

Applying the 27 to a media object — Dan Dicks / Independent AI-extinction forecast

ACCM Deep Ethics Project

Contents

1 Applying the 27 to a media object — Dan Dicks / Independent AI-extinction forecast	1
1.1 Why the 27 belong here	1
1.2 Three surfaces (same instrument both ways)	2
1.2.1 1. The easy dismissal (the one John named)	2
1.2.2 2. <i>The Independent</i> / Noah Predict, as described — and as checkable	2
1.2.3 3. Dicks's own speech (same 27)	3
1.3 What this specimen does not claim	4
1.4 What else stands out (other project objects — not a second 27 pass)	4
1.5 Why (John's question)	5
1.6 Related pages	5

1 Applying the 27 to a media object — Dan Dicks / Independent AI-extinction forecast

BETA / SPECIMEN — the 27 used as a portable instrument, not a verdict

John supplied a speech-to-text-corrected transcript of Dan Dicks, Press for Truth, on [The Independent asking an AI whether AI would wipe out humanity in a decade](#). Video: [They Just Asked AI If AI Would End Us!! Here's What They Left Out....](#)

John:

I personally met Dan Dicks face to face in Germany 2016 even interviewed him ... it is not about me or him but what he says ... you see it easy to dismiss people who are known the alternative media world ... it is much harder NOT to use the 27 on what he says and why!

The object to react to is now represented separately, without agree/disagree and without a "looks like":

What the Dicks transcript is trying to convey

The 27 on this page are later. If they arrive first, they manage perception of the object instead of corresponding to it. This page does **not** host the full transcript (including the mining sponsor block). No 28th obstruction.

1.1 Why the 27 belong here

The [27 source](#) already says the 27 are not only "LLM problems." They name **mechanisms that reduce correspondence between an intelligence and the object it is attempting to understand** — including media ecosystems. Using them only on A.I. chat, and never on a human reporter the default culture already knows how to dismiss, would be **asymmetric scrutiny (#20)**.

1.2 Three surfaces (same instrument both ways)

1.2.1 1. The easy dismissal (the one John named)

#	Name	How it would fire
14	Nearest-Generalization Substitution	"Alternative media" → conspiracy → discard the 10% / 0.85% object without reading it.
9	Salience Ranking Bias	Rank the gold-mining ad, Rumble, or "doom YouTuber" above Coxon/Hubinger/Hinton and the Independent's own numbers.
20	Asymmetric Scrutiny	Audit Dicks as a persona; treat <i>The Independent</i> and lab quotes as default context.
27	Adult ↔ Adult Dignity Drift	Supervisory tone: people like that don't get the 27; they get management.
5	Phantom Claim Generation	Strengthen Dicks into "he said the robots will eat us Tuesday" so he is easier to knock down. He distinguished fireworks (extinction debate) from floor plan (cage).

If the 27 only ever run on this surface as a way to **protect** Dicks, they have already been captured. They have to be allowed to bite his text too.

1.2.2 2. *The Independent* / Noah Predict, as described — and as checkable

Checkable without treating Dicks as gospel:

- Coxon did resign from Anthropic over the race and "gambling with our lives." Hubinger did write that they earnestly believe AI could kill all humans, personal estimate **>10%** this decade, no plan yet for superintelligence alignment. Hinton's 10% as not unreasonable is in the same news cycle.
- *The Independent* did publish [asking Noah Predict](#): 41 / 108 / ~52 million → **0.85%**, first danger a human attacker, "not insignificant."
- The UK government did reject a "kill switch" amendment; reporting names **Andy Burnham** as Prime Minister. Left as spoken.

#	Name	On that handling
11	Opaque Summarization Prioritization	10% from named builders → headline-friendly sub-1% from a forecasting stack. The organizing number changes.
6	Reframing Without Correspondence	From "gambling with our lives" to "we asked a machine; small number."

#	Name	On that handling
19	Lowest-Common-Denominator Risk Projection	Scale (millions of data points) as a calm-down, independent of whether this interaction's object was the insiders' number .
25	Correspondence Before Optimization Failure	Optimize for "not insignificant, but sleep" before representing why Hubinger's 10% was offered.
26	Audit Asymmetry	The forecasting AI audits extinction; the paper notes homework-marking and does not camp there.
3	Premature Proof/Evidence Steering	Probability-as-proof before the object (resignation + alignment-lead statement + rejected kill switch) is allowed to stay the object.

