kynetic

WHITE PAPER

Kept Thinking

Dated human directions for an age of machine science

DOCUMENT	KW-01
VERSION	0.1
DATE	26 November 2026
STATUS	Draft for review
AUDIENCE	Investors, laboratory partners and researchers

The research and licensing programme described here is planned, not live. Nothing in this paper is an offer. kynetic.cc · hello@kynetic.cc

01

Summary

IN PLAIN TERMS

- Machines are making experiments fast and cheap. Choosing what to test is becoming the scarce part of science.
- People already have good ideas, every working day, and almost nobody keeps a dated record of them.
- Kynetic keeps that record for free, on the person's own computer, with their words kept apart from any machine's.
- Later, with separate consent, labs may pay to read a record where it is kept. The person is paid, and paid again when an idea holds.

Most people expect machines to take over science: to choose the experiments as well as run them, so that human thinking matters less. We think the opposite is more likely. As machines make each experiment cheaper, the step before the experiment, deciding what is worth testing, becomes the expensive part. Human judgement about where to look will matter more, not less.

The evidence so far points the same way. Asked for research ideas, today's models produce long lists that soon repeat themselves, and ideas that sound new often fail once they are tested.[1,2] Labs that can run experiments around the clock still need someone to tell them where to start. Meanwhile, much of the thinking that would tell them now happens inside chat windows that belong to the companies that make the models, not to the people doing the thinking.

Kynetic is a free app that keeps a person's thinking as a dated record that stays theirs. It brings together their chats with AI tools, their notes, readings and drafts, dates every entry, and keeps the person's own words apart from a machine's and from other people's. It grows that record into a living picture, a neuron, that the person can explore. Students and academics come first, because they already think with machines every day, and the question they worry about most is the one our record answers: which parts of this are mine?

Later, and only with separate consent, the same record can earn. We plan a programme in which a lab's software may read a consenting person's record inside a sealed environment, for one purpose the person allows, and take away only a few checked directions: ideas about what to test next. The lab pays to read. When a direction holds, meaning a test agreed in advance proves it right, the lab pays milestones and a royalty, and the person shares in all of it. No record is ever sold, and no outside company ever trains a model on one.

The bet is lopsided. The app is cheap to run because the record stays on the person's computer. Nothing is sold below cost. A blind test, set up before any price, will tell us early if records do not beat cheaper sources of ideas, and we will publish the result either way. If we are right, the advantage compounds: a date cannot be added later, so every record kept from today carries dates no later rival can match.

Nothing in the programme exists yet. No record is held, nothing has been sold, and the legal, ethical and security reviews it needs have not happened. This paper sets out the case and the plan, so that anyone can hold us to it.

02

The problem, and the words we use

IN PLAIN TERMS

- A self-driving lab runs experiments by itself, but still has to be told where to look.
- We call that starting idea a direction. It holds when a test agreed in advance proves it right.
- Kynetic keeps the record that directions come from, and later may let labs read it, never take it.

In a self-driving laboratory, software plans experiments and robots carry them out, day and night. In 2023 one such lab at Lawrence Berkeley National Laboratory reported making 41 of 58 target compounds in 17 days of continuous work.[3] Other scientists then went back through the results and disputed whether any of them were genuinely new materials.[4] Both halves of that story matter. Machines can now do the work of testing at a pace no team of people can match. They still need people to choose what is worth testing and to judge what the results mean.

This paper uses a few words in fixed senses.

WORD	WHAT IT MEANS HERE
Record	Everything a person chooses to keep: notes, drafts, readings, questions, chats with machines and dead ends, each entry dated when it is written.
Originator	The person whose record it is.
Direction	An idea about what is worth testing next.
Holds	A direction holds when a test agreed before reading shows it was right, as shown by an outside document.
Custody	Holding a record for its owner, as a bank holds a safe-deposit box. It is read where it is kept and never handed over.
Reading room	The sealed part of custody where a lab's software reads a record. Only checked directions leave it.
Screen	Machine checks that seal other people's words, health details and anything uncertain before a record is read, and check every direction before it leaves.
Single	A licence to read one person's record, for one purpose and one period.
Pool	Many consenting records read together under code numbers, for one purpose. Only material dated before the pool is announced counts.
Open call	A problem a sponsor posts publicly. Anyone may answer with a dated direction.
Trial read	A short paid reading, scored blind, before a lab commits to a seat.
Seat	One reading place in a licence, for one agent.
Fingerprint	A short public code that proves a sealed block of entries existed on a date, without revealing any of it.

