Tracking artificial intelligence in climate inventions with patent data

Citation for the original published paper (version of record):
Verendel, V. (2023). Tracking artificial intelligence in climate inventions with patent data. Nature Climate Change, 13(1): 40-47. http://dx.doi.org/10.1038/s41558-022-01536-w

nature climate change

Article https://doi.org/10.1038/s41558-022-01536-w

Tracking artificial intelligence in climate inventions with patent data

Received: 8 December 2021

Accepted: 9 October 2022

Published online: 12 January 2023

Check for updates

Vilhelm Verendel

Artificial intelligence (AI) is spreading rapidly in many technology areas, and AI inventions may help climate change mitigation and adaptation. Previous studies of climate-related AI mainly rely on expert studies of literature, not large-scale data. Here I present an approach to track the relation between AI and climate inventions on an economy-wide scale. Analysis of over 6 million US patents, 1976 to 2019, shows that within climate patents, AI is referred to most often in transportation, energy and industrial production technologies. In highly cited patents, AI occurs more frequently in adaptation and transport than in other climate mitigation areas. AI in climate patents was associated with around 30–100% more subsequent inventions when counting all technologies. Yet AI-climate patents led to a greater share of citations from outside the climate field than non-AI-climate patents. This suggests the importance of tracking both increased invention activity and the areas where subsequent inventions emerge.

A range of artificial intelligence (AI) technologies are rapidly being developed with high expectations of technological innovation and economic growth[1–3]. AI could contribute to increasingly effective climate change mitigation and adaptation technologies in multiple areas[4–6]. However, an increasing capability to automate and transform production, equip industries with new tools and draw increasing commercial support also means that AI technologies could lead to a higher demand for computing power, larger carbon footprints, shifts in patterns of electricity demand and an accelerated depletion of natural resources[7–10]. High expectations of new technologies with limited experience suggest a risk of unjustified techno-optimism, which could delay effective climate policy[11]. Whether the net effect of AI on the climate system will be ameliorative or detrimental is currently an open question, and concerns about the impact of AI have been followed by calls for new regulations and increased international oversight[12–16]. This suggests a need for improved capabilities to track, examine and analyse these emerging technologies. Here I use large-scale patent data to track AI inventions in technologies that can contribute to climate adaptation and mitigation.

The initial research into the connection between AI and climate change has often been framed in terms of the United Nations Sustainable Development Goals and conducted as expert studies. These have indicated both positive and negative effects of AI[4,17–19]. For climate change, it has been suggested that machine learning could have broad potential in both mitigation and adaptation strategies, with a mixed message regarding the potential net effect on the climate system[20–24]. An advantage of expert-based reviews is the possibility to integrate knowledge from different domains, even when data are scarce. However, experts often find it challenging to unpack and fully explain their partially automatic judgement processes[25]. Moreover, expertise tends to be difficult to translate from one domain into another[26]. Scaling up to cover a larger literature is a challenge for any team of experts, and an interesting option would be to complement the analysis with other data sources. Here I investigate how to use large data sources from national patent offices and intellectual property organizations regularly used to monitor inventions and innovation in large economies.

Patents are possibly the most detailed track record of modern technological inventions[27–29], allowing individuals and organizations to protect the use of their patented inventions typically for years. National patent offices have organized and classified millions of patents using international classification systems. The resulting classification codes provide the primary means to group patents and make them searchable: patent offices need to examine the prior art to judge whether claimed inventions are sufficiently novel before granting patents. Patents have previously been used as a data source and a proxy to study trends in both AI[1,2,30] and climate technologies[31–34] separately. Here I combined classification data for climate inventions and AI technologies to find patents that are both. First, I used the Y02

Department of Physics, Chalmers University of Technology, Gothenburg, Sweden. e-mail: vive@chalmers.se

