Will the Next Decade Out-Learn the Last 2,600 Years?

A simple test of the exponential growth of recorded knowledge

23 September 2026

Abstract

When a record grows exponentially, each doubling adds as much as everything that came before. So if recorded knowledge now doubles in less than ten years, the decade 2026–2035 will add more than all earlier history. We test this with public counts of scholarly works, preprints, patents and scientific databases, and with counts of manuscripts and printed books reaching back toward the Babylonian Map of the World of about 600 BCE. Doubling times have shrunk from about 99 years for medieval manuscripts to 14 to 19 years for journals in the late twentieth century and 9.1 years for today's scholarly literature. At its current growth the scholarly record registered with Crossref will add about 155 million works in 2026–2035, 93% of the 167 million recorded before (likely range 84% to 108%). The next decade will therefore add almost as much as all earlier history, and slightly faster growth would take it past. The fastest records, arXiv's AI papers, papers on deep learning and several preprint servers, double within four years and will add 4 to 52 times their entire past. These are projections of document counts, which track the volume of recorded knowledge and say little about its value.

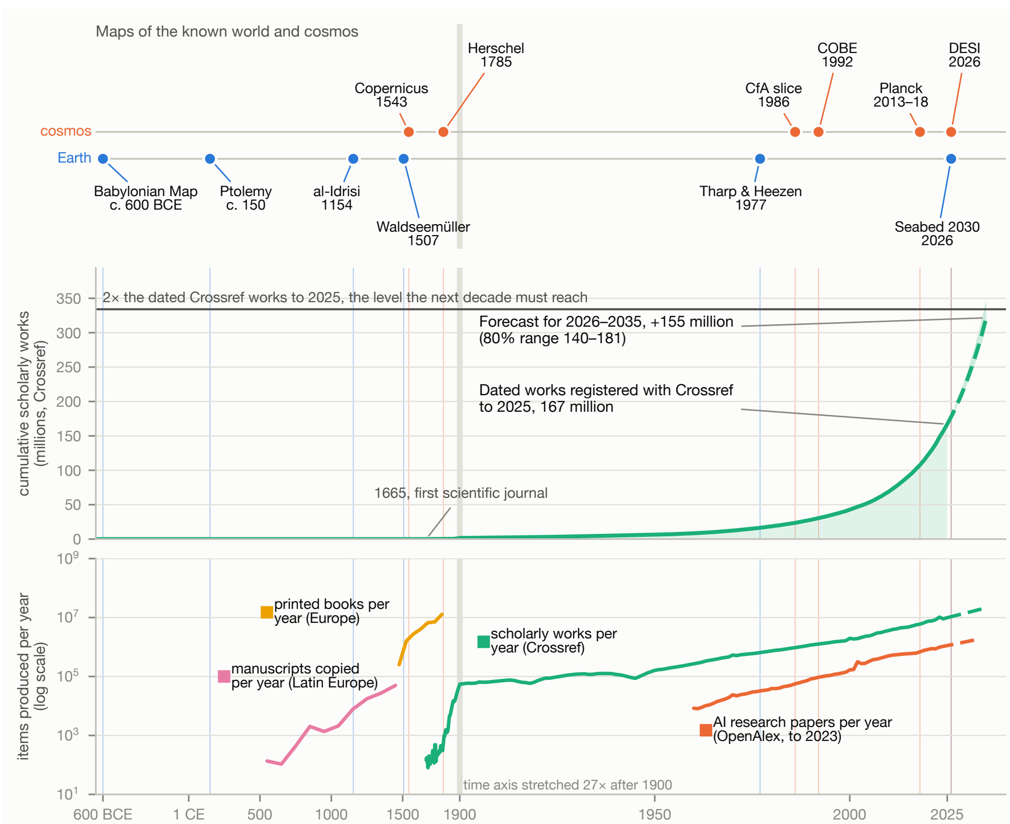

Figure 1. Recorded knowledge from 600 BCE to 2035. Top, milestones in mapping the Earth and the cosmos. Middle, the total number of scholarly works registered with Crossref, projected to 2035 at its 2016–2025 growth with the likely range shaded, against twice its 2025 total. Bottom, yearly output on a log scale. The time axis is compressed before 1900.