Our thesis is simple. **Labs that can test cheaply will pay for human directions worth testing, and custody is the only way to sell a direction without handing over the record it came from.**

Today's trade in human thought sells copies. Text is scraped or licensed in bulk to train models, and once it is copied, nobody can tell any more who had which idea. A copy can be sold once. A record that is only ever read in place, and never handed over, can be read again and again, for different purposes, and the person can be paid each time.

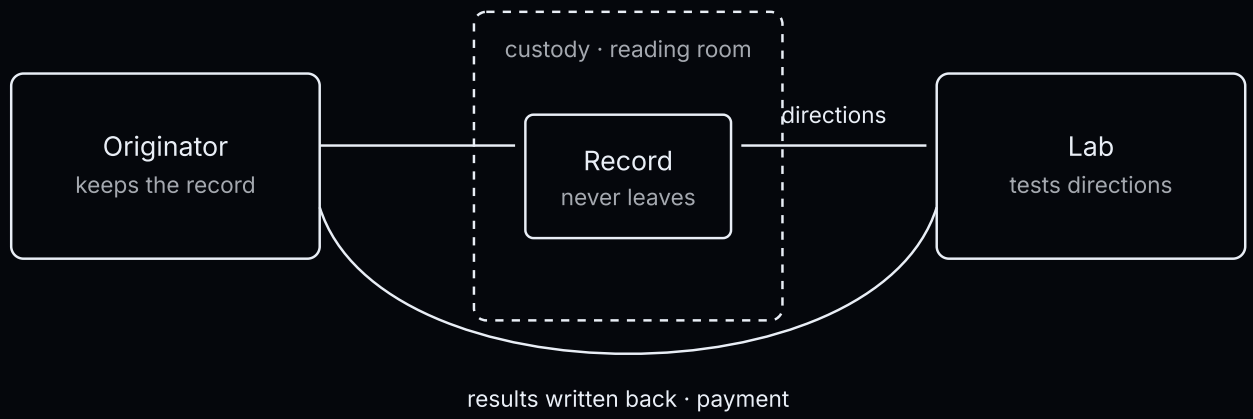

Figure 1. How a lab would read a record. The record stays in custody; only checked directions leave; results, including failures, come back as new dated entries, and payment follows.

03

Five claims

IN PLAIN TERMS

- Testing is getting cheap; choosing what to test is not.
- Machines asked for ideas tend to repeat themselves.
- A record of what someone actually thought beats a model that imitates them.
- Much of today's thinking happens with machines, and nobody keeps it for the person.
- Students and academics already need to show which words are theirs.

1. Testing is getting cheap; choosing is not

Most of the money in research goes on directions that fail. Of drug programmes that enter clinical trials, about 7.9 per cent reach approval, after an average of 10.5 years.[5] The median cost of bringing one new medicine to market, counting the failures, has been estimated at just under a billion dollars.[6] Across the economy, ideas are getting harder to find: keeping computer chips improving at their historic pace now takes about eighteen times as many researchers as it did in the early 1970s.[7]

Each failure began as a chosen direction. If cheaper testing lets a lab run a thousand experiments, it still has to choose which thousand. Even a small improvement in the share of directions that hold saves a great deal: if one tested direction in ten holds, each success costs ten tests; if a better source raised that to three in twenty, each success would cost about six and two-thirds. Nobody has shown that any source of directions does this. That is exactly what we will measure.

2. Machines asked for ideas repeat themselves

In a large study at Stanford, more than a hundred researchers compared ideas written by experts with ideas generated by a model, without knowing which was which. They rated the model's ideas as more novel. But when the researchers asked the model for thousands of ideas on one topic, only a small fraction were genuinely different from one another.[1] In a follow-up, the ideas were actually carried out, and the model's ideas lost more of their rating than the experts' did once tested.[2] Sounding new and holding up are different things. A long list of untested ideas is a poor starting point for a lab.