https://doi.org/10.1038/s41558-022-01536-w

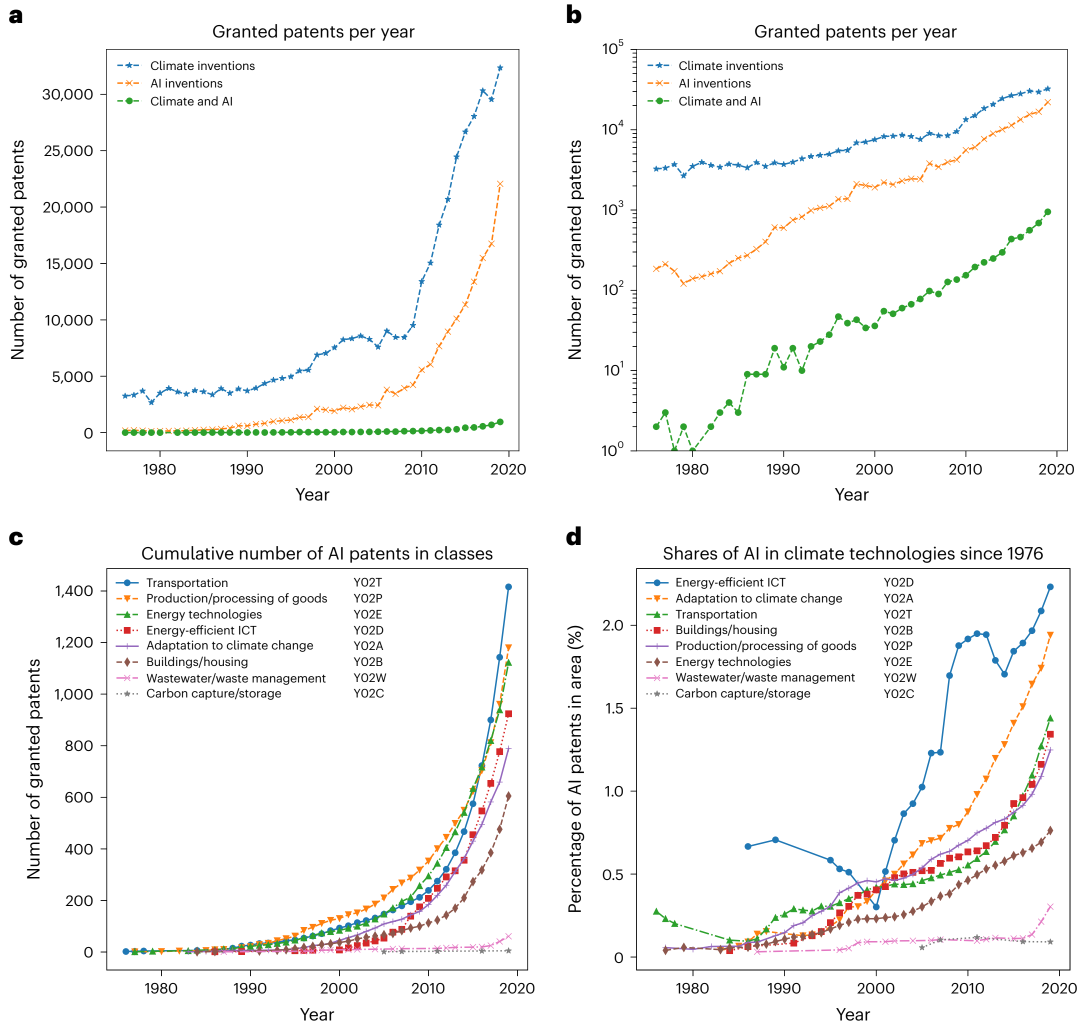

Fig. 1 | AI and climate patent counts and shares from 1976 and onwards. **a**, Granted patents per year, with a steeper rise starting around 2010. **b**, The rise in **a** can be seen as exponential growth in climate AI patents (linear on a log scale), and this holds for climate patents and AI patents separately. Within climate patents, however, AI patents are not growing exponentially: AI is associated with an approximately linear growth in shares (Supplementary Information). **c**,**d**, Transportation, energy and industrial production mitigation technologies have accumulated the most AI patents, while the smaller classes of energy-efficient ICT and adaptation patents have larger shares of AI inventions. Energy-efficient ICT cover inventions that reduce energy use within ICT equipment, but not ICT used to reduce energy use in a further piece of equipment. For readability, I have shortened the official names[35]. The official names and classification codes can be found in the Supplementary Information.

classification system initiated by the European Patent Office[35] to monitor selected technologies related to mitigating or adapting to climate change[31]. Second, I found AI patents with a recent method developed by the World Intellectual Property Organization (WIPO) that can be automated computationally[1,36]. The WIPO method classifies patents as AI on the basis of patent classification codes and by matching certain keywords from key sections in the patent texts, including terms such as 'machine learning', 'deep learning' and 'natural language processing'; more details can be found in the Supplementary Information. Third, I combined both of these classification approaches to find patents that are labelled as both. A few example patents that are classified as both AI and climate inventions are referred to and presented in the Supplementary Information.

As useful as patent data can be, it is also essential to understand some of the limitations of using patents and avoid unwarranted generalization from patented inventions to the population of all inventions, for reasons that follow. First, certain types of inventions may not even be possible to patent: it is currently not possible to patent entirely abstract inventions (for example, pure mathematical results). In the United States, patent claims that include abstract inventions for algorithms and computer software require a link to a practical application[37]. In the European Union, patents with abstract invention claims need to have a technical character—for example, controlling some physical process or providing an implementation or function that solves a particular technical problem[38]. The differences between what patent laws permit can sometimes be subtle, so patterns based on patentability might be specific to a country under consideration. Here I used data from the United States, which should be seen as one case study, although an interesting one. The approach used here is possible to extend for analysing patents from other countries. Besides variation in national patent laws, other factors could need to be controlled when investigating and comparing patenting trends. Such differences include technical expertise and specialization in countries and industries[33].

Second, inventors do not always apply for patents, even when patenting is possible. Some AI inventions are being made available through an open-source culture where inventions are neither patented nor protected by secrecy, which can be seen in academia and public-interest AI research. However, an initial investigation of patented

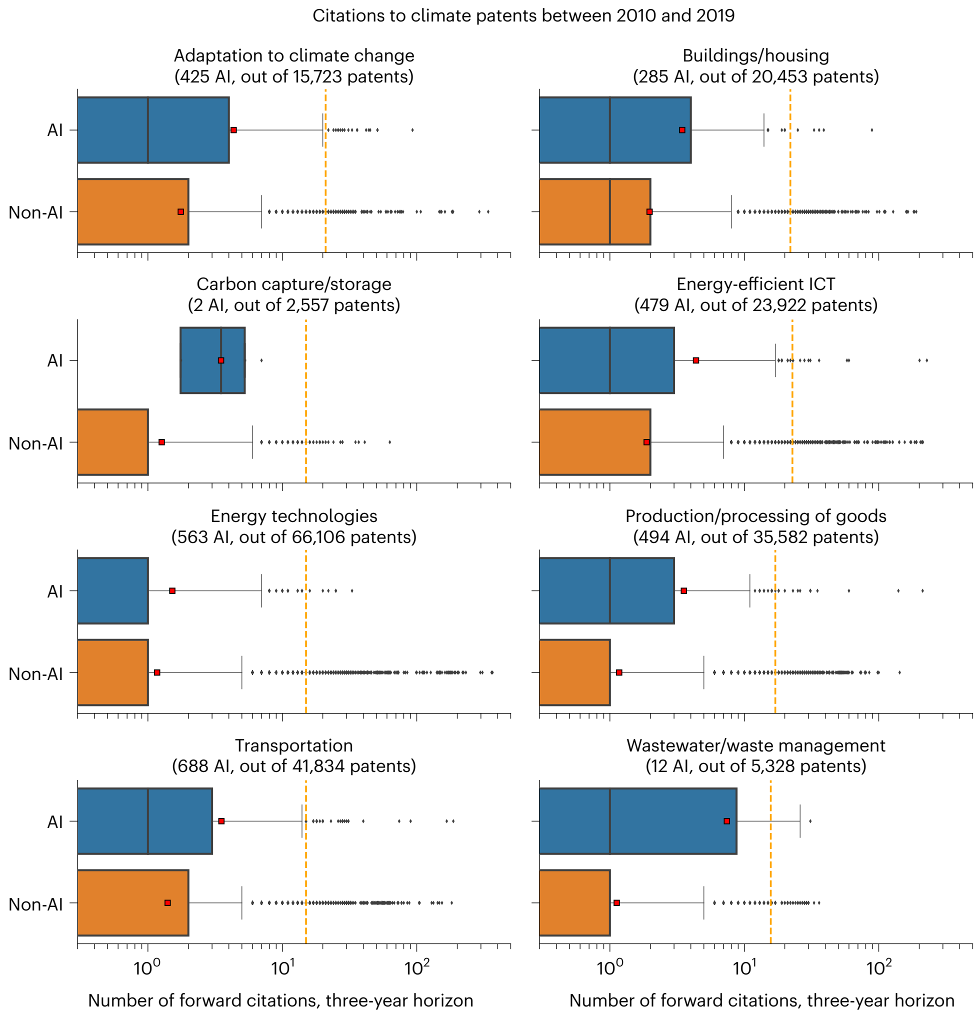

Fig. 2 | Comparing AI and non-AI citation counts in climate patents (citations to patents granted between 2010 and 2017). AI patents are on average associated with more subsequent inventions when counting citations from all subsequent technologies. The total count of AI breakthroughs is smaller, as expected because of fewer AI patents overall. In each plot, the centre line is the median, and the red square shows the mean. The dashed vertical lines show the boundaries for breakthroughs (for forward citation counts above the 99th percentile) over all years. The box limits show the 25th and 75th percentiles, with whiskers at the 5th and 95th percentiles. For plotting, a log-plus-one transformation was used.