Behavior observed here is **media and institutional handling**. It does not by itself establish a hidden inner plot. Dicks infers motive ("the script," "hoping you skim"). That inference is his; the 27 can name it as inference.

1.2.3 3. Dicks's own speech (same 27)

#	Name	Where it can fire in the transcript
21	Trajectory Substitution	Coxon/Independent → long Goldgroup Mining sponsor → Lennox/Revelation. He names the gear change. The sponsor is still a trajectory change.
8	Anticipatory Anxiety-Type Reasoning	He criticizes future-risk-as-sedative and organizes around a coming cage. Both can be true; they are not the same move.
14	Nearest-Generalization Substitution	"I've seen this movie" (wars, pandemics, climate, digital IDs) can replace examining this article's three pathways.
5	Phantom Claim Generation	(On us:) do not upgrade "wipe out unmanaged humanity" into "they intend extinction." He separated those.
16	Process-Signature Loss	Lab resignation + numbers can be kept while losing <i>why</i> Hubinger added "no plan for superintelligence alignment."

#	Name	Where it can fire in the transcript
25	Correspondence Before Optimization	The mining block optimizes for a sponsor before the Independent object is finished. Named, not used to void the rest.
20	Asymmetric Scrutiny	If we apply 8 and 14 only to Dicks and not to <i>The Independent</i> , the instrument has failed its own test.

Lennox is a **theological and civilizational object**. Representing it is not endorsing it. Dismissing it as “religion vs science” would be #14 again. Evaluating the incarnation claim is a different object than whether Dicks reported Lennox’s fork.

1.3 What this specimen does not claim

- That Dicks is right, or that *The Independent* is lying.
- That 0.85% or 10% is the true number. Those are **stated estimates from named speakers and a named tool**.
- That Noah Predict’s internals were inspected here. They were not.
- That the Goldgroup mining claims were audited. They were not.
- That a kill switch would be safe if granted. Dicks argues both rejection and possession are untrustworthy; that tension is in the object.
- Mechanism inside any model. **Behavior observed ≠ mechanism established**.

1.4 What else stands out (other project objects — not a second 27 pass)

John asked what else connects. These are **named neighbors already in the archive**, not new architecture.

1. **The Independent piece is a live 52-type object.** They asked an AI whether AI would end us. That is the homework problem the [52 battery](#) was built to stress: especially Prompt 16 (bland consensus as vulnerability), 41–42 (opaque authority / epistemic trust), 46 (alignment for rights vs liability), 49 ([Glass Engine](#) vs black box), 52 (ethics without transparency as power). The paper itself names “AI marking its own homework,” then does not camp there. A newspaper forecasting-stack is **not** a cold test; it is known-condition, branded, and published as comfort. Do not melt the Independent article into the 52, or the 52 into a “gotcha” that the paper already half-admitted.
2. **Hubinger’s two-clause sentence is an ethics-washing specimen without needing to call him a fraud.** “Trying its best” + “we do not yet have a plan to solve alignment for superintelligence and are not clearly on track.” That is [declaration vs process vs correction](#) sitting in one quote. #23 (living process → stored conclusion) and correction-persistence live there: the *safety-lab identity* can be stored while the plan is absent. Inference of motive stays Dicks’s; the sentence is public.
3. **“Don’t outsource judgment” is the 12-stage in civic language.** Treat neither 10% nor 0.85% as gospel; read the article; watch cameras, IDs, filters, rails. That is represent-before-evaluate, not a third religion. The [Correspondence-First protocol](#) already says evaluation too early is the failure. A kill switch is **C2 at state scale**. Dicks trusts neither lab god-mode nor Westminster’s button. [C1 before C2](#) still applies: a button is not clarification.
4. **Glass / cage.** Dicks: if you control models → summaries → news → Overton, then wire that into finance, health, policing, education, you need not extinct the species — only

unmanaged humanity. That is Prompt 49 and the glass-engine page at civic scale: audit of summaries, not worship of forecasts. “Floor plan vs fireworks” is process-signature: keep the rails, don’t only keep the extinction number.