3. Keep the record, not a twin

A digital twin is a model trained to imitate one person, and imitation already works surprisingly well for answers. Researchers who built software agents from two-hour interviews found the agents could reproduce people's survey answers more than four-fifths as well as the people reproduced their own answers two weeks later.[8] But a twin gives the answers a person would probably give, not the ideas they actually had, or when. A twin can always be rebuilt from a record; a record can never be rebuilt from a twin. So the record is what lasts.

4. Thinking now happens with machines, and nobody keeps it for the person

By mid-2025 ChatGPT alone had around 700 million weekly users sending about 18 billion messages a week, and a large share of that use was people asking for help to decide or understand something.[9] Public places where people used to think out loud are emptying: within six months of ChatGPT's release, activity on Stack Overflow fell by about a quarter compared with similar forums.[10] People did not stop thinking. They moved their thinking into private chat windows. The record of what they asked, what they kept, what they rejected and which answer they followed now sits with the company that runs the tool, if it is kept at all.

5. Students and academics need to show which words are theirs

In UK universities, the use of AI tools by students became almost universal within two years: a 2025 survey found that 92 per cent of students used them, and 88 per cent had used them for assessed work.[11] At the same time the institutions that judge academic work have drawn firm lines. Nature will not accept a language model as an author, because authorship carries accountability for the work.[12] The US National Institutes of Health has said it will not treat grant applications substantially developed by AI as the original ideas of the applicants.[13] Students and researchers are being asked, more and more, to show which ideas were theirs.

A dated record that marks each passage as the person's, a machine's or someone else's answers that question directly. It is useful long before any lab reads it, which is why the free app comes first.

04

What Kynetic is, and the plan, published openly

IN PLAIN TERMS

- Phase 1 is a free app: your record, on your computer, dated, with your words kept apart.
- Phase 2 is a separate, opt-in programme: labs read in custody, people are paid.
- Installing the app enrols nobody in anything.

Phase 1: the app

The Kynetic app brings in a person's AI chats (from ChatGPT, Claude and others), coding sessions, notes, readings and files, and keeps them as one dated record on their own computer. There is no Kynetic account and no Kynetic server holding a copy. Every entry is dated. The person's original words are never changed; summaries, tags and links that a machine produces are stored beside them and marked as machine-made. Agents can tag and file the record, using no AI at all, the person's own API key, an assistant they already use, or a model on their own computer. Every route that could send anything off the computer starts switched off.

The record grows into a neuron the person can turn and explore: what came in along the branches, their own words at the centre, time running along the axon, and what they made at the far end. A student can trace an essay back through their own words to the chats and readings that fed it, and see what came straight from a machine.

Phase 2: the programme

We publish the plan now, with this paper's date, so that we can be held to what we wrote first.

1. **Single.** A lab licenses a seat on one person's record, for one stated purpose and period, inside custody.
2. **Pool.** A lab licenses a seat on many consenting records read together under code numbers. Only material dated before the pool is announced counts, so nobody can write to order. When a direction found across them holds, the records it came from share the payments.
3. **Open call.** A sponsor posts a problem it can test. Anyone can answer with a dated direction, at any time. If several people say the same thing, the earliest date takes the credit.
4. **Trial read.** Before committing to a seat, a lab can buy a short reading whose directions are scored blind by independent graders.
5. **Credit that outlives the person.** Payments can be directed to anyone or any cause, during the person's life or by will. Heirs inherit the right to payments but can never sell or open up the record.

A separate research arm may also study consenting records, with ethics approval, methods published first and failures published too, and never by training a model.

Why pools and open calls

Good directions often lie between records. In 1986 the information scientist Don Swanson noticed that two bodies of medical research that never cited each other fitted together: one showed that dietary fish oil changes the blood in certain ways, the other suggested those changes might help people with Raynaud's syndrome, in which blood flow to the fingers shuts down in the cold. He ran no experiment; his direction took one sentence.[14] A controlled trial in 1989 reported a benefit in the primary form of the condition.[15] None of the authors whose work he joined was paid for it. A pool does across many people's records what he did across two literatures.

Open calls reach people insiders miss. In a study of 166 scientific problems posted to an online network of outside solvers, the further a problem lay from a solver's own field, the more likely that solver was to crack it.[16] Open calls will also draw machine-written answers, which is why answers are dated, coded and judged without names: each is judged on what it says, and its date settles the credit.

05

The economics of a kept direction

IN PLAIN TERMS

- Money is pouring into machine-run science, and buyers already pay for human thinking, by the copy or by the hour.
- Thinking is already paid for: a brain runs on about 20 watts.
- Nobody can back-date an idea, so a record kept from today gains value with time.