AI inventions shows that these are increasingly related to commercial patent rights, and at the same time the share of patents that depend on public government support is getting lower (Supplementary Information). Furthermore, firms and individuals also choose to protect some inventions with secrecy instead of filing patents. The incentives for secrecy vary between technological fields. For example, in the United States, the pharmaceutical and biomedical industries with high cost, high uncertainty and long innovation cycles rely more on patents than the software industry[39]. Software inventions often have lower costs and can have innovation cycles on a timescale of days or weeks rather than months or years. When patenting processes take longer than the innovation cycle, patenting might lose some of the appeal of getting expected rewards from innovation. For these reasons, a share of AI inventions can be expected not to be found in patents. To the best of my knowledge, this share is unknown and is a knowledge gap in the literature. AI technologies are being invented and used in various industries[1,2], so this share for AI probably depends on the incentives in several technological domains that may differ. That we do not know the share of inventions that are protected by secrecy suggests a need to be cautious about generalizing from patents to other non-patented inventions.

The data are as follows. First, I collected historical data on over six million granted patents publicly available[40,41] from the US Patent and Trademark Office for the period from 1976 to 2019, up to when the WIPO method for finding AI patents was developed and evaluated. I worked with data from the United States because it is a leading economy and because the US institutions have made patent full-text data readily available[41]. Previous work[42] indicates that US patents have been found to well represent the frontier of technological innovation in low-carbon energy innovation, which is part of the scope. Second, I extracted technology classification data for the patents, including current labels for climate inventions[43]: the Cooperative Patent Classification (CPC Y02) code "covers selected technologies, which control, reduce or prevent anthropogenic emissions of greenhouse gases in the framework of the Kyoto Protocol and the Paris Agreement, and also technologies which allow adapting to the adverse effects of climate change".[35,44] Third, I applied the WIPO method[1,36] described above to label the same patents as AI or non-AI for further analysis. For details about the classifications, see the Methods and the Supplementary Information. Precise summary metrics for AI and climate patents are also found in the Supplementary Information.

It is natural to first look at the data by aggregating mitigation and adaptation technologies, but then later disaggregate these into separate groups. Both climate and AI patents have seen clear growth in the past decade (Fig. 1a). AI and climate invention patent counts, as well

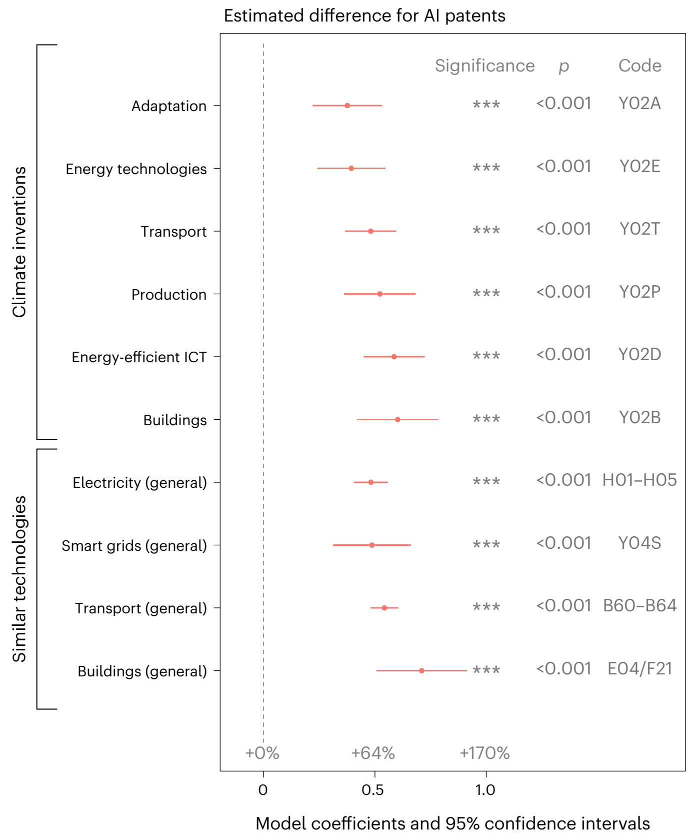

Fig. 3 | The estimated predictive difference of AI is in the range of 30–100% more subsequent inventions in the groups of climate patents, as in technologies with a similar function. The analysis estimates AI in climate invention areas and in more general technological areas using count regression models. When estimating the same predictive difference on similar groups of technologies (but without a clear connection to climate), the rankings of the AI coefficients are similar. The dot-and-whisker plot shows estimated coefficients with confidence intervals[55].

as the climate AI patents, have undergone exponential growth during the past decade (Fig. 1b). However, note that this does not mean that the share of AI within climate patents is growing exponentially: the growth of the share of AI within climate patents has been approximately linear, and AI climate patent counts are actually *lower* than expected if AI and climate innovations had been statistically independent (Supplementary Information). More than half of all AI inventions in climate patents since 1976 are found in technologies for transportation, energy and production (Fig. 1c). Climate adaptation and building/housing mitigation technology patents involving AI are somewhat lower in absolute numbers. For waste management and carbon capture/storage, there are very little data on AI. Energy-efficient information and communication technologies (ICT) and adaptation patents are areas where AI has had larger overall shares in the past few years (Fig. 1d). The number of unique patents found to involve both climate and AI between 1976 and 2019 is 4,390. This is around 1.5% of the total climate patents and 2.7% of the AI patents.