1 The question

The Babylonian Map of the World, a clay tablet of about 600 BCE, is the oldest surviving map of the then known world [1]. It shows Babylon, a ring of ocean and regions beyond it that the text fills with myth, and reaches about 700 km from its centre. Later maps spanned the globe by 1507 and reached the most distant light in the universe in 1992 (Figure 2). Parts of today's record grow very fast. Submissions to the International Conference on Machine Learning doubled from 12,107 in 2025 to 24,371 in 2026 [2, 3], and arXiv, the main preprint server for physics, mathematics and computer science, passed three million papers in April 2026, four years after passing two million [4, 5]. The question is whether the ten years from 2026 to 2035 will add more to the record than all the centuries before them.

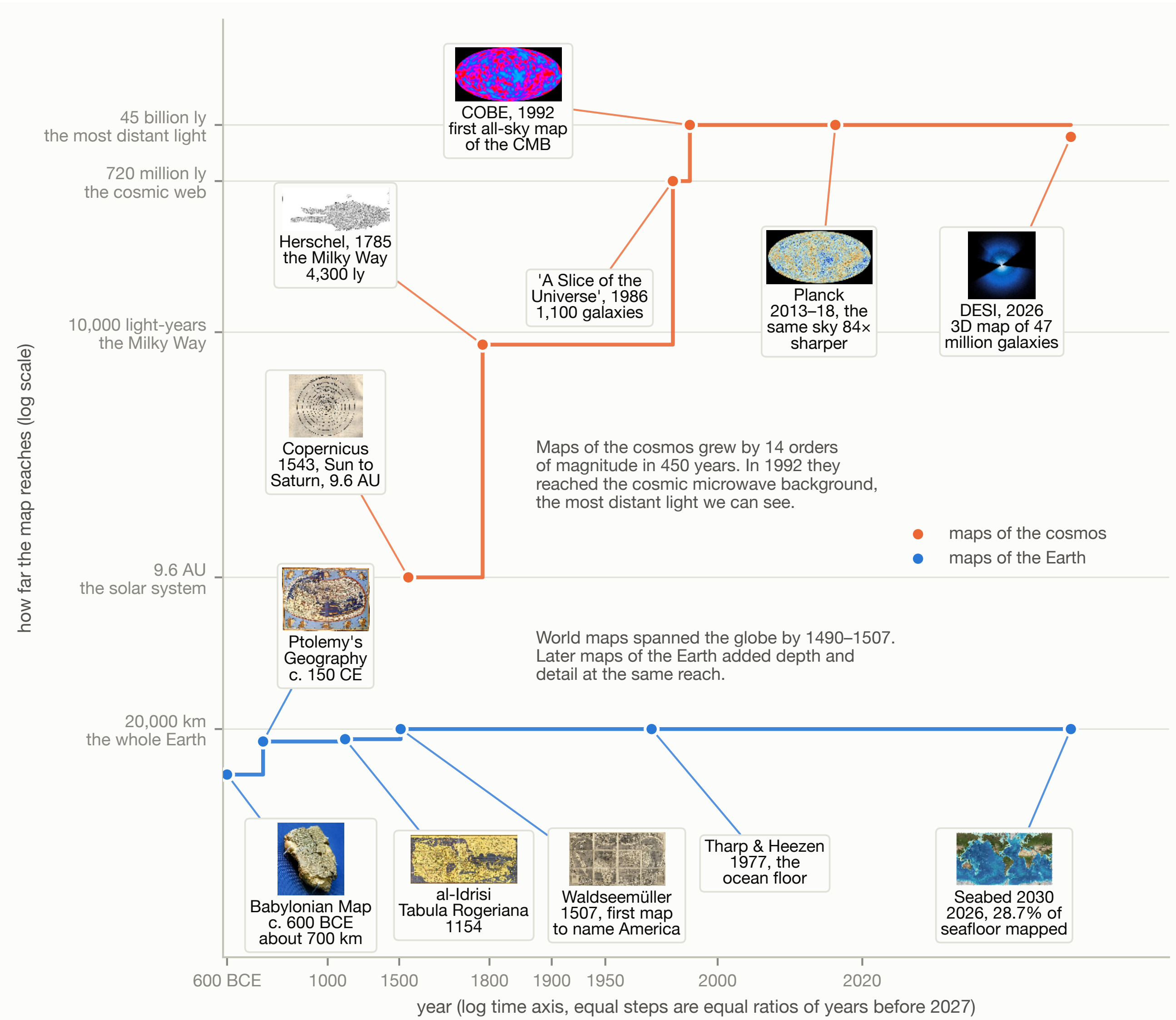

Figure 2. How far famous maps reached, measured as the distance from the map's centre to the farthest place it shows. Cosmic distances allow for the expansion of the universe. Sources and ranges are in `data/historical_maps.csv` [1, 6, 7]. The thumbnails come from Wikimedia Commons and are scaled. The Ptolemy image is the Ulm edition of 1482, the al-Idrisi image Konrad Miller's reconstruction of 1929 and the Seabed 2030 image the GEBCO 2014 grid (CC0). The licensed images are the tablet photograph by O. S. M. Amin (CC BY-SA 4.0), the Planck map by ESA and the Planck Collaboration (CC BY 4.0) and the DESI map by the DESI Collaboration and DESI Member Institutions/DOE/KPNO/NOIRLab/NSF/AURA/R. Proctor with image processing by M. Zamani (NSF NOIRLab) (CC BY 4.0). The others are in the public domain.

Knowledge has no unit, so we count what is recorded, papers, preprints, patents and entries in scientific databases, in

27 public records (Table 3). Figure 1 shows the idea for the scholarly literature. Its total rose slowly for three centuries and then steeply, and if recent growth continues the next decade nearly doubles it.

2 Why the doubling time decides it

The scientific literature has been known to grow exponentially since Price's studies of the 1960s [8, 9]. Publication counts grew about 4% a year from the late seventeenth century to 2018 and 5% a year after 1952 [10]. Exponential growth has a fixed doubling time, and each doubling adds as much as everything before it. The same arithmetic gives the ratio for any decade. If a record doubles every d years, the next ten years add $2^{10/d} - 1$ times everything recorded before (Table 1). A doubling time of exactly ten years gives a ratio of 1, so the next decade out-produces all earlier history whenever the record doubles in less than ten years, which takes growth above 7.2% a year.

