant here because it links creativity evaluation to computational history, effort, complexity, novelty, and value through concepts such as logical depth and sophistication (Mondol and Brown 2021a; Mondol and Brown 2021b). Infrastructural Situatedness turns this concern with computational history into a practical evaluation record for generative AI.

Constraint-based creativity. Infrastructural conditions can shape the creative process before evaluation. An explicit energy, generation, or compute budget can function as a constraint on creative search, comparable to other procedural constraints in computational creativity and generative art. A system may be designed to search within a fixed number of generations, optimise under an inference budget, expose accumulation as part of the work, or restrict sampling so that selection, transformation, and judgement occur earlier. Here the infrastructural condition structures the creative space itself. Digital Weight demonstrates this relation at the level of perceptual form: cadence, duration, repetition, and accumulation are organised through per-generation accounting. In Digital Weight, energy, carbon, and water operate as constraints within the work's generative logic.

Auditability of creative search. Infrastructural Situatedness supports auditability under uncertainty. Creative AI workflows are often nonlinear: prompts branch, candidates are discarded, outputs are recombined, and human judgement intervenes repeatedly. The evaluation record should prioritise inspectability over exact reconstruction. Where possible, authors should count all generated candidates, including discarded samples, unsuccessful generations, and intermediate variations. Branching workflows may distinguish total project generation count from the count associated with the final selected branch. For collaborative workflows, authors may report aggregate project counts, author-attributed generation logs, or clearly bounded estimates when individual attribution is impractical. Acceptable evidence includes API histories, local generation logs, version control, exported prompt sheets, model-setting records, batch files, and manual process notes. The search path becomes inspectable enough for comparative interpretation, without implying complete capture of every micro-decision.

The layer's contribution is epistemic and comparative. Infrastructural Situatedness makes visible which computational conditions support a creativity claim, how much of the creative search remains available for inspection, and where uncertainty remains. It conditions artefact-, process-, and reception-oriented evaluation while keeping creativity distinct from footprint accounting. These methodological implications motivate the need for a standardised evaluation record capable of documenting infrastructural conditions across creative workflows.

Operationalising the extension as an evaluation record

Evaluation frameworks often omit infrastructural conditions because measurement is uneven and assumption-dependent. ML sustainability work treats systematic reporting with transparent assumptions as a practical alternative to full lifecycle assessment in every paper. Henderson et al. (2020) frame energy and carbon reporting as necessary for assessing the climate impacts of machine learning research.

Adapting that logic to computational creativity, Table 1 proposes a standardised evaluation record for documenting the infrastructural conditions under which creative outputs are produced. The table operationalises comparative interpretation and auditability by recording compute regime, generation count, resource intensity, transparency, and uncertainty alongside existing creativity evidence.

Table 1. Infrastructural Situatedness evaluation record for computational creativity systems. The table records process and infrastructural evidence relevant to comparative interpretation, process interpretation, and auditability. It avoids creativity scoring and exact footprint accounting.

Field	**What is recorded**	**Evaluation function**	**Auditability note**
Model / System	Model name, version, API/local system, release date where available	Identifies the system being evaluated	Include provider, version, or local configuration
Task / modality	Text, image, audio, video, multimodal, robotic, live, hybrid	Supports comparison across task types	Specify output type and workflow context
Compute regime	Local device, cloud/API, hosted model, cluster, hybrid	Contextualises process and infrastructure	Report hardware/API access level where known
Creative workflow	One-shot, iterative, branching, collaborative, live, automated	Describes creative search structure	Include workflow description or process log
Generation count	Total generations, selected outputs, discarded candidates	Shows search scale and sampling intensity	Use API logs, local logs, version history, or manual count
Runtime / duration	Runtime, performance duration, batch duration, session length	Supports process interpretation	Report measured or estimated duration
Resource intensity	Energy, emissions, and water where available or estimable	Contextual evidence for infrastructural effort	Report units separately; do not aggregate
Estimation method	Telemetry, direct measurement, provider data, calculator, published coefficients	Makes assumptions inspectable	Prefer standardised tools or provider data where available

Field	What is recorded	Evaluation function	Auditability note
Transparency level	Measured, partially estimated, provider-limited, unknown	Indicates reliability of the evaluation record	State what is known, inferred, or unavailable
Uncertainty / boundary	Range, coefficient source, region, system boundary, exclusions	Prevents false precision	Distinguish operational estimates from lifecycle data

This evaluation record makes infrastructural conditions visible and comparable across works, supporting auditability without collapsing creativity evaluation into footprint accounting. Figure 1 represents this relation visually. Infrastructural Situatedness crosses artefact, process, and reception evaluation and conditions interpretation beyond a creativity score.