To examine whether AI makes a difference in climate inventions, I chose to analyse the number of citations from subsequent patents that cite back to previously granted patents. For a given patent, the forward citation count reflects the number of subsequent patents that relate to or build on it—or, put differently, cite it. Forward citation counts have been considered to be important indicators of the technological impact of a patented invention[45]. Harhoff et al.[46] found that the economic value of individual patents, measured through a survey with patent assignees, is positively correlated with the number of forward citations. Hall et al.[47] also showed that the number of forward citations per patent correlated positively with the market value of firms, and they estimated that if a firm's quality of patents increases so that the patents receive one additional citation, on average, the firm's market value increases by 3%. Moreover, forward citations are positively correlated with patent assignees' willingness to pay renewal fees[48], which indicates the economic value of cited patents. Furthermore, forward citations can also be used to investigate knowledge spillovers, or how knowledge from technologies in one area is useful in different areas[49,50]. In the analysis that follows, I distinguished between technological domains that cite back to previously granted AI and non-AI patents. Finally, forward citations have been used to investigate highly cited technological breakthroughs by using the accumulated forward citations in the years after which a patent was granted. Squicciarini et al.[51] define breakthrough inventions as the top 1% cited documents for each year, and they use a three-year window from the patent grant date to accumulate forward citation counts. Benson and Magee[27] constructed a metric that they term "immediate importance" as the average number of citations that a patent receives within three years of publication. Consistent with the previous literature, I examined the predictive difference associated with AI on a three-year horizon after patents have been granted. The breakthrough inventions were defined to be the top 1% cited patents in a technical domain per year.

To examine whether AI is associated with a difference in forward citations, it is natural to distinguish between AI and non-AI in the groups of climate patents described in Fig. 1. An initial exploration revealed two aspects about the target variable, shown in Fig. 2. First, the average

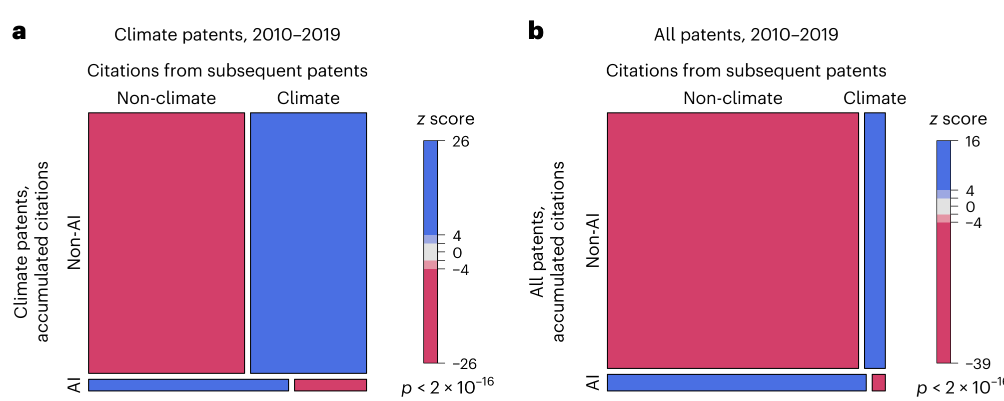

Fig. 4 | For the group of climate patents, AI is associated with a smaller share of spillovers to subsequent climate patents than non-AI technologies. **a**,**b**, The results in the mosaic plots[56] are statistically significant under a null model given by the hypergeometric distribution for citation counts in technological networks[53,54] and hold in both the aggregate populations of climate patents (**a**) and all patents (**b**). The data are from the period 2010–2019. Similar patterns hold when disaggregating the analysis into several more specific groups of climate inventions.

https://doi.org/10.1038/s41558-022-01536-w

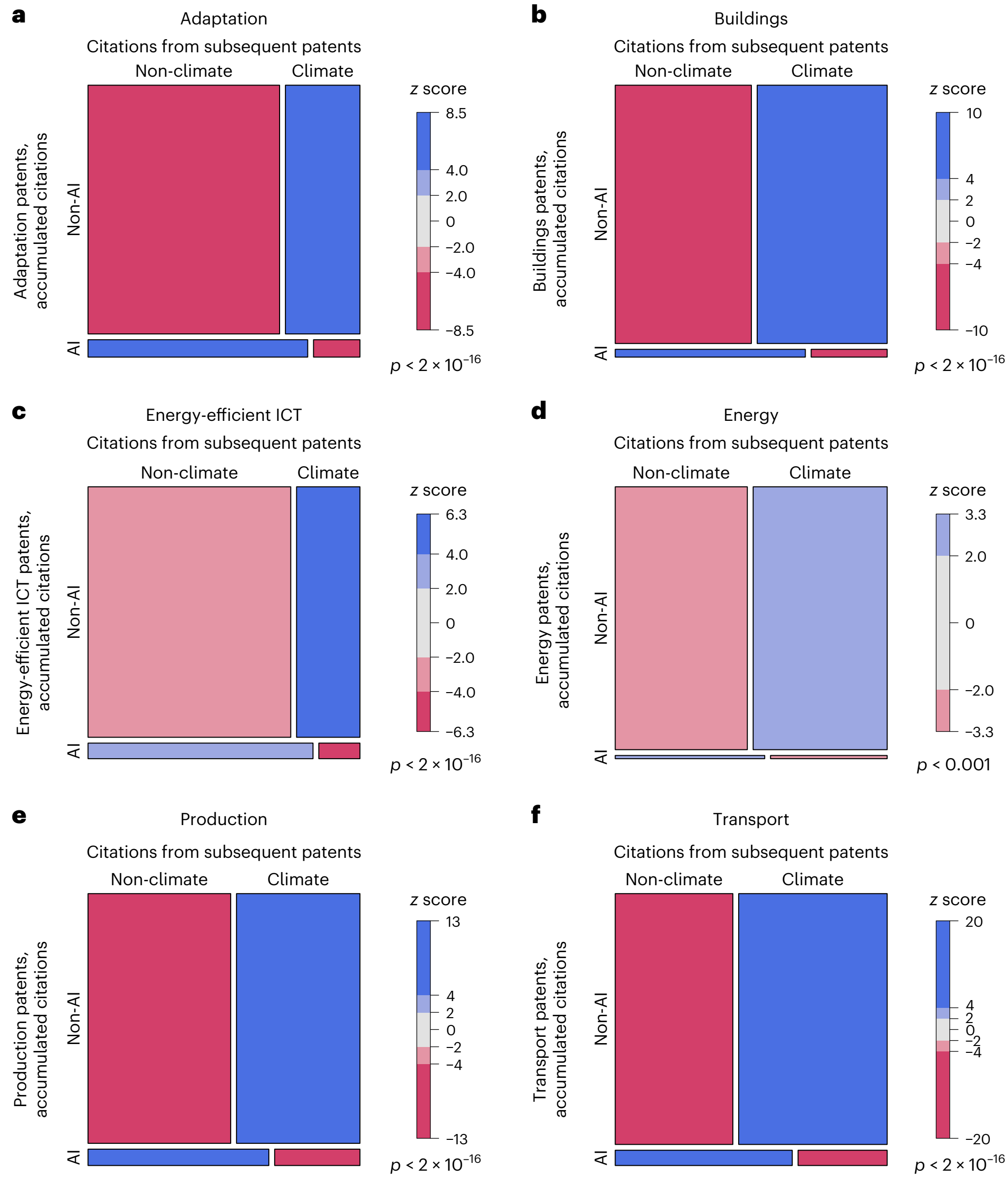

Fig. 5 | For the groups of climate patents used in the analysis, AI is associated with a smaller share of spillovers to climate patents than from non-AI technologies. a–f, Here climate patents have been disaggregated into the different Y02 areas of the CPC. The data are from the period 2010–2019.