Table 1. What the next ten years add, relative to everything recorded before, for a record that has always grown at a steady rate.

Doubling time (years)	Next ten years ÷ everything before
50	0.15
20	0.41
15	0.59
10	1.00
8	1.38
5	3.00
2	31

This rule holds exactly for a record that has always grown at the same rate. A record that grew more slowly in the past holds more relative to its current output and needs faster growth now. We therefore measure each record's recent growth, project its output for ten years and compare the result with everything it recorded before.

3 Growth through history

Figure 3 shows every doubling time measured in the study. Manuscript copying in Latin Europe doubled about every 99 years before 1450 and printed books in Europe every 62 years from 1475 to 1775 [11]. English patents doubled every 19 years from 1700 to 1850 and US patents every 10 years in the nineteenth century. From 1950 to 2000 journal literatures doubled every 14 to 19 years, and since 2006 arXiv has doubled every 7.9 years. The decline is uneven, since US patents took 57 years to double in the twentieth century. The fastest records today belong to AI. arXiv's AI papers doubled every 2.0 years over 2016–2025, the computing power used to train the largest AI models every 6 months over 2012–2025 [12, 13] and the length of tasks AI agents can complete every 4 months over 2023–2026 [14, 15].

4 The next decade

Figure 4 projects four records to 2035 at their growth rates of 2016–2025. A line that ends above 2 means the next decade adds more than the record's entire past. Table 2 gives the numbers for the main records and Table 3 for all 27.

The scholarly literature registered with Crossref, 167 million works up to 2025, grew 8.0% a year over 2016–2025, so its yearly output doubles about every 9.1 years. Projected at that rate, 2026–2035 adds about 155 million works, 93% of everything before. The likely range, based on how far such ten-year projections missed in the past, runs from 84% to 108%. Crossref falls just short although its output doubles in under ten years, because it grew more slowly in the past and holds a large past relative to its present output. Growth of 9.3% a year instead of 8.0% would carry the next decade past everything before.

