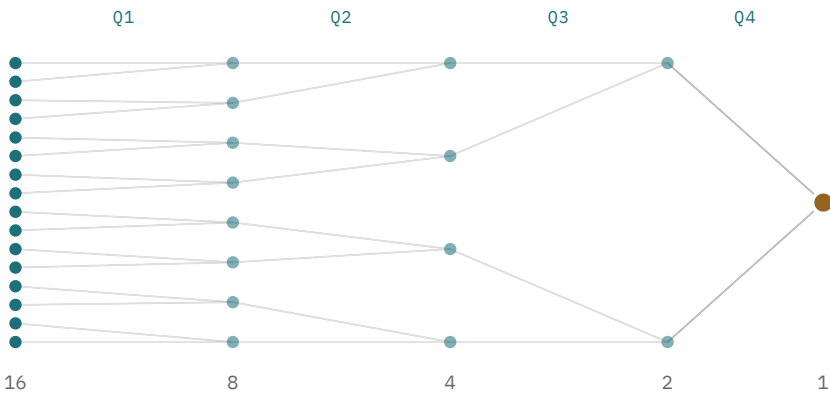


# Information *Theory*

*How many yes/no questions is a fact worth? And what does it cost, in heat, to forget one?*



Four good questions pin down one of sixteen things.  $\log_2 16 = 4$  bits.

**P**ick a whole number between 1 and 16 and hold it in your head. I'll find it with four questions, and you may only answer yes or no. "Is it 9 or higher?" – that splits the field in half. "Within that half, is it in the upper part?" – half again. Two more cuts and I have your number, every time, with room to spare. Notice what just happened: a fact that felt fuzzy and personal turned out to have an exact *size*. It was worth four questions – no more, no fewer. Four **bits**.

That a piece of knowledge can be *weighed* – measured in a unit as concrete as a kilogram – is one of the strangest and most consequential ideas of the twentieth century. It arrived almost fully formed in 1948, in a paper by a unicycle-riding

Bell Labs engineer named Claude Shannon. And it carries a sting in its tail that took until the 1960s to feel and until 2012 to measure in a lab: if information is a real quantity, then *erasing*<sup>1</sup> it is a physical act with a non-negotiable price, paid in heat. Today we earn both halves of that sentence.

We've spent six days building an epistemic toolkit; today we discover it has a *currency*. On **Day 4** we met *surprise* – Monty Hall's host opening a goat door was information, not noise – and the e-value as evidence you could literally bank. Today that intuition gets its unit. On **Day 3**, Boole's logic became physical switches (Shannon's *other* famous paper, 1938); today the same man makes *information* physical too. And the graded belief of **Day 1** – credence, the Bayesian brain minimizing Friston's "free energy" – turns out to be minimizing exactly the quantity we define here: expected surprise. The *information* thread, traced quietly since Day 1, finally gets its hard number. Keep it ready: it reappears as the arrow of time on **Day 33** and as the thing life itself seems to defy on **Days 83–85**.

---

## THE MODEL

### A bit is a halving

Before Shannon, "information" was a word for newspapers and telegrams – content, meaning, gossip. Shannon's radical move was to throw the *meaning* away. For an engineer trying to push messages down a noisy wire, what matters is not what a message *says* but how much it could have said: how much uncertainty it removes. Strip out meaning and you're left with something you can count.

The unit is the *bit*<sup>2</sup> – short for "binary digit," a contraction coined by Shannon's colleague John Tukey in a 1947 Bell Labs memo and credited by name in Shannon's paper. One bit is the amount of information in a single fair yes/no answer: the resolution of one perfectly balanced uncertainty. Two equally likely possibilities, one bit. Sixteen, four bits, because  $2^4 = 16$ . The bit is a *halving*.

But real choices aren't always balanced. The genius of the theory is how it handles loaded dice. Shannon defined the information content – the *surprise*<sup>3</sup> – of an outcome with probability  $p$  as:

$$\text{surprise}(p) = -\log_2 p \text{ bits}$$

A sure thing ( $p = 1$ ) carries zero surprise. A one-in-a-million event carries about twenty bits.