forward citation count for climate AI inventions is greater than for climate non-AI inventions. Second, zooming in on the highly cited breakthroughs (the highest counts in Fig. 2), a large majority of the most highly cited breakthroughs appear to be non-AI inventions. In other words, AI patents are related to more subsequent inventions on average but seemingly fewer highly cited breakthroughs. However, this does not address the fact that AI has a much smaller share of the total patents. This suggests estimating the predictive difference of AI on the average forward citation counts by including controls and testing for differences in breakthroughs after adjusting for the group size. However, carbon capture/storage and waste technology patents were left out, as the number of AI patents is too small for reasonable statistical analysis.

To estimate the predictive difference of AI on patent forward citation counts, I used count regression modelling of the forward citations on a three-year horizon, limited to patents granted in the previous decade (a three-year horizon, so patents from 2010 to 2017). I controlled for the difference in year, technological areas that patents are from and other factors in line with previous work on modelling forward citations for patents[29] (for details, see the Methods and Supplementary Information). To use control groups for wider context, I repeated the regression modelling for the climate invention areas and control groups given by similar technological domains (based on CPC classifications). The control groups have related technological functions but broader than climate inventions: buildings, electricity, smart grids and transport technologies in general, not restricted to climate relevance.

For the climate patents as well as the control groups, AI seems associated with more subsequent inventions even after controlling for other factors (Fig. 3). In groups of climate inventions, AI was associated with a 30–100% increase in forward citations, with predictive differences being statistically significant. Among the technologies with climate inventions, buildings and energy-efficient ICT showed the greatest increases related to AI, with adaptation and energy technologies on average showing a weaker difference. For the control groups, the ranking and effect sizes are similar to the groups with climate inventions: the coefficients in Fig. 3 are in a similar range as those found for related

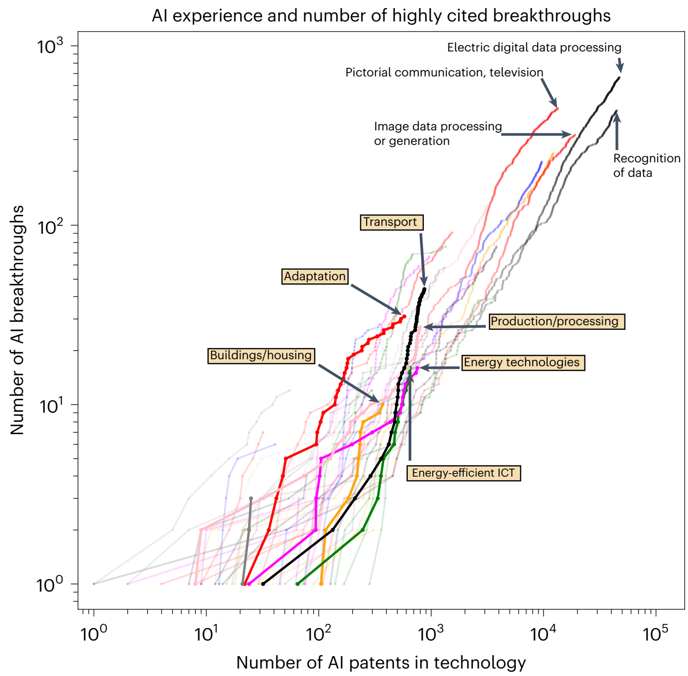

Fig. 6 | In climate adaptation technologies and transport technologies, the highly-cited breakthroughs have made up a larger share of AI patents than in other groups of mitigation technologies. Cumulative experience of AI inventions in technological areas and AI breakthrough counts in the same area. The trajectories represent adaptation and mitigation technologies, as well as the 50 largest technologies (largest by patent volume). Transport technologies have the most observed breakthroughs, in line with having the most AI patents. Patents up to the end of 2017.

technologies such as electricity, transport and building technologies in general. An analysis for other control groups less related to climate inventions showed that differences between AI and climate inventions can be stronger or weaker than in other technological domains (Supplementary Information).

So far the results show that AI inventions in climate patents were on average cited more from the group of all subsequent patents. However, this does not show us in which technological domains these subsequent inventions emerge. The idea that inventions can benefit some technologies by supporting subsequent inventions in certain technological areas more than in others can be framed in terms of knowledge spillovers[50,52,53]. In this case, one can examine whether AI patents get cited from technology domains that are either climate inventions or not. I examined whether AI and non-AI patents get cited to any different degree by climate patents as follows. For cited patents, I distinguished whether these are AI or non-AI climate innovations. For citing patents (the spillovers), I distinguished between climate and non-climate inventions. Aggregate patent citation count networks between technology domains are known to depend on several factors, such as technology domain size and average age[54]. Recent work[53,54] has shown that it is possible to control for domain size and average patent age and to test whether differences would arise just because of random patent citations with a null model based on the hypergeometric distribution.[53]

The results show that AI in climate patents is associated with a smaller share of spillovers to climate inventions than spillovers from other non-AI technologies (Fig. 4a). Within patents in general, it would be reasonable to expect AI spillovers to primarily be related to AI technologies rather than climate inventions (Fig. 4b). However, within climate patents, a larger share of spillovers than non-AI technologies are non-climate, suggesting that knowledge from the climate AI patents was more useful in other areas. Similar results hold when disaggregating patents into groups of climate adaptation and mitigation (Fig. 5). Taken together with the results above, this suggests a double association for AI in climate inventions: AI has been related to increased activity in subsequent inventions but also a smaller share of spillovers to climate patents than from non-AI technologies. Citation counts cannot be expected to always reflect the actual usefulness of individual technologies in practice, and citations are one of several ways to track the overall importance of new technologies. Therefore, tracking AI in climate patents will require distinguishing between the direct impact from more subsequent inventions and how knowledge spillovers are distributed between areas to improve our knowledge about the net effects of AI inventions.

Finally, I examined whether AI has any relation to highly cited patents, also termed as breakthroughs. An association between AI and highly cited breakthroughs would indicate where applications of AI have been more interesting or particularly valuable. Figure 2 shows fewer highly cited AI breakthroughs in total, possibly because of the smaller number of AI patents overall. I considered in each group of climate inventions the 1% patents per year with the most forward citations in the three years following publication. I then took the cumulative experience of AI inventions as the total count of AI patents in the area to compare technological domains of different sizes and the accumulated experience of patenting with AI.