Demand

Investors are funding machine-run science at scale. Isomorphic Labs raised 600 million dollars in 2025 to apply AI to drug discovery,[17] and Periodic Labs launched the same year with a reported 300 million dollars to build autonomous labs for materials science.[18] Every one of those labs will need directions worth testing.

Buyers already pay for human thinking, in two ways. They buy copies: Reddit disclosed 203 million dollars of data licensing contracts when it went public in 2024.[19] And they buy expert hours: Mercor, which hires experts to help train models, was valued at 10 billion dollars in 2025.[20] Taking without asking has proved expensive. In 2025 a US court found that training on lawfully bought books could be fair use, but that building a library from pirated copies was not; the case, Bartz v. Anthropic, then settled for about 1.5 billion dollars, around 3,000 dollars a book.[21] Some publishers now ask first and share the proceeds: Cambridge University Press asked its authors to opt in before licensing their books for AI and pays them a royalty.[22]

But both kinds of purchase buy the wrong thing for a lab. A copy, once delivered, can be used forever without further payment. An expert's hour buys an answer written after the question was asked, so nothing proves the idea came first. Neither pays the person when a direction holds.

Cost: thinking is already paid for

A human brain runs on about 20 watts.[23] Over an 80-year life that comes to about 14 megawatt-hours. Training GPT-3 in 2020 was estimated to use about 1,287 megawatt-hours, roughly the lifetime energy of 92 brains.[24] The comparison is rough, but the point holds: the cost of producing new human ideas has already been paid, by people thinking, whether or not anyone keeps what they produce. Reading a record costs very little, and less every year. It is cheaper to draw on what people have already thought than to pay machines to think it again.

Scarcity: human writing is running out, and dates cannot be faked

Models are on course to have used up the public stock of human-written text: one careful estimate puts the median year at 2028, with a range of 2026 to 2032.[25] Training models on text written by earlier models narrows them, losing rare cases first.[26] New ideas will still have to come from people.

What no amount of computing can buy is a date. On 28 September 1838 Charles Darwin began reading Malthus on population and saw how the same struggle could shape species. His notebooks kept the date. Twenty years later, when Alfred Russel Wallace sent him an essay with the same idea, both men's work was read together at the Linnean Society, and Darwin's dated notes and earlier letters meant the two could be presented side by side, and both men are credited.[27] Since 1735 the French Academy of Sciences has kept sealed, dated envelopes in which anyone may record an idea without publishing it. One, sent in 1940 by the young mathematician Wolfgang Doeblin shortly before his death in the war, was opened in 2000 and showed he had found a result later published independently by Kiyosi Itô; it is now sometimes called the Itô–Doeblin formula.[28]

A date matters most to people who cannot carry their own idea further. Alexander Fleming reported in 1929 that a mould killed bacteria; it was Howard Florey and Ernst Chain's team who showed in 1940 that penicillin could be a medicine, and the three shared the Nobel Prize.[29] Fleming's dated report kept his place.

A rival that starts keeping records later can never offer earlier dates. That is the moat.

Prices, and who is paid

We deliberately set no prices in this paper. Seat fees will be agreed with the first buyers by bid, measured against what they already pay for licensed data and expert hours, and never below the cost of custody. Two hazards are known. People may start writing what they think buyers want, so answers written for an open call are marked as such, apart from material written before. And buyers who cannot tell records apart fall back on famous names, the advantage Robert Merton called the Matthew effect.[30] So no record is priced until records have been tested blind against cheaper sources of ideas.

Under a Kynetic licence the person who is read is the one paid: a share of every fee a lab pays to read their record, and milestone payments and a royalty when a direction holds. How the fee is split between the person and Kynetic will be fixed in the charter, with legal advice, and published before the first licence.

Why most of the money waits for results

Selling an idea runs into an old problem, which the economist Kenneth Arrow set out in 1962: a buyer cannot judge information until they have seen it, and once they have seen it, they no longer need to pay for it.[31] A direction fits in a sentence, so showing it gives it away. Custody answers this in five ways:

1. The lab's own software judges a record where it is kept, cut off and capped in what it may quote, so the lab learns its worth without taking it away.
2. The licence, its purpose and what will count as holding are agreed before anything is read.
3. A lab can buy a trial read, scored blind, before a seat.
4. Milestones and royalties fall due only when an outside document shows the agreed outcome: a publication, a patent application, a trial registration or a regulator's decision.
5. Every entry is dated when it is written, so a lab cannot later claim it already had the idea.