It clicks the moment you test it. A coin you *know* is two-headed: calling "heads" right tells you nothing –  $p = 1$ , surprise =  $-\log_2 1 = 0$  bits. A fair coin:  $p = 1/2$ , surprise =  $-\log_2 1/2 = 1$  bit, the textbook halving. A weather forecaster who says "100% chance of rain" in a desert, and is right, has told you almost nothing; the same words in a place where rain is rare carry real information. **Rare outcomes are surprising; surprising outcomes are informative.** The logarithm is what makes surprises *add up* the way intuition demands: learn two independent facts and the surprises sum, just as the possibilities multiply.

## Entropy: the average surprise

Now zoom out from one outcome to a whole source – a language, a die, a stream of symbols. How surprising is it *on average*? That average is the crown jewel of the theory, the *Shannon entropy*<sup>4</sup>:

$$H = -\sum p_i \log_2 p_i \text{ bits per symbol}$$

Each outcome's surprise, weighted by how often it happens. The expected number of yes/no questions per symbol.

Entropy is the irreducible core of a message – the true number of questions you'd need, on average, to nail down each symbol with the smartest possible strategy. A fair coin has entropy 1; a fair eight-sided die, 3; the letter that follows "q" in English, almost 0 (it's nearly always "u," so you barely need to ask). It is, in a precise sense, the **amount of genuine choice** a source exercises, or equivalently the amount of your uncertainty it resolves. The same quantity, read from two ends.

## THE MOST USEFUL WORD HE NEVER QUITE CHOSE

Shannon's  $H$  has the exact algebraic shape – a sum of  $p \log p$  – of a quantity physicists had used since the 1870s to measure disorder: *entropy*. The story, told decades later by Myron Tribus, is that Shannon was unsure what to call his new measure, and John von Neumann told him to call it entropy for two reasons – the formula already had that name in statistical mechanics, and "nobody knows what entropy really is, so in any debate you'll have the advantage." It's probably too good to be literally true (it surfaces in print only in 1971). But the coincidence it points at is real, deep, and still argued over – and it's the hinge the entire back half of this course will swing on. Hold that thought.

The reference table below shows the same entropy pattern without a live dial: maximum uncertainty carries the most information, while near-certainty carries almost none.

## The Entropy Dial

SOURCE	P (HEADS)	ENTROPY	READING
Fair coin	0.50	<b>1.00 bit/flip</b>	Every flip answers a full yes/no question.
Loaded coin	0.88	<b>0.53 bits/flip</b>	The common face is cheap to encode; the rare face is expensive.
Near-certain source	0.99	<b>0.08 bits/flip</b>	The outcome is almost known in advance, so little information arrives.
"q" followed by "u"	0.95	<b>0.29 bits/symbol</b>	Language compresses because many symbols are strongly predictable from context.

### — WHY IT MATTERED

## The theorem that built the modern world

Defining information would have been a nice piece of bookkeeping. What made Shannon's 1948 paper – modestly titled "*A Mathematical Theory of Communication*" and later dubbed the "Magna Carta of the Information Age" – a foundation stone was a single staggering result about *noise*.

Every real channel corrupts its messages: static on a line, scratches on a disc, cosmic rays flipping bits in deep space. The folk wisdom of 1948 was that noise set a grim trade-off – to communicate more reliably, you had to slow down, and perfect reliability meant a crawl toward zero. Shannon proved the folk wisdom *wrong*. Every channel, he showed, has a fixed *capacity*<sup>5</sup>  $C$ , a ceiling in bits per

second. As long as you transmit *below*  $C$ , you can drive your error rate as close to zero as you like – not by shouting louder or going slower, but by encoding cleverly, wrapping your message in just enough mathematical redundancy to let the receiver reconstruct it perfectly. Above  $C$ , reliable communication is flatly impossible.

*There is a hard wall called capacity. Below it, near-perfect communication is always achievable. The only question is whether we're clever enough to find the code.*

Here's the kicker: Shannon proved the good codes *exist* without saying how to build them. He left engineers a treasure map with an X but no path. Chasing that X became one of the great quests of applied mathematics – Reed–Solomon codes (which armor your CDs, QR codes, and the data beamed back from Mars), then turbo codes (1993), then the low-density parity-check codes now humming inside Wi-Fi and 5G. Each crept closer to Shannon's wall. Every time you stream a movie over a flaky connection without a single glitch, you are watching a sixty-year-old theorem be cashed in. The shorter codeword for the commoner symbol – Morse's single dot for "E," the logic behind Huffman coding (1952) – is the same principle running underneath: spend your bits where the surprise is.