Most groups of climate mitigation technologies have been associated with similar AI breakthrough shares (defined as the number of AI-related breakthroughs per AI invention in that technology) as other technologies (Fig. 6). For adaptation technologies, AI breakthroughs were initially higher compared with other technologies; among mitigation technologies, transport is clearly leading in shares. A quantile test (Methods and Supplementary Information) also suggests that the share of AI breakthroughs has been higher in climate adaptation and transport technologies. Estimates for the other groups are more uncertain: wide confidence intervals (Supplementary Information) suggest that the current evidence is too weak to strongly rule one way or the other about AI breakthroughs in most areas of climate mitigation, besides transport patents and in contrast to climate adaptation technologies. For most groups of climate mitigation technologies, the uncertainty means that the jury is still out with respect to the role of AI in breakthroughs.

Taken together, the results show that AI in climate patents is associated with more subsequent patents but also a larger share of knowledge spillovers to non-climate technical domains than non-AI technologies. This suggests that the analysis of AI in climate inventions should also consider the impact on other technological fields that may benefit more from these inventions than from non-AI climate inventions. AI has been associated with a higher share of breakthroughs in climate adaptation and transportation patents than in other groups of technologies with a potential for climate mitigation. These results are for one country, but the approach can be used to study other countries and regions. Caution is also needed to avoid unjustified generalization to inventions and innovations beyond those covered by patents and being aware that the criteria for patentability can vary between countries.

New inventions and technological breakthroughs may meaningfully contribute to addressing climate change. More capabilities are necessary to track the emerging technologies for which both risks and promises exists but where the use of large-scale data is still scarce. Using patent data, we can better track AI in technologies to adapt and mitigate climate change.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41558-022-01536-w.

References

1. *Technology Trends 2019: Artificial Intelligence* (World Intellectual Property Organization, 2019).
2. Cockburn, I., Henderson, R. & Stern S. in *The Economics of Artificial Intelligence: An Agenda* (eds Agrawal A. K., Gans J. & Goldfarb A.) 115–146 (Univ. of Chicago Press, 2019).
3. Aghion, P., Jones, B. F. & Jones, C. I. *Artificial Intelligence and Economic Growth* (Univ. of Chicago Press, 2019).
4. Vinuesa, R. et al. The role of artificial intelligence in achieving the Sustainable Development Goals. *Nat. Commun.* **11**, 233 (2020).
5. Tomašev, N. et al. AI for social good: unlocking the opportunity for positive impact. *Nat. Commun.* **11**, 2468 (2020).
6. Ardabili, S., Mosavi, A., Dehghani, M. & Várkonyi-Kóczy, A.R. Deep learning and machine learning in hydrological processes climate change and Earth systems: a systematic review. In *Engineering for Sustainable Future. INTER-ACADEMIA 2019* (ed. Várkonyi-Kóczy, A.) 52–62 (Springer, 2020).
7. Mora, C. et al. Bitcoin emissions alone could push global warming above 2 C. *Nat. Clim. Change* **8**, 931–933 (2018).
8. Jones, N. How to stop data centres from gobbling up the world's electricity. *Nature* **561**, 163–167 (2018).
9. Brevini, B. Black boxes, not green: mythologizing artificial intelligence and omitting the environment. *Big Data Soc.* **7** https://doi.org/10.1177/2053951720935141 (2020).
10. Masanet, E., Shehabi, A., Lei, N., Smith, S. & Koomey, J. Recalibrating global data center energy-use estimates. *Science* **367**, 984–986 (2020).
11. McLaren, D. & Markusson, N. The co-evolution of technological promises, modelling, policies and climate change targets. *Nat. Clim. Change* **10**, 392–397 (2020).
12. Dauvergne, P. Is artificial intelligence greening global supply chains? Exposing the political economy of environmental costs. *Rev. Int. Polit. Econ.* **29**, 696–718 (2022).
13. Cihon, P., Maas, M. M. & Kemp, L. Fragmentation and the future: investigating architectures for international AI governance. *Glob. Policy* **11**, 545–556 (2020).
14. Jelinek, T., Wallach, W. & Kerimi, D. Policy brief: the creation of a G20 coordinating committee for the governance of artificial intelligence. *AI Ethics* **1**, 141–150 (2021).
15. Coeckelbergh, M. AI for climate: freedom, justice, and other ethical and political challenges. *AI Ethics* **1**, 67–72 (2021).
16. Gailhofer, P. et al. *The Role of Artificial Intelligence in the European Green Deal* (Policy Department for Economic, Scientific and Quality of Life Policies, European Parliament, 2021).
17. Khakurel, J., Penzenstadler, B., Porras, J., Knutas, A. & Zhang, W. The rise of artificial intelligence under the lens of sustainability. *Technologies* **6**, 100 (2018).
18. Sovacool, B. K. & Griffiths, S. Culture and low-carbon energy transitions. *Nat. Sustain.* **3**, 685–693 (2020).
19. Sætra, H. S. AI in context and the Sustainable Development Goals: factoring in the unsustainability of the sociotechnical system. *Sustainability* **13**, 1738 (2021).
20. Rolnick, D. et al. Tackling climate change with machine learning. *ACM Computing Surveys* **55**, 42 (2022).
21. Lacoste, A., Luccioni, A., Schmidt, V. & Dandres, T. Quantifying the carbon emissions of machine learning. Preprint at *arXiv*. https://doi.org/10.48550/arXiv.1910.09700 (2019).
22. Henderson, P. et al. Towards the systematic reporting of the energy and carbon footprints of machine learning. *J. Mach. Learn. Res.* **21**, 1–43 (2020).
23. Schwartz, R., Dodge, J., Smith, N. A. & Etzioni, O. Green AI. *Commun. ACM* **63**, 54–63 (2020).
24. Kaack, L. H. et al. Aligning artificial intelligence with climate change mitigation. *Nat. Clim. Chang.* **12**, 518–527 (2022).
25. Hastie, R. & Dawes, R. M. *Rational Choice in an Uncertain World: The Psychology of Judgment and Decision Making* (Sage, 2009).
26. Bedard, J. & Chi, M. T. Expertise. *Curr. Dir. Psychol. Sci.* **1**, 135–139 (1992).
27. Benson, C. L. & Magee, C. L. Quantitative determination of technological improvement from patent data. *PLoS ONE* **10**, e0121635 (2015).
28. Bettencourt, L. M., Trancik, J. E. & Kaur, J. Determinants of the pace of global innovation in energy technologies. *PLoS ONE* **8**, e67864 (2013).
29. Barbieri, N., Marzucchi, A. & Rizzo, U. Knowledge sources and impacts on subsequent inventions: do green technologies differ from non-green ones? *Res. Policy* **49**, 103901 (2020).
30. Toole, A. et al. *Inventing AI: Tracing the Diffusion of Artificial Intelligence with US Patents* (US Patent and Trademark Office, 2020).
31. Popp, D. *Environmental Policy and Innovation: A Decade of Research* Working Paper (National Bureau of Economic Research, 2019).
32. Haščič, I. & Migotto, M. *Measuring Environmental Innovation Using Patent Data* OECD Environment Working Papers No. 89 (OECD, 2015).
33. Probst, B., Touboul, S., Glachant, M. & Dechezleprêtre, A. Global trends in the invention and diffusion of climate change mitigation technologies. *Nat. Energy* **6**, 1077–1086 (2021).
34. Pasimeni, F., Fiorini, A. & Georgakaki, A. International landscape of the inventive activity on climate change mitigation technologies. a patent analysis. *Energy Strategy Rev.* **36**, 100677 (2021).
35. Angelucci, S., Hurtado-Albir, F. J. & Volpe, A. Supporting global initiatives on climate change: the EPO's 'Y02-Y04S' tagging scheme. *World Pat. Inf.* **54**, S85–S92 (2018).
36. *Technology Trends 2019: Artificial Intelligence: Data Collection Method and Clustering Scheme: Background Paper* (World Intellectual Property Organization, 2019).
37. *Revised Patent Subject Matter Eligibility Guidance* (US Patent and Trademark Office, 2019); https://www.uspto.gov/sites/default/files/documents/peg_oct_2019_update.pdf
38. *EPO Case Law of the Boards of Appeal: 9.1.1 Technical Character of an Invention* (EPO, accessed 22 March 2022); https://www.epo.org/law-practice/legal-texts/html/caselaw/2019/e/clr_i_d_9_1_1.htm
39. Graham, S. J., Merges, R. P., Samuelson, P. & Sichelman, T. High technology entrepreneurs and the patent system: results of the 2008 Berkeley Patent Survey. *Berkeley Technol. Law J.* **24**, 1255–1327 (2009).
40. Marco, A. C., Myers, A., Graham, S. J., D'Agostino, P. & Apple, K. *The USPTO Patent Assignment Dataset: Descriptions and Analysis* (US Patent and Trademark Office, 2015).
41. Marco, A. C., Carley, M., Jackson, S. & Myers, A. *The USPTO Historical Patent Data Files: Two Centuries of Innovation* (US Patent and Trademark Office, 2015); https://bulkdata.uspto.gov/data/patent/grant/redbook/fulltext/1976,...,2019
42. Albino, V., Ardito, L., Dangelico, R. M. & Petruzzelli, A. M. Understanding the development trends of low-carbon energy technologies: a patent analysis. *Appl. Energy* **135**, 836–854 (2014).
43. *Cooperative Patent Classification (CPC) Master Classification File (MCF) for U.S. Patent Grants (Jul 1790–Dec 2022)* (US Patent and Trademark Office, accessed 1 August 2021); https://bulkdata.uspto.gov/data/patent/classification/cpc/
44. *CPC Y02 Description* (US Patent and Trademark Office, accessed 1 December 2021); https://www.uspto.gov/web/patents/classification/cpc/html/cpc-Y.html