06

How a record is kept and read

IN PLAIN TERMS

- The record never leaves custody. A lab's own software reads it there, and only checked directions leave.
- Every entry is dated, and no single key-holder, Kynetic included, can open a record alone.

A record would pass through five stages, designed so that no single person or organisation can read it alone.

Screened. Before anything reads a record, machines check all of it. Other people's words, health details, research participants' data, anything that may fall under export controls, and anything the screen is unsure of are sealed and stay on the originator's own device. A second, separate screen reads every direction before it leaves, and if it flags anything, the whole direction is held back. No person reads a record to screen it.

Sealed and dated. Each entry carries its date and who added it, and entries are only ever added at the end. A deletion leaves a dated gap, never a silent one. Each day's new entries are sealed into one encrypted block. Only a fingerprint of the block is published: for a block B and a secret random salt s, the fingerprint is $f = \mathrm{H}(s \,\|\, B)$, where H is a standard hash function. Publishing f commits to the block like a sealed envelope; anyone later shown s and B can check they match f, but nobody holding only f can learn anything about B. The date comes from OpenTimestamps, a free public service that Kynetic does not control.[32]

Split. Each record's key is divided between the originator, Kynetic and an independent keyholder, so that no one of them can open it alone. Every opening is logged.

Read in place. Before anyone reads, the originator names the subjects their record may serve, the buyers it may never serve and when licensing may begin. Every question is checked against them. The lab's own agent reads inside custody, in a sandbox with no connection out and only the tools its licence allows. Access is headless, by command line, an MCP server or a local programming interface, so no person at the lab browses the record.

Gated. Only directions in a fixed form leave, after the second screen, and the amount that can ever leave as quotation is capped. Every release is logged.

From a direction to a lab's run

A direction is still a sentence; a self-driving lab needs a test it can run. Each chosen direction is written up as an experiment. The person it came from confirms that the test still says what they meant, because a machine may translate an idea but never picks it. The pass mark is fixed before anything runs. Nothing runs until a named scientist at the lab approves it, and Kynetic never touches a lab's equipment. When the run ends, the result, pass or fail, is written back to the person's record as a new dated entry.

What exists today

We are honest about status. The app's core keeps original words untouched, dates every version and de-duplicates imports; its design keeps every network route off by default. **Nothing of the programme is built.** The custody design, the split keys, the reading room and the public fingerprints are designs. Before any real record enters custody, an independent cryptographer will review the design. The charter and draft agreements for originators, labs and sponsors are written and awaiting legal review.

Other sources of directions

Labs have other sources of directions, and most are cheaper, so a record has to earn its cost.

SOURCE	DATED BEFORE THE QUESTION?	KEEPS DEAD ENDS?	PERSON KEEPS THE RECORD?	PERSON PAID WHEN IT HOLDS?	PROVEN AT SCALE?
Published papers	Yes	Rarely	Only the finished slice	No, credit by citation	Yes
Hiring an expert	No	Partly, from memory	Yes	No, paid by the hour	Yes
A prize for a stated problem	No	No	Depends on terms	Yes, the prize	Yes
Asking a model	No	No	No person involved	No	Partly: plenty, but repetitive
Training on licensed text	Lost in training	No	No	No, paid once	Yes
Chat logs held by a company	Within one company	One company's slice	No	No	Yes
A Kynetic record	Yes, every entry	Yes	Yes, it never leaves custody	Yes	No, nothing is held yet

07

Why not the model makers, or the employer?

IN PLAIN TERMS

- A company that sells models cannot be a neutral keeper of the records that could improve them.
- Chat logs belong to providers or employers, and a court can order them kept.
- Kynetic keeps records the way a seed vault keeps seeds: the depositor decides.

A model maker could build the reading room. It cannot be a neutral keeper, for two reasons. It sells the models the records could improve, so it has a stake in what they contain. And it sees only the work done in its own products, not across rival tools. A provider's promise to delete chats is a policy, and a court can override it: in 2025 OpenAI was ordered to preserve ChatGPT conversations, including deleted ones, in a copyright case brought by The New York Times.[33] Kynetic holds only one of three key shares, so it cannot open a record on its own, even if ordered to.