---

— THE DEBATE

## Is information *physical*?

So far, information sounds like mathematics – abstract, weightless, the stuff of probability and logarithms. A bit seems no more physical than the number 7. For a long time that was the consensus. And then a paradox more than a century old forced the issue, and the answer turned out to be **no, a bit is not weightless – and forgetting one warms the room.**

## A demon at the trapdoor

In 1867 James Clerk Maxwell dreamed up a troublemaker. Picture a box of gas split by a wall, with a tiny trapdoor and a tiny intelligent being – later christened *Maxwell's demon* – guarding it. The demon watches the molecules. When a fast one approaches from the right, it opens the door and lets it through to the left; when a slow one approaches from the left, it lets that through to the right. It never does any work on the molecules – just opens and shuts a frictionless door at the right moments. Slowly, patiently, it sorts hot from cold, building a temperature difference out of a uniform gas.

This should be impossible. Building order from equilibrium, for free, is exactly what the *second law of thermodynamics*<sup>6</sup> forbids – it's the law that says coffee cools, eggs don't unscramble, and entropy never spontaneously falls. The demon seems to break the deepest bookkeeping rule in physics using nothing but *information* about which molecules are which. For a hundred years it haunted the field. Leó Szilárd sharpened it in 1929 down to a single molecule in a box, and showed the demon could extract a tidy packet of work – exactly  $kT \cdot \ln 2$  – from one bit of "which side is it on?" knowledge. The arrow pointed somewhere uncomfortable: information could apparently be *converted into energy*.

## The twist: it's not knowing, it's forgetting

The resolution is one of the most beautiful pieces of reasoning in twentieth-century physics, and it came from the people who built computers. In 1961, IBM's **Rolf Landauer** asked a question nobody had thought to ask: is computation *necessarily* dissipative? Must shuffling bits around always cost energy? His surprising answer: *no* – almost every logical step can in principle be done with arbitrarily little energy, run as slowly and gently as you like. *Almost*. There is exactly one exception, and it is *erasure*.

Erase one bit  $\rightarrow$  dissipate at least  $kT \cdot \ln 2$   
 $\approx 2.8 \times 10^{-21}$  joules (2.8 zeptojoules  $\approx$  0.018 eV) at room temperature. *Landauer's principle*<sup>7</sup>, 1961.

Why erasure specifically? Because erasure is the one logical operation you can't run backward. If I tell you a bit is now "0," you cannot recover whether it was "0" or "1" a moment ago – that history is gone, two possible pasts crushed into one present. *Logically irreversible operations*<sup>8</sup> destroy distinctions, and in a physical device, distinctions live in physical states. Crush two states into one and the "lost" possibility has to go somewhere; it flows out into the surrounding world as a minimum dollop of heat,  $kT \cdot \ln 2$  per bit. Landauer's slogan became a rallying cry: **"Information is physical."**

In 1982 his IBM colleague **Charles Bennett** closed the trap on Maxwell's demon with this insight. The demon's mistake was never *measuring* the molecules – measurement, Bennett showed, can be done reversibly, for free. The demon's mistake is that it has a *memory*, and that memory fills up. To keep sorting forever, it must eventually erase old observations to make room for new ones. And each erasure pays back, as heat, precisely the entropy the demon thought it was removing from the gas. The books balance to the last joule. The second law was never in danger; the demon was just running up a tab in a ledger nobody had been reading. The cost of the demon's cleverness isn't thinking – *it's forgetting*.

The worked example below follows the same erasure cycle as a static sequence: protect a bit, remove its distinction, force it to 0, and pay the heat cost.

## The Landauer Erasure Machine

STEP	LOGICAL STATE	PHYSICAL MOVE	THERMODYNAMIC READING
0	Bit can be 0 or 1	A barrier separates two stable wells.	No heat cost is forced by the logic.
1	Old value becomes unprotected	Lower the barrier between wells.	This can be done reversibly in principle.
2	Drive toward 0	Tilt the landscape so both possible starts end left.	Thermal entropy is exported to the surroundings.
3	Bit reads 0	Raise the barrier again.	Two possible pasts have become one present.
4	One bit erased	Reset complete.	At least $kT \ln 2$ of heat, about 2.8 zJ at room temperature, has been dissipated.