45. Archibugi, D. & Pianta, M. Innovation surveys and patents as technology indicators: the state of the art. In *Innovation, Patents and Technological Strategies* 17–56 (OECD, 1996).
46. Harhoff, D et al. *Citation Frequency and the Value of Patented Innovation* Tech. Rep. (Wissenschaftszentrum Berlin, Research Unit: Competition and Innovation, 1997).
47. Hall, B. H., Jaffe, A. & Trajtenberg, M. Market value and patent citations. *RAND J. Econ.* **36**, 16–38 (2005).
48. Hegde, D. & Sampat, B. Examiner citations, applicant citations, and the private value of patents. *Econ. Lett.* **105**, 287–289 (2009).
49. Jaffe, A. B., Trajtenberg, M. & Henderson, R. Geographic localization of knowledge spillovers as evidenced by patent citations. *Q. J. Econ.* **108**, 577–598 (1993).
50. Hall, B. H., Mairesse, J., & Mohnen, P. In *Handbook of the Economics of Innovation* (eds Hall, B. H. & Rosenberg N.) 1033–1082 (North-Holland, 2010).
51. Squicciarini, M., Dernis, H. & Criscuolo, C. *Measuring Patent Quality: Indicators of Technological and Economic Value* OECD Science, Technology and Industry Working Paper No. 2013/03 (OECD, 2013).
52. Ertur, C. & Koch, W. Growth, technological interdependence and spatial externalities: theory and evidence. *J. Appl. Econometr.* **22**, 1033–1062 (2007).
53. Pichler, A., Lafond, F. & Farmer, J. D. *Technological Interdependencies Predict Innovation Dynamics* INET Oxford Working Paper 2020-04 (Institute for New Economic Thinking at the Oxford Martin School, University of Oxford, 2020).
54. Alstott, J., Triulzi, G., Yan, B. & Luo, J. Mapping technology space by normalizing patent networks. *Scientometrics* **110**, 443–479 (2017).
55. Solt, F. & Hu, Y. dotwhisker: Dot-and-whisker plots of regression results, Comprehensive R Archive Network (CRAN) (2015).
56. Meyer, D., Zeileis, A. & Hornik, K. The strucplot framework: visualizing multi-way contingency tables with vcd. *J. Stat. Soft.* **17**, 1–48 (2007).

Methods

The raw data are the texts of over six million US patents from 1976 to 2019, from which most covariates were extracted and computed for the regression model. The conclusions crucially depend on patent classifications: patents were labelled as climate inventions (climate mitigation or adaptation technologies) if they were classified with a CPC code of Y02, and the WIPO method was applied to the patent classification codes and raw texts to classify patents as AI. I used current and multiple CPC classifications per patent (where possible). Classifications were further disaggregated with one or more labels according to what part of climate change they relate to (the CPC Y02 subclasses, described in the Supplementary Information). The analysis also included classifications of patented technologies according to their wider functions[43].

For a statistical analysis with regression modelling, the following restrictions were made by pre-processing the data. First, CPC subgroups so small that they lack patents in at least one of the years between 2010 and 2017 were removed (this is around 1% of subgroups, representing technology areas with very little patenting activity). Second, patent citation data between 2010 and 2020 was used to consider the most recent period that represents a rise in AI patenting out of which a majority emerged in the past decade. Finally, it should also be noted that the same (specified below) regression model was fit to the separate groups of technologies when I estimated and compared the predictive difference of AI between technologies. This allows the influence of the different parameters to vary across different technologies, as expected.