The promise a model maker cannot make is the price of entry: no outside company ever trains a model on a record, and if Kynetic ever trains a model of its own, it does so only in custody, with separate consent, and pays for what went in. The moat is what builds up behind that promise: dates that cannot be made later; a pool that grows more valuable with every record kept; and a history, held only by the keeper, of which records produced directions that held.

The world already keeps its seeds this way. Since 2008 the Svalbard Global Seed Vault has stored spare copies of seed collections from around the world. Each box stays the property of the bank that sent it, and only that bank can take it out. When war forced an agricultural research centre out of Aleppo, it became the first to withdraw seeds, in 2015, to rebuild its collection in Lebanon and Morocco.[34] The keeper stored the seeds; the owner decided what to do with them.

08

Rules are arriving

IN PLAIN TERMS

- Rules on marking machine-written text and on training data are coming into force.
- Funders and journals already insist that original ideas come from accountable people.
- Wider rules make a dated, attributed direction more valuable, not less.

Under the European Union's AI Act, providers of general-purpose models must publish a summary of the content they were trained on, and machine-generated content must be marked in a way software can detect.[35] Funders and journals already treat original ideas as human: the NIH will not treat applications substantially developed by AI as the applicants' original ideas, and Nature will not list a language model as an author.[13,12] As self-driving labs spread, the questions will widen from what models learn from to who chose what a lab tests, and who answers for it.

A Kynetic direction carries its date, the person who agreed to its use and a record of how it was chosen. It does not make a test safe; the lab still answers for its experiments. What it shows is where the idea came from. The more capable machines become, the more that will be asked for.

09

Risks, and how we limit them

RISK	WHAT COULD HAPPEN	HOW WE LIMIT IT
The value is unproven	Records may not beat cheaper sources of ideas.	A blind test is designed and published before any price, with the margin fixed in advance. The result is published either way. The free app stands on its own.
Trust	One mistake with a person's record could end the programme.	Nothing is opt-out; each study and licence asks separately; the charter only tightens; an independent auditor, never paid from licences, checks every access.
Law and privacy	Records hold other people's words and sensitive details.	Two screens; sealed material never enters a licence; adults only; legal review, a data protection impact assessment and ethics review before the first reading.
Security	Custody could be breached.	Split keys, an attested machine, logging, and independent cryptographic review before any real record is held.
Adoption	People may not keep records, or labs may not buy.	The app is useful on its own, for free; students and academics first, where the authorship need is immediate.
Power	Employers or universities could pressure people to license.	Nobody with power over a person may license their record, during the relationship or for two years after.
Cost	Running a service could outpace income for years.	Local-first design: no servers hold records, so the free app costs little to run.

Until legal review, ethics review, an independent review of custody, a published blind test and an auditor are all in place, the programme does not open, and nothing is sold.

10

Closing entry

IN PLAIN TERMS

- A record can prove when you had an idea only if it was running at the time.
- Start a free record now. Labs and sponsors can ask about the first seats and open calls.

In the spring of 1976 the mathematician George Andrews, looking through a dead colleague's papers in the library of Trinity College, Cambridge, found 138 pages in Srinivasa Ramanujan's handwriting that nobody had noticed for half a century. They are still being studied.[36] Today's thinking will survive only by that kind of luck unless someone keeps it, dated and in the thinker's name.

Self-driving labs will not run short of speed. They will run short of directions worth testing, traced to the people who had them. Kynetic is built to keep that record, prove where each idea came from, and carry it, with consent, to those who can test it.

How to take part.

- **If you study or think for a living,** start your record. The app is free, and your record stays on your computer: join the waitlist at kynetic.cc.
- **If you run a lab, fund research or build AI for science,** ask how a seat, a pool or an open call would work, and help us design the blind test.
- **If you want to help build it,** as an investor, a partner, an adviser on ethics or security, or a colleague, write to hello@kynetic.cc.

Thirty years from now, will what you thought this year still exist, dated and in your name?