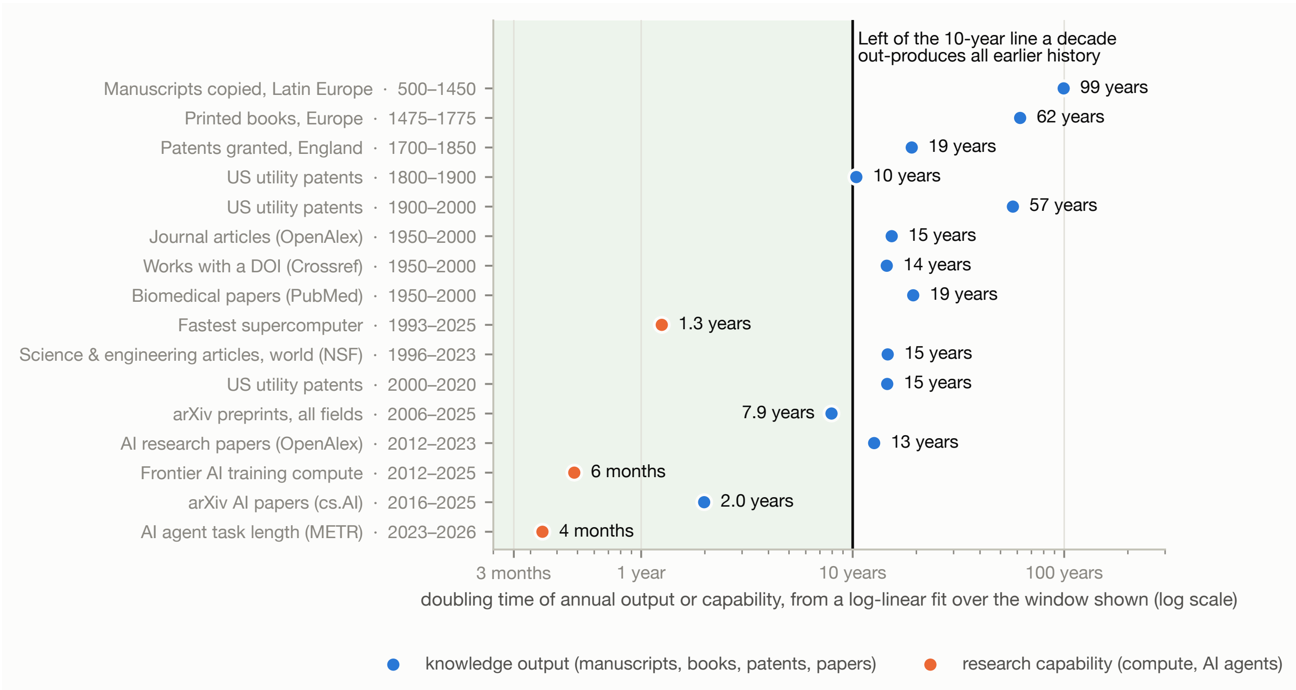

Figure 3. Doubling time of yearly output for every period and record measured in the study, oldest at the top. Left of the ten-year line, a record growing at a steady rate out-produces its entire past every decade. Data in `results/hero3_data.csv`.