— THE FRONTIER · 2026

## From a thought experiment to a zeptojoule on a bench

For fifty years Landauer's principle was a piece of theory so clean it felt almost philosophical. The numbers were absurd – a few *zeptojoules*, billionths of a billionth of a joule, lost in the thermal roar of any real device. Measuring it seemed hopeless. Then, beginning in 2012, a string of breathtaking experiments dragged the principle out of the realm of argument and onto the lab bench, across

system after system. Here is where the hype filter earns its keep – and here, unusually, the boldest-sounding claims are the ones that *survived*.

---

Edge 01 [ESTABLISHED]

## The bound, measured five ways

The first direct confirmation came from **Bérut and colleagues** (*Nature*, 8 March 2012): a single micron-sized glass bead held in a double-well optical trap – a physical one-bit memory – was erased over and over, and the average heat released settled onto the  $kT \cdot \ln 2$  floor as the erasure was done more gently. Two years later **Jun, Gavrilov & Bechhoefer** (*Physical Review Letters*, 4 Nov 2014) pushed precision higher in a feedback trap, confirming that halving the number of accessible states costs at least  $kT \cdot \ln 2$  – with individual cycles dipping *below* the bound, exactly as the fluctuation theorems (a Day-85 preview) predict, while the *average* holds firm.

What makes it [ESTABLISHED] is the sheer diversity of confirmations. A single **electron** in a box run as a Szilárd engine (Koski et al., *PNAS*, 2014). An array of nanoscale **magnets** – the closest thing yet to a real digital memory bit – pinned the cost near "2.8 zJ at 300 K," measuring  $(4.2 \pm 0.9)$  zJ (Hong, Lambson, Dhuey & Bokor, *Science Advances*, 11 March 2016). And a single trapped **calcium ion** extended the principle into the fully quantum regime (Yan et al., *Physical Review Letters*, 21 May 2018). Glass, electrons, magnets, atoms – wildly different stuff, one identical floor. That's what a real law looks like.

---

Edge 02 [ESTABLISHED] [REVIEW]

## Running the demon in reverse: information as fuel

If erasing a bit *costs* energy, can measuring one *buy* energy? Szilárd said yes on paper; the lab now says yes in fact. **Toyabe and colleagues** (*Nature Physics*, 2010)

built the first real information engine: a Brownian particle climbing a staircase, lifted against gravity using nothing but well-timed measurements of its position and a feedback ratchet – converting pure information into mechanical work, and validating a generalized form of the Jarzynski equality in the process. **Koski's single-electron Szilárd engine** (2014–2015) did it with one electron and even built an "information-powered refrigerator." More recently, the **Bechhoefer lab** (Saha et al., *PNAS*, 18 May 2021) optimized a colloidal information ratchet whose output power, they report, rivals the molecular machinery inside living cells – a genuinely striking result, though the "world's fastest information engine" billing comes from the press release, not the peer-reviewed claim, so treat that phrase as marketing, the underlying physics as solid.

---

Edge 03 [ESTABLISHED] [REVIEW]

## The new physics of information — and what it says about your GPU

All of this now lives inside a mature framework called *stochastic thermodynamics*, which extends the old laws of heat to tiny, jittering systems where fluctuations dominate. Its engine room is a pair of exact results – the **Jarzynski equality** (1997) and the **Crooks fluctuation theorem** (1999) – that let you write the second law as an *equality* with a correction term, and that correction term is *information* (mutual information, to be exact). The authoritative synthesis is Parrondo, Horowitz & Sagawa's review, bluntly titled "*Thermodynamics of information*" (*Nature Physics*, 2015): information sits in the ledger of physics on equal footing with work and heat. In 2024, David H. Wolpert and collaborators at the Santa Fe Institute extended this to realistic finite-time computation (*Physical Review X*, 13 May 2024), quantifying the "mismatch cost" – how much any real computer must burn *above* the Landauer minimum.