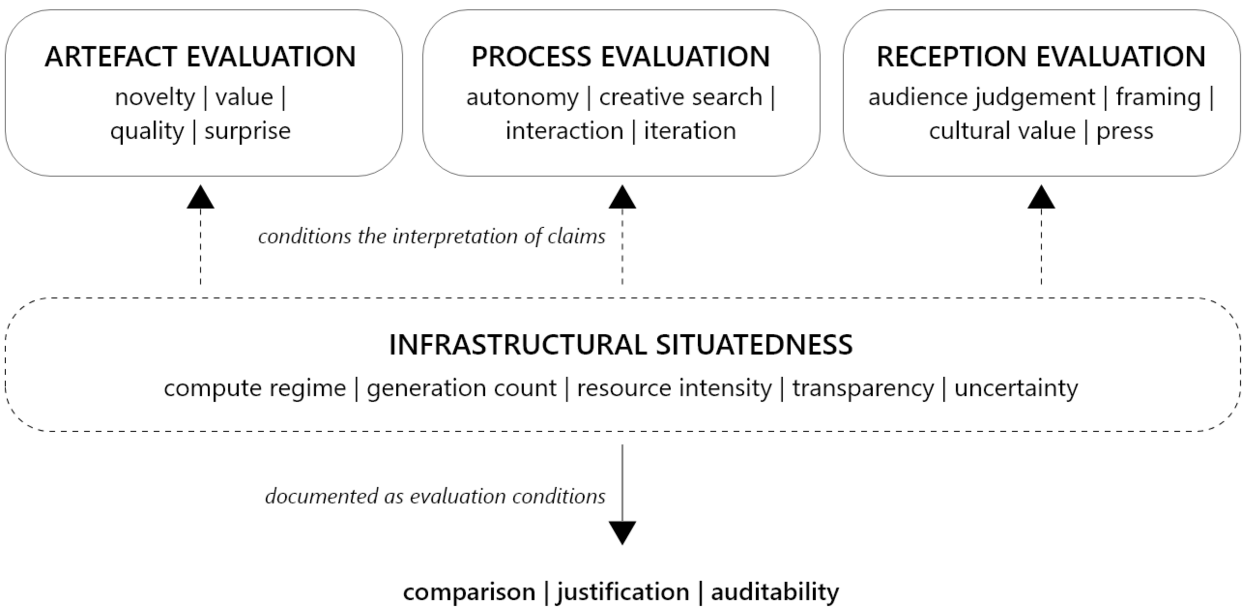

Figure 1. Infrastructural Situatedness as a cross-cutting conditioning layer within computational creativity evaluation. The layer leaves artefact-, process-, and reception-oriented criteria intact and is not an additional creativity score. Instead, it records infrastructural conditions including compute regime, generation count, resource intensity, transparency, uncertainty, and, where available, energy use, emissions, water consumption, and discarded candidates. This enables comparison, justification, and auditability under explicit infrastructural assumptions.

Digital Weight as an operational demonstration

Digital Weight provides a compact methodological demonstration: Infrastructural Situatedness can be operationalised and made perceptible within an artwork's generative logic.

Assume a cadence of 2 images/s for 8 minutes, yielding 960 generative events. Using Luccioni et al.'s task-level mean for image generation (2.907 kWh per 1,000 queries), a back-of-the-envelope estimate gives 0.0029 kWh per generated image and ~2.79 kWh for 960 images (Luccioni, Jernite and Strubell 2024). Using a global-average emissions factor example of ~480 gCO_2/kWh (Ember 2024) (explicitly parameterised), this corresponds to ~1.34 $kgCO_2$. Using a simple cooling proxy of ~2 L/kWh (Zewe 2025) yields ~5.6

L of cooling water. These values are included as illustrative and should be reported as parameterised coefficients.

The inference-cost range also shows scale sensitivity: Luccioni et al. note that multimodal tasks such as image captioning and image generation are on the high end of the spectrum (0.06-2.9 kWh per 1,000 inferences), and their analysis indicates that image generation can be orders of magnitude more energy-intensive than many text tasks.

Figure 2. Two consecutive frames of the artwork (001,002) showing the cumulative infrastructure counters associated with each AI generation cycle. Electricity, CO_2 emissions, and water consumption increase as the generative process continues.