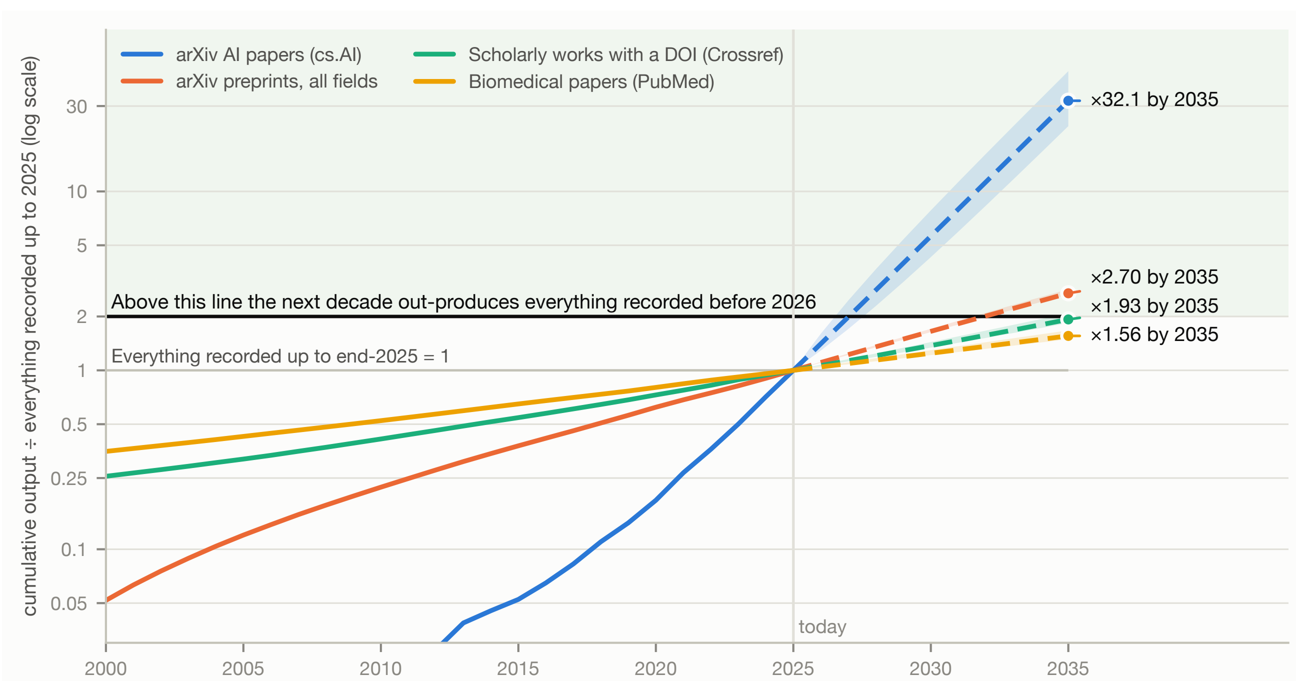

Figure 4. Four records divided by their total at the end of 2025 and projected to 2035 at their growth rates of 2016–2025, with likely ranges shaded. A record that ends above 2 adds more in 2026–2035 than in its entire past. arXiv's AI papers include papers cross-listed from other categories. Data in `results/hero2_data.csv`.

Table 2. The next decade for the main records. Growth is the yearly growth of output over the last ten recorded years, and the doubling time is that of yearly output. The ratio divides the output of the next ten years, projected at that growth, by everything recorded before. The likely range is an 80% range based on how far such projections missed in the past. Records that end before 2025 are projected for the ten years after their last year.

Record	Counts	Total so far	Growth a year	Doubling time (years)	Next decade ÷ all before	Likely range
Crossref	scholarly works with a DOI	167 million	8.0%	9.1	0.93	0.84–1.08
OpenAlex	indexed scholarly works	293 million	1.9%	36	0.56	0.47–0.71
PubMed	biomedical papers	40.1 million	4.9%	14.5	0.56	0.49–0.67
US patents (to 2020)	patents granted	10.8 million	4.4%	16.1	0.41	0.33–0.56
arXiv	preprints	2.9 million	9.9%	7.4	1.70	1.65–1.84
bioRxiv	biology preprints	312,221	23.1%	3.3	5.88	4.19–8.70
arXiv AI (cs.AI)	AI papers	156,573	41.8%	2.0	31.1	22.2–46.0
OpenAlex AI (to 2023)	papers tagged AI	14.9 million	6.4%	11.2	0.95	0.73–1.32
Protein Data Bank	protein structures	247,225	5.3%	13.3	0.96	0.07–1.64
GenBank	DNA bases	6.7 trillion bases	75.3%	1.2	150	10.2–256
NASA Exoplanet Archive	confirmed planets	6,089	−5.6%	falling	0.30	0.02–0.72
English Wikipedia	articles	7.1 million	−3.9%	falling	0.23	0.02–0.56

Records that grow more slowly fall well short. The biomedical literature in PubMed grows 4.9% a year and adds 56% of its past, and US patents add 41%. Faster records pass easily. arXiv adds 1.70 times its past and its AI category 31.1 times. Of the 27 records, 13 are projected to add more than their past. All but one are preprint, AI or DNA sequence records, and the exception is the count of near-Earth asteroids. Scientific databases vary most, from DNA sequence archives that add many times their past to the catalogue of confirmed exoplanets, which adds 30%.