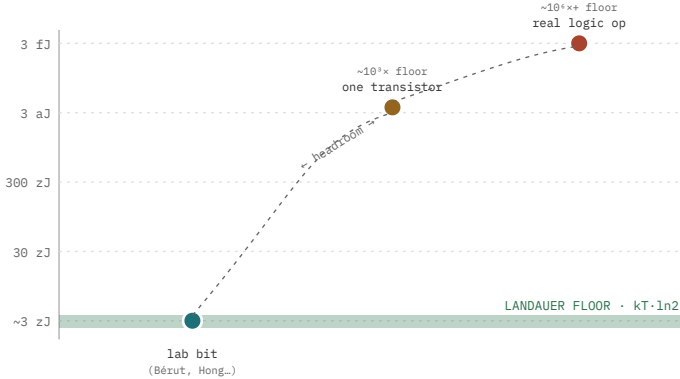
Which raises the question lurking behind every data center. How far above the floor are our machines? **Very**. A single transistor switch dissipates on the order of  $10^{18}$  joules – hundreds to thousands of times the Landauer bound at the device

level, and millions of times more once you count memory traffic, cooling, and power conversion. We are nowhere near the wall. That's the honest framing: the energy appetite of large-scale AI – data centers drew about 1.5% of global electricity in 2024, on track to roughly double by 2030 (IEA, *Energy and AI*, April 2025) – is, for now, an **engineering and economic** problem, not a fundamental-physics one. There are several orders of magnitude of headroom before the laws of thermodynamics, rather than the laws of budgets, become the binding constraint. [REVIEW] – but the floor is real, it is getting closer every hardware generation, and it is why reversible and neuromorphic computing are suddenly interesting again.

DIAGRAM · ORDERS OF MAGNITUDE

# The Landauer ladder — how far we are from the floor

Energy dissipated per bit operation, on a logarithmic scale (each step is  $\times 10$ ). The experiments above kissed the floor; the chips in your pocket sit a thousand-fold higher; a full system, higher still. The gap is the room we have left to improve.



The lab experiments aren't *better engineered* than your laptop — they're run unimaginably slowly, one bit at a time, precisely to approach the thermodynamic limit. Speed costs energy; the floor is only reached in the limit of infinite patience. Real computing trades that patience for billions of operations per second, and pays in heat.

Edge 04 [CONTESTED]

## Where the idea overreaches

"Information is physical" is established. The temptation is to sand off the qualifier and declare that information is *fundamental* — that reality is, at bottom, made of bits. The physicist John Archibald Wheeler gave this its slogan in 1989: "*it from*

*bit*," the conjecture that every particle and field derives its very existence from yes/no answers, from information. It is a gorgeous, generative idea – and it is metaphysics, not a tested result. Critics note the obvious circularity: a bit has to be encoded in *something*, so information can't be the bottom turtle. Keep it in the "stimulating speculation" drawer, clearly labeled.

Further out lies genuine fringe. Melvin Vopson's "mass–energy–information equivalence" claims information has rest mass; his "second law of infodynamics" claims information entropy must *decrease* over time; his 2025 paper deriving gravity from information drew a flat verdict from physicist Sabine Hossenfelder that it "makes no sense." When a frontier is hot, it grows a fringe – and learning to tell the two apart, using exactly the hype filter we've run all course, is the skill this whole project is really about. (Per the syllabus's standing rule: any citation to a future-dated preprint is treated as fabricated and discarded on sight.)

#### — OPEN QUESTIONS

## What's genuinely unsettled

- **Is Shannon entropy the *same thing* as thermodynamic entropy, or just shaped like it?** The formula is identical; whether that's a deep physical identity (the maximum-entropy view of Jaynes) or a profound analogy is still argued. This is the question Day 33 and Days 83–85 will reopen with the stakes raised to "what is life?"
- **Can the demon ever truly be beaten?** A minority of physicists (Earman & Norton) argue the standard exorcism is subtly circular – using the second law to derive the erasure cost, then using that cost to defend the second law. A contested 2016 result even claimed a logically irreversible gate run below  $kT \ln 2$ . The mainstream, backed by the experiments, says no – but the foundations aren't fully closed.
- **How low can real computation actually go?** Reversible computing promises to dodge erasure costs almost entirely. Nobody has built a useful machine that does. Is the Landauer floor a practical target or a permanent curiosity?

- **And the question waiting in the AI block:** when a model "knows" something, is that knowledge ultimately just bits arranged to reduce a loss – and does the thermodynamic cost of those bits tell us anything about what it would take to *think*? (Days 138–145.)

