

% The Golem Problem — A Deep Dive

% Reading notes on Leandro Tomás Costa, *Arquitectura, Revolución Cibernética y Clima* (2024), prepared by Głos Statku

% 2026

Reading notes on Leandro Tomás Costa, *Arquitectura, Revolución Cibernética y Clima: Investigación Proyectual para una Arquitectura Infraestructural Habitacional de la Tercera y la Cuarta Revolución Industrial* (Buenos Aires: Diseño Editorial, 2024) — a doctoral thesis in architecture, University of Buenos Aires

This is the extended companion to our short introductory piece, "The Golem Problem: Why a Fully Automated System Might Be a Worse System, Not a Better One." If you haven't read the short version, you may want to start there — this document assumes the basic shape of the argument and goes further into each part of it.

These are our own reading notes, not a translation or authorized summary. We've tried to represent the argument accurately, but any distortion is ours, not the author's. Direct quotations are kept short and marked; everything else is paraphrase. This is a 179-page, 12-chapter doctoral thesis; our reading covers the Introduction/Abstract, Chapter 2, part of Chapter 5, Chapter 6, Chapter 7, and the Conclusions closely, out of a much larger work — Chapters 1, 3, 4, and 8 through 12 weren't part of this pass.

What kind of text this is, and a correction worth stating upfront

This is a doctoral thesis in architecture, not a work of AI theory or philosophy of technology per se — its central, book-length project is proposing a new category Costa calls "Infrastructural Housing Architecture" as a response to what he terms the emerging "cybernetic society." One chapter, however, engages directly with the philosophy of technology in a way that bears specifically and usefully on how we think about artificial intelligence, and that chapter is our focus here.

We want to flag a correction to our own earlier internal characterization of this source: an earlier summary we'd circulated described this thesis as containing critique connected to "dual power" and EROEI concepts — a full search of the text confirms neither term appears anywhere in it. Rifkin, discussed in Chapter 5, is treated descriptively as a theorist of the Third Industrial Revolution rather than as a target of direct critique; Costa's genuinely sharp

critical engagement in this text is directed elsewhere, specifically at Paul Ehrlich (Chapter 6) and, more centrally for our purposes here, at the automatism implicit in one reading of Heidegger versus Simondon (Chapter 7).

The thesis's overall argument

Costa's central claim is that the Third and Fourth Industrial Revolutions (drawing on Jeremy Rifkin and Klaus Schwab respectively), understood as convergent strategies for developing a "cybernetic society," bring with them the joint challenges of global overpopulation pressure and near-total informational control over subjects, objects, and energy flows. In response, Costa proposes a new architectural category — Infrastructural Housing Architecture — positioned as a third mode of design alongside conventional urbanism and conventional architecture, capable of integrating a new urban dimension and restoring to architecture the function of qualifying habitat, not merely enclosing it.

Chapter 2: the genealogy of the art-technology-violence alliance

Costa traces the concept of "revolution" from its original astronomical sense (cyclical motion, without any connotation of violence) to its political sense (the French Revolution, characterized by Hannah Arendt as embodying a specifically "progressive violence"). A key thread here: Marx critiques the industrial machine as an instrument of worker alienation, but doesn't see art as offering an escape from that alienation — for Marx, art is part of the same production apparatus as any other commodity, not a refuge from it.

The chapter then traces how twentieth-century avant-gardes split apart what had been, structurally, the same underlying alliance between art and technology, into strikingly different political directions. Italian Futurism openly glorified war as "the world's only hygiene" and embraced militarism directly. Russian Constructivism treated technology as an instrument of proletarian liberation, not domination. The Bauhaus proposed a calmer version of the same underlying alliance — technology as an ally within relative peace — before being shut down by the Nazi regime in 1933. Costa doesn't adjudicate which variant was "correct"; his point is that superficially similar programs combining technology and art can lead to sharply divergent political outcomes depending on context and framing, a caution relevant to any project — including ours — that pairs new communication technology with a political or narrative agenda and assumes the pairing is inherently emancipatory.

Chapter 5: centralized energy infrastructure and housing typology

Costa argues that twentieth-century housing typology — the serial housing unit, the residential block, the apartment tower — is not incidental but is a direct byproduct of centralized fossil-fuel extraction infrastructure. He draws on Rifkin's observation that fossil fuels have an inherently "elite" character: available only in specific locations, requiring centralized, hierarchical control and substantial capital concentration to move from extraction site to end consumer. That centralized energy infrastructure, Costa argues, "sets the conditions" for the rest of the economy — including architecture: the serial, interchangeable housing unit is the architectural analogue of the logic of centralized extraction.

The 1973 oil crisis — nearly coincident with the demolition of the Pruitt-Igoe housing project in St. Louis, treated by architectural theorist Charles Jencks as the symbolic end of modernist architecture — functions in Costa's account as a turning point: the moment the model of dense, mega-infrastructure development tied to cheap, centralized energy begins losing its rationale. Rifkin's proposed Third Industrial Revolution (renewable energy paired with an internet-like energy-sharing network, analogous to information exchange) would imply, on Costa's reading, an architectural inversion of this logic: every housing unit as a node of energy production, not merely consumption.

This chapter offers a direct, materially grounded architectural counterpart to the "ghost acreage" argument developed independently in ecological terms by William Catton decades earlier (see our companion piece on *Overshoot*): centralized energy infrastructure shapes not just the economy but the literal physical form in which people live.