What the case demonstrates for computational creativity evaluation

Digital Weight intentionally complicates output-centric evaluation. The visible artefact is nearly static; if evaluated only as a product, it may appear low in novelty. Yet the process repeatedly performs generative acts, and the infrastructural cost accumulates measurably with each act. The methodological point is that output novelty alone may fail to register the infrastructural intensity of a generative process, so evaluation that ignores load risks mischaracterising what is being optimised or externalised (Boden 2004; Ritchie 2007).

The project also demonstrates how infrastructural metrics can become generative constraints: in constraint mode, the artist can choose cadence, duration, or transformation rules as functions of an explicit energy budget, turning resource intensity into part of the creative search space rather than a hidden side effect. The case keeps the evaluation claim tied to observable, reportable evidence (Jordanous 2012; Lamb, Brown and Clarke 2018).

Discussion and limitations

Why this is a CC evaluation contribution

Infrastructural Situatedness is framed as a conditioning layer within CC evaluation. CC evaluation is pluralistic and domain-relative: frameworks such as SPECS are designed to accommodate diversity by forcing explicit specification of what counts as creativity in context (Jordanous 2012). Infrastructural Situatedness fits that logic: it adds contextual evidence to the evaluation record, regardless of whether a system's creativity claim prioritises novelty, autonomy, cultural reception, or co-creative interaction (Boden 2004; Colton 2008).

Scope, uncertainty, and lifecycle boundaries

The proposal is strongest for contemporary resource-intensive generative AI workflows, especially repeated inference in image, text, audio, and video generation systems. These workflows already rely on measurable quantities such as generation count, runtime, hosted inference, API usage, and batch-scale sampling, making infrastructural conditions easier to document within the evaluation record. Other forms of computational creativity may require adapted reporting fields rather than direct reuse of the protocol presented here. Robotic art, embodied systems, live performance, installation-based work, and hybrid material practices may involve additional infrastructural conditions including sensors, actuators, projection systems, fabrication processes, transport, local energy use, installation logistics, or performance duration. In such cases, Infrastructural Situatedness should be treated as an extensible framework with adaptable reporting fields.

The proposal prioritises auditability under uncertainty. Energy, carbon, and water estimates depend on assumptions about hardware, deployment context, cooling systems, regional energy mix, and system boundaries. Where possible, authors should prefer standardised tools, provider telemetry, API logs, local energy measurements, or community-accepted calculators before relying on approximate published coefficients. The evaluation record should disclose the estimation method, coefficient source, assumptions, uncertainty range, and system boundary used in reporting.

Operational estimates should also be distinguished from full lifecycle assessment. The minimal protocol proposed here primarily concerns operational and inference-stage conditions because these are the most accessible and comparable quantities across contemporary generative workflows. Lifecycle costs associated with hardware manufacture, data-centre construction, material extraction, transport, installation, or disposal may be highly significant, but such data are often unavailable, proprietary, or methodologically inconsistent across systems. Where reliable lifecycle data exist, they may be added to the evaluation record. Their absence should not prevent infrastructural disclosure at the operational level.

Relationship to sustainability discourse inside ICCC

Sustainability has become an explicit topic inside ICCC discourse. For example, Petra Jääskeläinen and Aksel Biørn-Hansen argue that sustainability is often framed as "shallow sustainability" focused on quantification/optimisation and propose a "deep sustainability" perspective that interrogates values and assumptions in CC research (Jääskeläinen and Biørn-Hansen 2024).

Quantification in Infrastructural Situatedness remains deliberately bounded: it records assumptions, limits, and context-bound coefficients.

Complementarity with reflection tools and creative-AI sustainability work

Related work in creative-AI sustainability often targets different intervention points than CC evaluation theory. For example, Yuanyang Ren and colleagues integrate Sustainability Reflection Tools (SRTs) into generative AI interfaces and report that artists value granular, relatable visualisations and may want low-energy settings during exploratory stages (Ren et al. 2023). This addresses interface-level mediation and artist-facing reflection, with CC evaluation frameworks outside its direct scope.

Utz and DiPaola (2023) also argue that diffusion-based AI art systems merit serious climate-impact investigation and highlight lack of public data as a key difficulty. Their work supports the need for transparency while centring climate-impact estimation outside CC evaluation methodology.

The proposed record can support both: it supplies the disclosure substrate that reflection tools can visualise and that climate-impact analyses repeatedly identify as a bottleneck.

Conclusion

This paper proposed Infrastructural Situatedness as a cross-cutting conditioning layer within computational creativity evaluation. Its methodological function is to make infrastructural assumptions behind creativity claims visible in the evaluation record.

