Analyzing dynamics of innovation networks in climate technologies

Malte Toetzke[a,c,b,*], Benedict S. Probst[c,b,f], Stefan Feuerriegel[d,e], Laura Diaz Anadón[f,g], Volker H. Hoffmann[b]

[a]*TU Munich, Munich, Germany*
[b]*ETH Zurich, Zurich, Switzerland*
[c]*Max Planck Institute for Innovation and Competition, Munich, Germany*
[d]*LMU Munich, Munich, Germany*
[e]*Munich Center for Machine Learning, Munich, Germany*
[f]*University of Cambridge, Cambridge, United Kingdom*
[g]*Harvard University, Cambridge, United States*

Abstract

Innovation networks play a crucial role in the development, commercialization, and adoption of climate technologies, yet their structure and dynamics remain poorly understood. To address this gap, we use large language models to analyze 26 million LinkedIn posts, thereby mapping a global network of 166,459 organizations and 442,250 collaborations ($\sim$1.4 million direct partnerships) across 189 countries. The dataset spans 27 climate technologies and diverse types of collaborations, such as demonstration projects, product launches, and equity investments. Our findings highlight the growing centrality of governmental organizations in these networks, particularly in emerging and scaling technologies that require substantial upfront investment, such as direct air capture, geothermal energy, and green hydrogen. Countries with the largest increases in governmental engagement exhibit the strongest growth in domestic partnerships, underscoring the role of public-sector actors as catalysts for industrial ecosystem formation. We show that 62% of all partnerships directly or indirectly depend on government

*Corresponding author.
Email address: malte.toetzke@tum.de (Malte Toetzke)

involvement, which also raises concerns about the resilience of climate-tech innovation networks under shifting political and fiscal priorities. Contrary to the US, the EU exhibits weaker geographical interconnectivity and regional specialization, potentially limiting industrial competitiveness. Our machine learning approach provides a scalable, real-time methodology for tracking innovation networks over time, with broad applications for industrial analysis and policy evaluation beyond climate-tech.

Keywords:
Machine learning, large language models, monitoring, innovation networks,
climate technologies, industrial policy

---

1. Main

Achieving net-zero emissions hinges on the rapid advancement of climate technologies (climate-tech), from development to commercialization and large-scale adoption. Although emerging industries in climate-tech hold considerable economic promise [1, 2], many technologies remain in early stages between development and commercialization, or lack the scale necessary for widespread market adoption [3]. In response, countries worldwide have implemented ambitious green industrial policies (GIPs) to spur climate-tech innovation and bolster domestic leadership in emerging industries, such as the US Inflation Reduction Act (IRA) [4] and the EU Net Zero Industry Act[5].

Climate technologies are developed, commercialized, and scaled through dense innovation networks that connect diverse actors, such as universities, startups, incumbent firms, financial institutions, and governmental organizations [6, 7, 8]. These networks provide access to complementary resources and capabilities [9, 10, 11] and enable knowledge spillover, which is crucial for technological refinement and scale-up [12, 13, 14, 15]. Despite the importance of the middle stages in the innovation process spanning demonstration, commercialization, and early adoption, existing innovation data disproportionately capture early-stage invention (e.g.,, via patents) [16, 17, 18, 19] and late diffusion stages (e.g., via trade and deployment indicators) [20, 21, 22, 23]. As a result, they fail to sufficiently track the critical transitions from research and development to demonstration, commercialization, and early adoption, which are particularly challenging for climate technologies and lead to the failure of many ventures due to a lack of financial capital, necessary infrastructure, demand, or profitability [24, 25, 26, 27].