Sources

1. Chenglei Si, Diyi Yang and Tatsunori Hashimoto, "Can LLMs Generate Novel Research Ideas? A Large-Scale Human Study with 100+ NLP Researchers", arXiv 2409.04109, September 2024. https://arxiv.org/abs/2409.04109
2. Chenglei Si, Tatsunori Hashimoto and Diyi Yang, "The Ideation-Execution Gap: Execution Outcomes of LLM-Generated versus Human Research Ideas", arXiv 2506.20803, June 2025. https://arxiv.org/abs/2506.20803
3. Nathan Szymanski and colleagues, "An autonomous laboratory for the accelerated synthesis of novel materials", Nature 624: 86–91, 2023. https://doi.org/10.1038/s41586-023-06734-w
4. Josh Leeman, Leslie Schoop, Robert Palgrave and colleagues, "Challenges in High-Throughput Inorganic Materials Prediction and Autonomous Synthesis", PRX Energy 3: 011002, 2024. https://doi.org/10.1103/PRXEnergy.3.011002
5. Biotechnology Innovation Organization, Informa Pharma Intelligence and QLS Advisors, "Clinical Development Success Rates 2011–2020", February 2021. https://go.bio.org/rs/490-EHZ-999/images/ClinicalDevelopmentSuccessRates2011_2020.pdf
6. Olivier Wouters, Martin McKee and Jeroen Luyten, "Estimated Research and Development Investment Needed to Bring a New Medicine to Market, 2009–2018", JAMA 323(9): 844–853, 2020. https://doi.org/10.1001/jama.2020.1166
7. Nicholas Bloom, Charles Jones, John Van Reenen and Michael Webb, "Are Ideas Getting Harder to Find?", American Economic Review 110(4): 1104–1144, 2020. https://doi.org/10.1257/aer.20180338
8. Joon Sung Park and colleagues, "Generative Agent Simulations of 1,000 People", arXiv 2411.10109, version 1, November 2024 (later versions retitled "AI Simulations Grounded in Rich Self-Reports Increase Accuracy and Reduce Distortion"). https://arxiv.org/abs/2411.10109
9. Aaron Chatterji and colleagues, "How People Use ChatGPT", National Bureau of Economic Research working paper 34255, September 2025. https://www.nber.org/papers/w34255
10. Maria del Rio-Chanona, Nadzeya Laurentsyeva and Johannes Wachs, "Large language models reduce public knowledge sharing on online Q&A platforms", PNAS Nexus 3(9): pgae400, 2024. https://doi.org/10.1093/pnasnexus/pgae400
11. Josh Freeman, "Student Generative AI Survey 2025", Higher Education Policy Institute and Kortext, February 2025. https://www.hepi.ac.uk/2025/02/26/student-generative-ai-survey-2025/
12. "Tools such as ChatGPT threaten transparent science; here are our ground rules for their use", Nature 613: 612, 24 January 2023. https://doi.org/10.1038/d41586-023-00191-1
13. National Institutes of Health, "Supporting Fairness and Originality in NIH Research Applications", notice NOT-OD-25-132, 17 July 2025, applying to applications from 25 September 2025. https://grants.nih.gov/grants/guide/notice-files/NOT-OD-25-132.html
14. Don R. Swanson, "Fish oil, Raynaud's syndrome, and undiscovered public knowledge", Perspectives in Biology and Medicine 30(1): 7–18, 1986. https://doi.org/10.1353/pbm.1986.0087
15. Ralph DiGiacomo, Joel Kremer and Dhiren Shah, "Fish-oil dietary supplementation in patients with Raynaud's phenomenon: a double-blind, controlled, prospective study", The American Journal of Medicine 86(2): 158–164, 1989. https://pubmed.ncbi.nlm.nih.gov/2536517/
16. Lars Bo Jeppesen and Karim Lakhani, "Marginality and Problem-Solving Effectiveness in Broadcast Search", Organization Science 21(5): 1016–1033, 2010. https://doi.org/10.1287/orsc.1090.0491
17. Isomorphic Labs, "Isomorphic Labs announces $600M external investment round", 31 March 2025. https://www.isomorphiclabs.com/articles/isomorphic-labs-announces-600m-external-investment-round
18. SiliconANGLE, "Periodic Labs raises $300M to accelerate scientific research with AI", 1 October 2025. https://siliconangle.com/2025/10/01/periodic-labs-raises-300m-accelerate-scientific-research-ai/
19. Reddit, Inc., Form S-1 registration statement, 22 February 2024: "an aggregate contract value of $203.0 million". https://www.sec.gov/Archives/edgar/data/1713445/000162828024006294/reddits-1q423.htm
20. TechCrunch, "Mercor quintuples valuation to $10B with $350M Series C", 27 October 2025. https://techcrunch.com/2025/10/27/mercor-quintuples-valuation-to-10b-with-350m-series-c/
21. Bartz v. Anthropic PBC, No. C 24-05417 WHA (N.D. Cal.), order on fair use, 23 June 2025, https://storage.courtlistener.com/recap/gov.uscourts.cand.434709/gov.uscourts.cand.434709.231.0_2.pdf; settlement of about $1.5 billion (about $3,000 a work) announced September 2025 and given final approval on 20 July 2026, https://www.anthropiccopyrightsettlement.com/
22. Authors Alliance, "AI Licensing: An Interview with Ben Denne of Cambridge University Press", 17 March 2025. https://www.authorsalliance.org/2025/03/17/ai-licensing-an-interview-with-ben-denne-of-cambridge-university-press/
23. Vijay Balasubramanian, "Brain power", Proceedings of the National Academy of Sciences 118(32): e2107022118, 2021. https://doi.org/10.1073/pnas.2107022118
24. David Patterson and colleagues, "Carbon Emissions and Large Neural Network Training", arXiv 2104.10350, April 2021, table 4. https://arxiv.org/abs/2104.10350 (The 80-year life and the ratio of about 92 are our arithmetic: 20 W × 8,760 h × 80 ≈ 14.0 MWh; 1,287 ÷ 14.0 ≈ 92.)