An investigation showed that the target count variable of forward citations is zero-inflated, heavy-tailed and related to grouping by different technologies. The Poisson model served as a natural starting point for count data, but a direct comparison of the mean and variance showed that the target variable is overdispersed. To adjust for this, I also fit negative binomial models. Here an additional size parameter controls the degree of overdispersion compared with a Poisson distribution, and a test showed it to be statistically significant. Despite this, an analysis of the residuals for model fits indicated a lack of fit as well as a grouping in the data. I therefore used a generalized inverse Gaussian distribution (a Sichel distribution), as this has been used to model highly dispersed count data[57,58] in other domains. To fit the model and adjust for grouping, the GAMLSS methodology[59] was used. Here an analysis with the randomized quantile residuals[58,60] indicated a good fit of the model to the data (for the details and model diagnostics see the Supplementary Information).

The covariates include indicator variables as controls for patents being AI inventions, whether organizations are found in the list of patent applicants and whether patents have been classified as chosen CPC group/subgroup codes. Furthermore, I controlled for grant year, the number of patent claims, the number of inventors and three variables for the number of citations to other publications: to other patents, to research literature[61] and to other literature. Finally, I included a technology cycle time[62] (TCT, the median age of cited patents) as a factor. An exploration of the data (Supplementary Information) suggests that the relationship between TCT and the target variable is nonlinear and varies between groups of technologies.

The regression model for the number of forward citations y on a 3-year horizon is specified as:

$\log \mathrm{E}[y_i] = \beta_0 + \mathrm{ai}_i \times \beta_1 + \mathrm{organizational}_i \times \beta_2 + \mathrm{claims\_log}_i \times \beta_3 + \mathrm{individual\_inventors\_log}_i \times \beta_4 + \mathrm{patent\_citations\_log}_i \times \beta_5 + \mathrm{research\_citations\_log}_i \times \beta_6 + \mathrm{other\_citations\_log}_i \times \beta_7 + \sum_j \mathrm{grantyear}_{i,j} \times \beta_j + \sum_k \mathrm{tct\_type}_{i,k} \times \beta_k + \sum_{l,m} \mathrm{tct\_type}_{i,l} \times \mathrm{grantyear}_{i,m} \times \beta_{l,m} + \sum_n \mathrm{cpc\_classification}_{i,n} \times \beta_n$ where the log transformation of the target captures the idea that the linear model describes a non-negative count variable. For further details about the covariates, see the Supplementary Information.

To compare shares of highly cited breakthroughs, I tested whether there is a difference for the climate invention areas. For this, a quantile comparison across two groups was made, adjusted for their sizes by resampling them using percentile bootstrap[63,64], to compute confidence intervals for the difference between the two groups. If one distribution has a greater quantile than the other, it indicates that inventions from this distribution more often lead to breakthroughs. The null hypothesis was taken to be H_0: $\phi_{q1} = \phi_{q2}$ for a specific quantile, where ϕ_{q1} and ϕ_{q2} are taken to represent the 99th percentile. By resampling the two distributions and examining the results under the null hypothesis[65], I estimated whether there is a difference $\phi_{q1} - \phi_{q2}$ for the 1% most cited patents (Supplementary Information).

Data availability

The datasets generated and/or analysed during the current study are available on Figshare at https://doi.org/10.6084/m9.figshare.21130173.v1 (ref. 66).

Code availability

The program scripts (in R and Python) used for the statistical data analysis are available in a GitHub/Zenodo repository under a Creative Commons Zero v.1.0 Universal license[67].

References

57. Stein, G. Z., Zucchini, W. & Juritz, J. M. Parameter estimation for the Sichel distribution and its multivariate extension. *J. Am. Stat. Assoc.* **82**, 938–944 (1987).
58. Rigby, R. A., Stasinopoulos, D. M. & Akantziliotou, C. A framework for modelling overdispersed count data, including the Poisson-shifted generalized inverse Gaussian distribution. *Comput. Stat. Data Anal.* **53**, 381–393 (2008).
59. Rigby, R. A. & Stasinopoulos, D. M. Generalized additive models for location, scale and shape. *J. R. Stat. Soc. C* **54**, 507–554 (2005).
60. Dunn, P. K. & Smyth, G. K. Randomized quantile residuals. *J. Comput. Graph. Stat.* **5**, 236–244 (1996).
61. Marx, M. & Fuegi, A. Reliance on science: worldwide front-page patent citations to scientific articles. *Strateg. Manage. J.* **41**, 1572–1594 (2020).
62. Kayal, A. A. & Waters, R. C. An empirical evaluation of the technology cycle time indicator as a measure of the pace of technological progress in superconductor technology. *IEEE Trans. Eng. Manage.* **46**, 127–131 (1999).
63. Wilcox, R. R., Erceg-Hurn, D. M., Clark, F. & Carlson, M. Comparing two independent groups via the lower and upper quantiles. *J. Stat. Comput. Simul.* **84**, 1543–1551 (2014).
64. Wilcox, R. R. *Introduction to Robust Estimation and Hypothesis Testing* (Academic Press, 2011).
65. Mair, P. & Wilcox, R. Robust statistical methods in R using the WRS2 package. *Behav. Res. Methods* **52**, 464–488 (2020).
66. Verendel, V. Tracking artificial intelligence in climate inventions with patent data: data for statistical analysis. *Figshare* https://doi.org/10.6084/m9.figshare.21749744.v1 (2022).
67. Verendel, V. Code for plots and statistical analysis of ai and climate patents. *Zenodo* https://doi.org/10.5281/zenodo.7473021 (2022).

Acknowledgements

I thank A. Pichler, K. Hötte, F. Lafond, P. Andres and E. Dugoua for valuable comments and discussions. Furthermore, I thank the seminar participants at the Institute of New Economic Thinking, Oxford, and at the Oxford Martin School. I acknowledge support from the Area of Advance Energy at Chalmers University of Technology and support from Intelligence Advanced Research Projects Activity (IARPA), via contract no. 2019-1902010003.

https://doi.org/10.1038/s41558-022-01536-w

Funding
Open access funding provided by Chalmers University of Technology.

Competing interests
The author declares no competing interests.

Additional information
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41558-022-01536-w.

Correspondence and requests for materials should be addressed to Vilhelm Verendel.

Peer review information *Nature Climate Change* thanks Nick Johnstone, Benedict Probst and David Rolnick for their contribution to the peer review of this work.