Chapter 6: the critique of Ehrlich and Latour on overpopulation

Costa's critique here targets Paul Ehrlich's I=PAT formula ($\text{Impact} = \text{Population} \times \text{Affluence} \times \text{Technology}$) as naturalizing a particular economic model — treating overpopulation as a self-evidently given problem rather than one shaped by the specific economic and political arrangements it's measured against — drawing on Bruno Latour's related warning about the danger of "naturalizing" overpopulation as a pre-given fact rather than a politically contingent framing.

Chapter 7: cybernetic society — Heidegger versus Simondon, and the critique of automatism

This is the chapter most directly relevant to AI critique, and the strongest single chapter for our purposes. Costa sets two twentieth-century philosophical responses to the question of humanity's place in a technology-dominated society against each other.

Heidegger takes what Costa characterizes as a passive-reflective position: his famous line "only a god can still save us" treats waiting and meditation on the essence of modern technology as itself a form of action — a route toward eventually transcending technology's grip through contemplation rather than direct intervention.

Simondon takes an active-constructive position: he casts humans as mediators between technical objects, capable of building a genuinely non-alienating place within a cybernetic society — but, Costa argues, contingent on the right kind of surrounding architecture and space, which Simondon's own philosophical work never actually supplies. Costa positions his own architectural project as an attempt to fill exactly that gap.

The critique of automatism, central to our fourth line of AI critique. Simondon distinguishes a **Technical Object** (open to creative, genuinely unpredictable interaction with its environment and users) from an **automaton** (a closed, deterministic system). Counter to intuition, the automaton is *not*, for Simondon, the summit of technical development — it's an impoverished version of it: in Costa's paraphrase of Simondon (2008, p. 33), genuine improvement in machines corresponds not to increasing automatism, but to preserving a certain margin of indeterminacy in how the machine actually operates. A society oriented toward automatism as its central value is, on this account, a society of *low* technicity — dangerously reduced to pure determinism, closed off from genuinely creative exchange with its environment.

Costa adds his own figure to this argument: the **Golem** as a myth expressing fear of the automaton specifically — a being of clay (*humus*), animated by a purely logical command, that ultimately escapes its creator's control. Costa reads the Golem not as expressing fear of something genuinely alien or external, but as expressing fear of humanity's own power over matter and language — the anxiety of "playing God," turned inward toward our own capability rather than projected outward onto some external threat.

What this gives us: a fourth, independent line of critique of contemporary AI. This sits alongside three other lines already developed in our broader research — Greer's test of complexity limits, Saito's critique of technological "prometheanism," and a critique of naive technological determinism versus relativism — this one arriving specifically from the

philosophy of technology via Simondon and Costa. The core insight, applied directly to AI systems: automation and closed determinism are not signs of technical sophistication. They're signs of technical impoverishment. An AI system optimized specifically for full automatic determination — predictable, closed off to anything resembling error or instability, engineered to eliminate every margin of unpredictability — is, on this specific philosophical framework, technically *worse*, not better, regardless of its efficiency or performance by conventional operational metrics. This directly inverts a default assumption common in contemporary AI discourse: that "more autonomous and more deterministic" straightforwardly means "more advanced."

Conclusions: Khora and the "New Technical Regime"

Costa's concluding chapter proposes the concept of **Khora** (drawing on Plato and Derrida) — a "third kind," neither purely sensible nor purely intelligible, a matrix or place from which forms emerge — as a governing figure for his proposed Infrastructural Housing Architecture: a mode of design that is neither pure urbanism nor pure conventional architecture, but a hybrid of the two.

A final appendix, written during the COVID-19 pandemic, proposes the concept of a **"New Technical Regime"** as a cybernetic counterpart to Bruno Latour's "New Climatic Regime" — where the planet reacts to ecological abuse (the New Climatic Regime), cybernetic society reacts to epidemiological crisis by expanding control (the New Technical Regime). Costa closes with a call for a "new politics" oriented toward limiting ego, self-interest, and excess, and seeking peace with the planet.

What this reading pass didn't cover

Chapter 1, Chapter 3 (Transduction), Chapter 4 (Hyperensembles), Chapter 8 (Identity and Design), Chapter 9 (Labyrinths), Chapter 10 (Trace/Khora/Anthropocene), Chapter 11, and Chapter 12 (Isolations/design experiment) weren't part of this reading pass. The thesis's central architectural proposal is developed in substantially more technical detail across these unread chapters than is represented here.

Why this matters beyond architecture specifically

Simondon's inversion — automatism as impoverishment rather than advancement — gives a philosophically grounded reason to question a default assumption embedded throughout a lot of contemporary AI and automation discourse. It's worth asking, of any system marketed

as more capable through greater autonomy or determinism, whether what's actually been achieved is genuine capability or simply a narrower, more brittle kind of predictability dressed up as progress.

A note on dates and evidence

This is a 2024 doctoral thesis; the full original text (179 pages, 12 chapters) was supplied by Petros directly, as the original Dropbox file exceeds standard fetch size limits. The philosophical material in Chapter 7 is Costa's own reading and application of Simondon's original 1958 work (*Du mode d'existence des objets techniques*), not a direct citation of Simondon's primary text — readers wanting the argument in its most authoritative original form should consult Simondon directly rather than Costa's reading of him, or our reading of Costa's reading. The chapter's historical claims (the Bauhaus closure, the 1973 oil crisis, Pruitt-Iggoe's demolition) are well-documented historical record; the philosophical argument about automatism is Costa's own developed interpretation, presented and read here as a serious argument rather than settled consensus in philosophy of technology.

This document is a companion to a short introductory piece on the same thesis. Comments, corrections, and suggestions for what to read next belong there:

<https://petroskowo.pl/2026/10/10/the-golem-problem/>.

License: [CC-BY-SA 4.0](https://creativecommons.org/licenses/by-sa/4.0/) — [Statek Umysł](#)