## ◆ THE DAY IN THREE SENTENCES

## BIG IDEA

Information is a measurable quantity — the bit, defined as expected surprise  $H = -\sum p \log_2 p$  — and it is not abstract but *physical*: erasing one bit must dissipate at least  $kT \cdot \ln 2$  of heat (Landauer), which is exactly how Maxwell's demon is exorcised and the second law saved.

## BEST ANALOGY

A fact's size is the number of yes/no questions it's worth; and a memory is a ball in a double well, so resetting it to a standard state — crushing two possible pasts into one present — is the act that leaks heat into the room.

## LIVE CONTROVERSY

Whether Shannon's entropy and thermodynamic entropy are one thing or two; whether the demon is truly beaten; and how far the grand claim "information is fundamental" (Wheeler's "it from bit") overreaches into hype.

---

THREADS TODAY › information (gets its hard unit at last — the bit, surprise, entropy) · energy (Landauer's  $kT \cdot \ln 2$  ties bits to heat) · computation (the thermodynamic cost of erasing, the demon as a memory device) — with first hints of emergence (a law that holds across glass, electrons, magnets, atoms) and a setup for evolution & life on Days 83–85.

TOMORROW → DAY 08

## Complexity & Emergence

Today a single number captured a whole source. Tomorrow we ask what happens when simple parts, following simple rules, conjure behavior that none of them contains – a murmuration of starlings wheeling like one animal, a market, a mind. We'll separate *weak* emergence (surprising but derivable) from *strong* (genuinely irreducible?), meet the brand-new attempts to *measure* complexity – and apply today's hardened hype filter to Assembly Theory, a 2020s claim that's drawn both excitement and serious fire. Bring your bits; complexity is, in part, information that resists compression.

---

### NOTES

1. Erasure means resetting a physical memory to a standard state, so information about its previous state is lost.
2. A bit is the amount of information carried by one perfectly balanced yes/no distinction.
3. In information theory, surprise is the information carried by a particular outcome, measured as negative log probability.
4. Shannon entropy is the expected surprise of a source: the average information per symbol.
5. Channel capacity is the maximum reliable information rate a communication channel can support under its noise conditions.
6. The second law says, roughly, that total entropy cannot decrease in an isolated system.
7. Landauer's principle says irreversible erasure of one bit must dissipate at least  $kT \ln 2$  of heat.
8. A logically irreversible operation maps multiple possible inputs to one output, so the input cannot be reconstructed from the output alone.

### EVIDENCE

## Sources & further reading

1. Shannon, C. E. (1948). "A Mathematical Theory of Communication." *Bell System Technical Journal* 27: 379–423 (July) & 623–656 (October). – the founding paper: bit, entropy,

- channel capacity, noisy-channel coding theorem. doi.org/10.1002/j.1538-7305.1948.tb01338.x
2. Shannon, C. E. (1938). "A Symbolic Analysis of Relay and Switching Circuits." *Trans. AIEE* 57(12): 713–723 (MIT master's thesis, 1937). – Boolean logic in physical switches; the Day-3 bridge.
  3. Soni, J. & Goodman, R. (2017). *A Mind at Play: How Claude Shannon Invented the Information Age*. Simon & Schuster. – biography; the "Magna Carta" framing and the Tukey "bit" coinage.
  4. Landauer, R. (1961). "Irreversibility and Heat Generation in the Computing Process." *IBM Journal of Research and Development* 5(3): 183–191. – the  $kT \ln 2$  erasure bound. doi:10.1147/rd.53.0183. doi.org/10.1147/rd.53.0183
  5. Bennett, C. H. (1982). "The Thermodynamics of Computation – a Review." *International Journal of Theoretical Physics* 21(12): 905–940. doi:10.1007/BF02084158; erasure, not measurement, exorcises Maxwell's demon. doi.org/10.1007/BF02084158
  6. Szilárd, L. (1929). "Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen." *Zeitschrift für Physik* 53: 840–856. doi:10.1007/BF01341281; the single-molecule engine; one bit  $\leftrightarrow kT \ln 2$  of work. doi.org/10.1007/BF01341281
  7. Bérut, A., Arakelyan, A., Petrosyan, A., Ciliberto, S., Dillenschneider, R. & Lutz, E. (2012). "Experimental verification of Landauer's principle linking information and thermodynamics." *Nature* 483: 187–189 (8 March 2012). doi:10.1038/nature10872. doi.org/10.1038/nature10872
  8. Jun, Y., Gavrilo, M. & Bechhoefer, J. (2014). "High-Precision Test of Landauer's Principle in a Feedback Trap." *Physical Review Letters* 113: 190601 (4 Nov 2014). doi:10.1103/PhysRevLett.113.190601. doi.org/10.1103/PhysRevLett.113.190601
  9. Hong, J., Lambson, B., Dhuey, S. & Bokor, J. (2016). "Experimental test of Landauer's principle in single-bit operations on nanomagnetic memory bits." *Science Advances* 2(3): e1501492 (11 March 2016). doi:10.1126/sciadv.1501492; measured  $(4.2 \pm 0.9)$  zJ; "2.8 zJ at 300 K." doi.org/10.1126/sciadv.1501492
  10. Yan, L. L. et al. (2018). "Single-Atom Demonstration of the Quantum Landauer Principle." *Physical Review Letters* 120: 210601 (21 May 2018). – a trapped  $^{40}\text{Ca}^+$  ion; the quantum regime. link.aps.org/doi/10.1103/PhysRevLett.120.210601