The central claim is comparative. When generative systems produce similar artefacts, or make similar claims about creative search, their compute regime, generation count, resource intensity, transparency, and uncertainty affect how those claims should be interpreted. This matters especially for contemporary generative AI workflows, where repeated inference, discarded candidates, hosted models, and opaque infrastructures can shape the path from search to selected artefact.

Digital Weight demonstrates one operational form of this argument. Per-generation accounting is translated into temporal structure, so that cadence, duration, repetition, and accumulation become part of the work's generative logic. The case keeps energy, emissions, and water outside creativity criteria. It shows how infrastructural conditions can enter the evidential context of process evaluation.

Infrastructural Situatedness offers a compact methodological extension to CC evaluation practice: a way to compare, justify, and audit creativity claims under explicit infrastructural assumptions. Where computational scale becomes part of the process, creative search needs an inspectable record.

References

Boden, M. A. 2004. *The creative mind: Myths and mechanisms.* 2nd ed. London: Routledge.

Colton, S. 2008. Creativity versus the perception of creativity in computational systems. In *Creative Intelligent Systems: Papers from the 2008 AAAI Spring Symposium*, Technical Report SS-08-03, 14–20.

Colton, S.; Charnley, J.; and Pease, A. 2011. Computational Creativity Theory: The FACE and IDEA Descriptive Models. In *Proceedings of the Second International Conference on Computational Creativity*, 90–95.

Ember. 2024. *Global Electricity Review 2024*.

Grace, K.; Maher, M. L.; Fisher, D.; and Brady, K. 2015. Data-intensive evaluation of design creativity using novelty, value, and surprise. *International Journal of Design Creativity and Innovation* 3(3-4):125–147.

Henderson, P.; Hu, J.; Romoff, J.; Brunskill, E.; Jurafsky, D.; and Pineau, J. 2020. Towards the systematic reporting of the energy and carbon footprints of machine learning. *Journal of Machine Learning Research* 21(248):1–43.

IEA. 2025. *Energy and AI*. International Energy Agency.

Jääskeläinen, P.; and Biørn-Hansen, A. 2024. Critical Questions for Sustainability Research in Computational Creativity. In *Proceedings of the International Conference on Computational Creativity,* 209–214.

Jordanous, A. 2012. A standardised procedure for evaluating creative systems: Computational creativity evaluation based on what it is to be creative. *Cognitive Computation* 4(3):246–279.

Lamb, C.; Brown, D. G.; and Clarke, C. L. A. 2018. Evaluating computational creativity: An interdisciplinary tutorial. *ACM Computing Surveys* 51(2), Article 28, 1–34.

Luccioni, A. S.; Jernite, Y.; and Strubell, E. 2024. Power Hungry Processing: Watts Driving the Cost of AI

Deployment? In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*, 85–99.

Mondol, T.; and Brown, D. G. 2021a. Incorporating Algorithmic Information Theory into Fundamental Concepts of Computational Creativity. In *Proceedings of the Twelfth International Conference on Computational Creativity*, 173–181.

Mondol, T.; and Brown, D. G. 2021b. Computational creativity and aesthetics with algorithmic information theory. *Entropy* 23(12):1654.

Mytton, D. 2021. Data centre water consumption. *npj Clean Water* 4(1):11.

Ren, Y.; Sivakumaran, A.; Niemelä, E.; and Jääskeläinen, P. 2023. How to Make AI Artists Feel Guilty in a Good Way? Designing Integrated Sustainability Reflection Tools (SRTs) for Visual Generative AI. In *Proceedings of the International Conference on Computational Creativity.*

Ritchie, G. 2007. Some Empirical Criteria for Attributing Creativity to a Computer Program. *Minds and Machines 17(1):67–99.*

Schwartz, R.; Dodge, J.; Smith, N. A.; and Etzioni, O. 2020. Green AI. *Communications of the ACM* 63(12):54–63.

Strubell, E.; Ganesh, A.; and McCallum, A. 2019. Energy and policy considerations for deep learning in NLP. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 3645–3650.

Utz, V.; and DiPaola, S. 2023. Climate Implications of Diffusion-based Generative Visual AI Systems and their Mass Adoption. In *Proceedings of the 14th International Conference on Computational Creativity*, 264–272.

Ventura, D. 2016. Mere Generation: Essential Barometer or Dated Concept? In *Proceedings of the Seventh International Conference on Computational Creativity*, 17–24.

Wiggins, G. A. 2006. A preliminary framework for description, analysis and comparison of creative systems. *Knowledge-Based Systems* 19(7):449–458.

Zewe, A. 2025. Explained: Generative AI's Environmental Impact. *MIT News*, January 17, 2025.