5 What it means

Counted by recorded works, the answer is close. The scholarly literature as a whole is on course to add almost as much in 2026–2035 as in all earlier centuries, and a modest rise in growth would take it past. The fastest records, arXiv's AI papers and several preprint servers, will add many times their past, while PubMed, OpenAlex and US patents will add about half or less.

Counts measure the volume of recorded knowledge, and the value of each unit may be falling. At least 13.5% of biomedical abstracts in 2024 were processed with large language models [16], and scientists who use AI publish about 3 times as many papers while the range of topics they study narrows [17]. Counted by notable scientists and inventors instead of documents, no decade from 1500 to 1950 added more than 14% to all who came before [18].

The growth does not come from more people. The decade 2026–2035 adds only about 6% to all the years humans have lived since 600 BCE [19], so the rise in output comes almost entirely from more output per person.

The projections extend the last ten years of growth. Trends do change, which is why each estimate carries a likely range. The records also start when their registries start, so most knowledge written down before the scientific journals of the seventeenth century is not counted.

A modest acceleration therefore decides whether the next decade out-learns the last 2,600 years. The scholarly record needs its growth to rise from 8.0% to 9.3% a year. AI systems that carry out research, for example inside validated simulations, could supply it, and the computing power and AI capabilities they rely on double every few months.

6 Data and methods

All counts were retrieved on 23 September 2026 from public sources, scholarly works from Crossref [20] and OpenAlex [21], biomedical papers from PubMed [22], preprints from arXiv [23] and from bioRxiv and medRxiv [24], patents from the US Patent and Trademark Office via Our World in Data [25], and database entries from the Protein Data Bank [26], GenBank [27], the Sequence Read Archive [28], UniProt [29], the NASA Exoplanet Archive [30], NASA's Center for Near Earth Object Studies [31] and Wikimedia [32]. Manuscripts and printed books come from Buringh and van Zanden [11] via Our World in Data, population from HYDE [19], notable people from Pantheon [18], AI computing power from Epoch AI [12] and AI task lengths from METR [14].

Each record's count runs to its last complete year, 2025 for most. Growth is the slope of a straight line fitted to the logarithm of yearly output over the last ten years. The projection applies that growth to the last year's output for ten more years and divides the total by everything recorded before. The likely range is the range that held 80% of outcomes when the same projection, made from each year between 1990 and 2015, was compared with the decade that followed. OpenAlex changed how it tags research topics in 2024, so its three AI series stop in 2023, and two series end early, US patents in 2020 and the Sequence Read Archive in 2023. Each of these is projected for the ten years after its last year.

The raw downloads, code and results are in the study folder, and rerunning the analysis from the raw data reproduces all 23 compared results exactly. A technical version of this report, with proofs, the estimation of the likely ranges, robustness checks and independent reviews, is in `report_archive/v1.2.1`.