Here, we address this gap by presenting a novel approach for capturing the dynamics of climate-tech innovation networks along the entire innovation process. We use large language models (LLMs) to map innovation networks in climate-tech by analyzing LinkedIn [28] announcements from key public and private organizations. Our methodology constructs a highly

comprehensive network of climate-tech collaborations in two main steps (Figure 1a): First, we classify the LinkedIn profiles of organizations into 18 organization types (e.g., universities, manufacturers, venture capital firms; Supplementary Table S4) and 46 sectors (e.g., automotive, chemicals, software; Supplementary Table S3). Second, we apply LLMs to (1) identify announcements of climate-tech innovation collaborations from the LinkedIn posts of organizations, (2) classify these announcements into 17 collaboration types (e.g., research and development, demonstrations and pilots, product launches, equity investments; Supplementary Table S8) and 27 climate technologies (e.g., offshore wind, electric vehicles, direct air capture; Supplementary Table S2), and (3) extract collaborating organizations from the posts. We ran these steps over three iterations, starting with an initial set of organizations from BloomberNEF [29] and expanding the network by adding new organizations that were identified through collaborations in the previous iterations (details are provided in the Methods). Our approach is fully replicable, enabling systematic, large-scale retrieval of innovation collaborations as well as continuous updates with new announcements.

Based on an input of 26 million LinkedIn posts, our final dataset comprises 442,250 innovation collaborations between 166,459 organizations, representing 1.4 million direct partnerships. As multi-actor collaborations connect more than two organizations, we transform all collaborations into direct links between organizations, hereafter called "partnerships" (Supplementary Figure S2; details are described under *Network graph* in the Methods). Collectively, the included organizations operate from 345,479 locations across 189 countries (Figure 1b) and employ around 136 million people (Supplementary Figure S3). The dataset was collected until January 2025. It spans collaborations from January 2017 to December 2024. Likely, organizations and collaborations from China, South Korea, and Japan are underrepresented due to lower LinkedIn activity in these countries (we discuss cross-country coverage in the Supplementary Discussion S1). Finally, our data allows us to track changes in climate-tech innovation net-

works over time and across wide range of technologies, geographies, and innovation activities, shedding light into the critical transition from lab-scale research to large scale diffusion.