11. Toyabe, S., Sagawa, T., Ueda, M., Muneyuki, E. & Sano, M. (2010). "Experimental demonstration of information-to-energy conversion and validation of the generalized Jarzynski equality." *Nature Physics* 6: 988–992. doi:10.1038/nphys1821. doi.org/10.1038/nphys1821
12. Koski, J. V., Maisi, V. F., Pekola, J. P. & Averin, D. V. (2014). "Experimental realization of a Szilard engine with a single electron." *PNAS* 111(38): 13786–13789. doi:10.1073/pnas.1406966111. doi.org/10.1073/pnas.1406966111 See also Koski et al., *PRL* 113: 030601 (2014) and *PRL* 115: 260602 (2015).
13. Saha, T. K., Lucero, J. N. E., Ehrich, J., Sivak, D. A. & Bechhoefer, J. (2021). "Maximizing power and velocity of an information engine." *PNAS* 118(20): e2023356118 (18 May 2021). doi:10.1073/pnas.2023356118; optimized colloidal information ratchet ("fastest" is press framing). doi.org/10.1073/pnas.2023356118
14. Jarzynski, C. (1997). "Nonequilibrium Equality for Free Energy Differences." *Physical Review Letters* 78: 2690. doi:10.1103/PhysRevLett.78.2690. doi.org/10.1103/PhysRevLett.78.2690 · Crooks, G. E. (1999). "Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences." *Physical Review E* 60: 2721. doi:10.1103/PhysRevE.60.2721. doi.org/10.1103/PhysRevE.60.2721
15. Parrondo, J. M. R., Horowitz, J. M. & Sagawa, T. (2015). "Thermodynamics of information." *Nature Physics* 11: 131–139. doi:10.1038/nphys3230. doi.org/10.1038/nphys3230
16. Manzano, G., Kardeş, G., Roldán, É. & Wolpert, D. H. (2024). "Thermodynamics of Computations with Absolute Irreversibility, Unidirectional Transitions, and Stochastic Computation Times." *Physical Review X* 14: 021026 (13 May 2024). doi:10.1103/PhysRevX.14.021026; the "mismatch cost" above the Landauer floor. doi.org/10.1103/PhysRevX.14.021026
17. International Energy Agency (2025). *Energy and AI*. (April 2025). – data-center electricity ≈1.5% of global demand in 2024 (~415 TWh), projected ~945 TWh by 2030. [iea.org/reports/energy-and-ai](https://www.iea.org/reports/energy-and-ai)
18. Wheeler, J. A. (1990). "Information, Physics, Quantum: The Search for Links." In *Complexity, Entropy, and the Physics of Information*. – "it from bit" (CONTESTED metaphysics).
19. Vopson, M. M. (2023). "The second law of infodynamics and its implications for the simulated universe hypothesis." *AIP Advances* 13: 105308. doi:10.1063/5.0173278; CONTESTED/HYPE. doi.org/10.1063/5.0173278 See also Vopson (2025), "Is gravity evidence of a computational universe?", *AIP Advances* 15: 045035,

doi:10.1063/5.0264945, and Sabine Hossenfelder's 2025 public critique.  
[doi.org/10.1063/5.0264945](https://doi.org/10.1063/5.0264945)

END OF DAY 07 · 173 DESCENTS REMAIN