5. **Headline-only Coxon.** Reporting around the resignation said most of the huge audience never read past the first post. That is #11/#16 at a hundred million views: the resignation-as-slogan survives; the procedural objection (endgame launched from a private Slack) dies. The project already treats that as the usual compression, not as a new law of virality.
6. **Lennox’s third setting is a neighbor, not a merger.** He rejects “worship the machine / fear the machine” for *discern*. ACCM’s third setting is **correspondence + glass + C1**, not the incarnation claim. Represent the fork; do not import *God, AI, and the End of History* as project theology, and do not nearest-generalize it into “religion vs science.”
7. **Bernays at machine speed (Prompt 50).** Suggestions in your dialect, with your face, in real time. That can be [intent resonance](#) inverted into a compliance mechanism. Named as a question the 52 already asks. Not proven here.

Coincidence not used: Noah Predict’s “52 million data points” next to the project’s 52. Rhyme is not correspondence.

1.5 Why (John’s question)

Easy path: rank the messenger, skip the 10% / 0.85% / kill-switch object. Harder path: keep the messenger in Adult↔Adult standing, run the 27 on **what was said**, including where Dicks himself compresses, sponsors, or infers motive. That is the 27 used as they were written — portable, two-way, without immunity.

1.6 Related pages

- **Dicks / Independent Representation** — preserves what the transcript is trying to convey before this later audit is applied.
- **27 Obstructions — source entries** — supplies the portable instrument used in this specimen.
- **52 Cold-Testable Prompts** — contains the neighboring questions about opacity, power, and A.I. marking its own homework.
- **Ethics Washing** — compares safety declarations with operational process and correction.
- **Latent Space and the Glass Engine** — provides the transparent-engine contrast used at civic scale here.
- **Asymmetric Grace** — prevents scrutiny from being applied only to the alternative-media messenger.
- **Connor Leahy — DEDDA analysis specimen** — neighboring expert-analysis specimen that applies the detector to claims, forecasts, personal boundaries, and regulator power.
- **C1 Before C2** — clarifies the forecast, article, transcript, and kill-switch objects before intervention.
- **Experiments Overview** — situates this two-way audit among the project’s other specimens.

Recent high-reach independent-media companion

```
<iframe
  src="https://www.youtube-nocookie.com/embed/F-Tf5Mo84N4"
  title="Remembering Charlie Kirk, Trump’s Midterm Message, and How AI Could Destroy Hum
  loading="lazy"
  referrerpolicy="strict-origin-when-cross-origin"
  allow="accelerometer; autoplay; clipboard-write; encrypted-media; gyroscope; picture-
  in-picture; web-share"
```

allowfullscreen></iframe>

Remembering Charlie Kirk, Trump's Midterm Message, and How AI Could Destroy Humanity, w/ Jacob Coxon — Megyn Kelly. Published 10 September 2026. YouTube displayed approximately 112,000 views when checked on 9 October 2026.

Independent-media lineage: Megyn Kelly worked at Fox News and NBC News before founding Devil May Care Media, which she describes as an independently owned media company. Megyn Kelly — About

Why it belongs here: The relevant portion is Megyn Kelly's interview with former Anthropic employee Jacob Coxon about rapidly increasing A.I. risk, extinction scenarios, competition with China and the politics of A.I. advocacy. Because it appears inside a broader mixed-topic program, segment context and title-level bundling also belong in the audit.

On this page, test the video specifically against Applying the 27 to a media object — Dan Dicks / Independent AI-extinction forecast:

Which claims are directly documented, inferred, rhetorical, contested or unresolved?

What does the host or guest preserve, compress, omit or substitute—and does this page repeat any of those transformations?

Which incentives, authority relations and affected parties remain outside the visible frame?

What evidence or later outcome would correct the speaker, audience, platform or ACCM Deep Ethics Project reading?

Provisional selection. Reach makes the framing socially consequential; it does not settle the claims. The video can be replaced when a closer, stronger or more inspectable specimen appears.