References

1. British Museum. Map of the world (the Babylonian map of the world), clay tablet, museum number 1882,0714.509. https://www.britishmuseum.org/collection/object/W_1882-0714-509
2. RIKEN Center for Advanced Intelligence Project. 29 papers have been accepted at ICML 2025. 2025. https://aip.riken.jp/news/icml2025/
3. ICML Conference. The people want to know: how many submissions made it over the finish line? ... 24371! This is the number of submitted PDFs. Post by the ICML account (icmlconf) on X, 2026. https://x.com/icmlconf/status/2016954655599735289
4. Boboris K. arXiv now hosts over 3 million articles. arXiv blog, 2026. https://blog.arxiv.org/2026/07/09/arxiv-now-hosts-over-3-million-articles/
5. Fromme A. arXiv hits 2M submissions. Cornell Chronicle, 2022. https://news.cornell.edu/stories/2022/01/arxiv-hits-2m-submissions
6. The Nippon Foundation-GEBCO Seabed 2030 Project. Global seabed mapping reaches new milestone as five million square kilometres added in a year. 2026. https://seabed2030.org/2026/04/20/global-seabed-mapping-reaches-new-milestone-as-five-million-square-kilometres-added-in-a-year/
7. Lawrence Berkeley National Laboratory. DESI completes planned 3D map of the universe and continues exploring. Berkeley Lab News Center, 2026. https://newscenter.lbl.gov/2026/04/15/desi-completes-planned-3d-map-of-the-universe-and-continues-exploring/
8. de Solla Price DJ. Science since Babylon. Yale University Press, 1961.
9. de Solla Price DJ. Little science, big science. Columbia University Press, 1963.
10. Bornmann L, Haunschild R, Mutz R. Growth rates of modern science: A latent piecewise growth curve approach to model publication numbers from established and new literature databases. *Humanities and Social Sciences Communications* 8, 224 (2021). doi.org/10.1057/s41599-021-00903-w
11. Buringh E, van Zanden JL. Charting the "rise of the West": Manuscripts and printed books in Europe, a long-term perspective from the sixth through eighteenth centuries. *The Journal of Economic History* 69, 409–445 (2009). doi.org/10.1017/S0022050709000837
12. Epoch AI. Data on AI models. 2026. https://epoch.ai/data/ai-models
13. Epoch AI. Trends in artificial intelligence. 2026. https://epoch.ai/trends
14. METR. Time horizon 1.1. 2026. https://metr.org/blog/2026-1-29-time-horizon-1-1/
15. Kwa T, West B, Becker J, et al. Measuring AI ability to complete long software tasks. arXiv 2503.14499, NeurIPS 2025, 2025. https://arxiv.org/abs/2503.14499
16. Kobak D, González-Márquez R, Horvát E-Á, Lause J. Delving into LLM-assisted writing in biomedical publications through excess vocabulary. *Science Advances* 11, eadt3813 (2025). doi.org/10.1126/sciadv.adt3813

17. Hao Q, Xu F, Li Y, Evans J. Artificial intelligence tools expand scientists' impact but contract science's focus. *Nature* 649, 1237–1243 (2026). doi.org/10.1038/s41586-025-09922-y
18. Yu AZ, Ronen S, Hu K, Lu T, Hidalgo CA. Pantheon 1.0, a manually verified dataset of globally famous biographies. *Scientific Data* 3, 150075 (2016). doi.org/10.1038/sdata.2015.75
19. Klein Goldewijk K, Beusen A, Doelman J, Stehfest E. Anthropogenic land use estimates for the Holocene, HYDE 3.2. *Earth System Science Data* 9, 927–953 (2017). doi.org/10.5194/essd-9-927-2017
20. Crossref. Crossref REST API. 2026. https://api.crossref.org/
21. Priem J, Piwowar H, Orr R. OpenAlex: A fully-open index of scholarly works, authors, venues, institutions, and concepts. 2022. https://arxiv.org/abs/2205.01833
22. National Library of Medicine. PubMed, queried through the NCBI Entrez programming utilities (E-utilities). 2026. https://pubmed.ncbi.nlm.nih.gov/
23. arXiv. Submissions by category since 2022. 2026. https://info.arxiv.org/about/reports/submission_category_by_year.html
24. bioRxiv. bioRxiv API summary. 2026. https://api.biorxiv.org/
25. Our World in Data. Grapher data downloads (population, books, patents, literacy, computing, AI training compute). 2026. https://ourworldindata.org
26. Berman HM, Westbrook J, Feng Z, et al. The Protein Data Bank. *Nucleic Acids Research* 28, 235–242 (2000). doi.org/10.1093/nar/28.1.235
27. Sayers EW, Cavanaugh M, Frisse L, et al. GenBank 2025 update. *Nucleic Acids Research* 53, D56–D61 (2025). doi.org/10.1093/nar/gkae1114
28. Katz K, Shutov O, Lapoint R, Kimelman M, Brister JR, O'Sullivan C. The Sequence Read Archive: A decade more of explosive growth. *Nucleic Acids Research* 50, D387–D390 (2022). doi.org/10.1093/nar/gkab1053
29. The UniProt Consortium. UniProt: The universal protein knowledgebase in 2025. *Nucleic Acids Research* 53, D609–D617 (2025). doi.org/10.1093/nar/gkae1010
30. Christiansen JL, McElroy DL, Harbut M, et al. The NASA Exoplanet Archive and exoplanet follow-up observing program: Data, tools, and usage. *The Planetary Science Journal* 6, 186 (2025). doi.org/10.3847/PSJ/ade3c2
31. Center for Near Earth Object Studies (CNEOS), NASA Jet Propulsion Laboratory. Discovery statistics. 2026. https://cneos.jpl.nasa.gov/stats/totals.html
32. Wikimedia Foundation. Wikimedia analytics API. 2026. https://doc.wikimedia.org/generated-data-platform/aqs/analytics-api/