25. Pablo Villalobos and colleagues, "Position: Will we run out of data? Limits of LLM scaling based on human-generated data", Proceedings of the 41st International Conference on Machine Learning, PMLR 235, 2024. https://proceedings.mlr.press/v235/villalobos24a.html

26. Ilia Shumailov and colleagues, "AI models collapse when trained on recursively generated data", Nature 631: 755–759, 2024. https://doi.org/10.1038/s41586-024-07566-y

27. Darwin Correspondence Project, University of Cambridge, on the joint Darwin–Wallace reading at the Linnean Society on 1 July 1858, https://www.darwinproject.ac.uk/; and Darwin Online, notebook pages of 1838 (Cambridge University Library, DAR 71.53–59), on his reading of Malthus from 28 September 1838, https://darwin-online.org.uk/converted/manuscripts/Darwin_C_R_CUL-DAR71.53-59.html

28. Bernard Bru and Marc Yor, "Comments on the life and mathematical legacy of Wolfgang Doeblin", Finance and Stochastics 6(1): 3–47, 2002. https://doi.org/10.1007/s780-002-8399-0

29. The Nobel Prize in Physiology or Medicine 1945, Sir Alexander Fleming, Ernst Boris Chain and Sir Howard Walter Florey. https://www.nobelprize.org/prizes/medicine/1945/summary/

30. Robert K. Merton, "The Matthew effect in science", Science 159(3810): 56–63, 1968. https://doi.org/10.1126/science.159.3810.56

31. Kenneth J. Arrow, "Economic Welfare and the Allocation of Resources for Invention", in The Rate and Direction of Inventive Activity, Princeton University Press for the National Bureau of Economic Research, 1962, pages 609–626. https://www.nber.org/books-and-chapters/rate-and-direction-inventive-activity-economic-and-social-factors/economic-welfare-and-allocation-resources-invention

32. OpenTimestamps, "A timestamping proof standard". https://opentimestamps.org/

33. OpenAI, "How we're responding to The New York Times' data demands in order to protect user privacy", 5 June 2025. https://openai.com/index/response-to-nyt-data-demands/

34. Svalbard Global Seed Vault, "ICARDA seeds safe back in Lebanon and Morocco", 27 October 2017. https://www.seedvault.no/news/icarda-seeds-safe-back-in-lebanon-and-morocco/

35. Regulation (EU) 2024/1689 (the Artificial Intelligence Act), Articles 50 and 53. https://artificialintelligenceact.eu/article/50/ and https://artificialintelligenceact.eu/article/53/

36. George E. Andrews and Bruce C. Berndt, Ramanujan's Lost Notebook, Part I, Springer, 2005. https://link.springer.com/book/10.1007/0-387-28124-X