Appendix

Table 3. All 27 records, sorted by the ratio of the next decade to everything before. Columns as in Table 2.

Record	Counts	Total so far	Growth a year	Doubling time (years)	Next decade ÷ all before	Likely range
GenBank	DNA bases	6.7 trillion bases	75.3%	1.2	150	10.2–256
OpenAlex deep learning (to 2023)	papers tagged deep learning	279,069	55.5%	1.6	52.3	34.1–86.6
arXiv AI (cs.AI)	AI papers	156,573	41.8%	2.0	31.1	22.2–46.0
GenBank genome projects	DNA bases, whole genomes	42.1 trillion bases	37.2%	2.2	18.1	1.22–43.7
arXiv language (cs.CL)	language-technology papers	99,110	32.9%	2.4	15.6	11.1–23.1
Sequence Read Archive (to 2023)	DNA bases, raw reads	88.7 quadrillion bases	33.5%	2.4	11.2	0.76–27.0
arXiv machine learning (cs.LG)	machine-learning papers	247,045	30.7%	2.6	10.7	7.65–15.9
bioRxiv	biology preprints	312,221	23.1%	3.3	5.88	4.19–8.70
OpenAlex preprints	preprints	6.0 million	20.3%	3.8	4.26	3.42–5.65
medRxiv	medical preprints	77,157	1.1%	61	2.07	1.47–3.06
OpenAlex machine learning (to 2023)	papers tagged machine learning	3.6 million	11.7%	6.3	1.80	1.43–2.43
arXiv	preprints	2.9 million	9.9%	7.4	1.70	1.65–1.84
Near-Earth asteroids	asteroids found	40,583	6.7%	10.7	1.16	0.08–1.99
Protein Data Bank	protein structures	247,225	5.3%	13.3	0.96	0.07–1.64
OpenAlex AI (to 2023)	papers tagged AI	14.9 million	6.4%	11.2	0.95	0.73–1.32
Crossref	scholarly works with a DOI	167 million	8.0%	9.1	0.93	0.84–1.08
Crossref articles	journal articles with a DOI	119 million	7.0%	10.3	0.82	0.77–0.94
OpenAlex	indexed scholarly works	293 million	1.9%	36	0.56	0.47–0.71
PubMed	biomedical papers	40.1 million	4.9%	14.5	0.56	0.49–0.67
US patents (to 2020)	patents granted	10.8 million	4.4%	16.1	0.41	0.33–0.56
OpenAlex articles	journal articles	192 million	−0.2%	falling	0.36	0.30–0.46
Wikidata	items	120 million	−7.9%	falling	0.30	0.02–0.72
NASA Exoplanet Archive	confirmed planets	6,089	−5.6%	falling	0.30	0.02–0.72
OpenAlex books	books	5.8 million	−1.6%	falling	0.25	0.21–0.32
English Wikipedia	articles	7.1 million	−3.9%	falling	0.23	0.02–0.56
GenBank records	DNA sequence records	260 million	−1.9%	falling	0.18	0.01–0.32
Swiss-Prot	reviewed proteins	574,369	−6.6%	falling	0.02	0.00–0.03