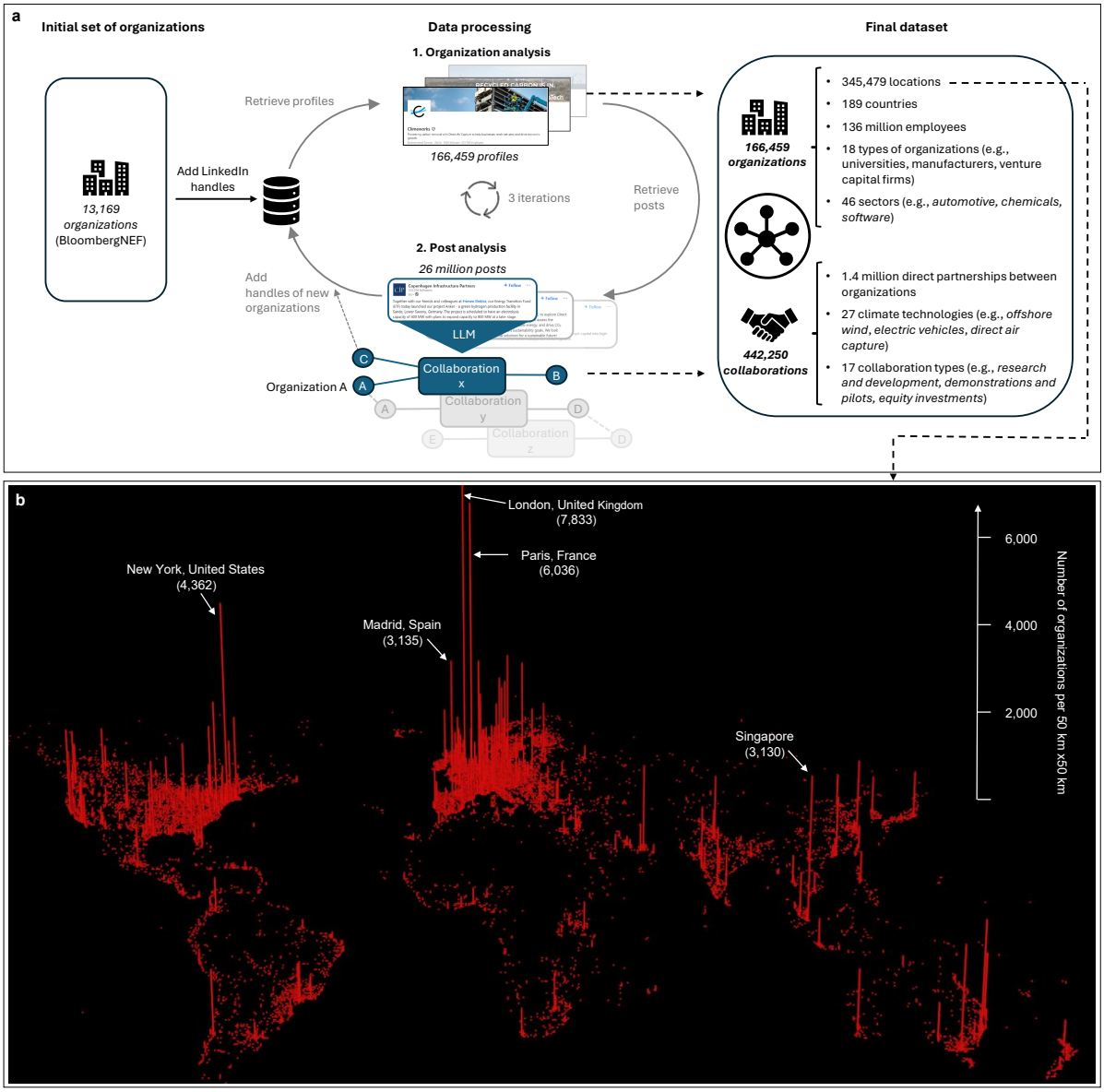


Figure 1: **Overview of information retrieval process and final dataset capturing innovation collaborations in climate-tech. a,** Overview of information retrieval process and final dataset spanning collaborations from 01.2017–12.2024. Note that collaborations can include multiple organizations, which lead to a higher number in direct partnerships (see Supplementary Figure S2). **b,** Geographic distribution of organizations engaging in climate-tech collaborations as of January 2025, based on 50 km x 50 km grids.

2. Evolution of global innovation networks in climate-tech

Structures and activities in climate-tech innovation networks are evolving dynamically. To compare innovation networks over time, we use the enactment of the Inflation Reduction Act (IRA) in July 2022 to distinguish between two periods of equal length: the Pre-IRA (January 2020–June 2022) and the Post-IRA period (July 2022–December 2024). Between the periods, our data show a substantial expansion of innovation networks across technologies, with an overall increase of 93% in newly formed partnerships and 49% in the number of participating organizations. These growth rates correlate with the development stages of technologies, and are highest for technologies around commercialization stages but seem to decrease along market uptake and large-scale diffusion.

In the following, we discuss the structure and dynamics of innovation networks for three exemplary climate-techs representing different development stages, namely, (i) nuclear fusion, (ii) direct air capture, and (iii) green hydrogen (Fig 2).

Nuclear fusion is a nascent technology in the early stages of technological development (i.e., technology readiness level (TRL): 1–3 [30]). Here, we observe how progress in demonstrations and pilots has catalyzed a substantial expansion in research and development (R&D) partnerships, while stimulating first commercialization efforts. The innovation network in the first period (01.2020–06.2022) was primarily shaped by the construction of large demonstration reactors (Fig. 2a), such as ITER in France [31] and SPARC in the United Kingdom [32]. Consequently, partnerships during this period are dominated by *demonstrations and pilots* (n=623 partnerships), with the international research organization ITER emerging as a central actor, linking 89 organizations from manufacturing (primarily manufacturing of conductor materials, such as ASG Superconductors and Elytt Energy), project management (e.g., Assystem), and governments (e.g., the Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA), United Kingdom Atomic Energy Authority). In the subse-

quent period (07.2022–12.2024), the number of partnerships around demonstration projects has remained constant (likely due to the long time horizon of the demonstration projects). However, we observe two major shifts: first, we observe a surge in R&D partnerships, which have increased by 491% (n=2,222), likely driven by rising interest in fusion technologies following several research breakthroughs at existing demonstration plants (e.g., the National Ignition Facility (US), Joint European Torus (UK), Wendelstein 7-X (Germany)). Second, first partnerships emerged with a focus on commercialization (n=98). This trend is accompanied by the entry of venture capital firms and the first recorded offtake agreements (n=7), committed to purchasing future electricity supply from Helion's planned first-of-its-kind fusion power plant (expected to launch in 2028). These trends suggest that technological milestones in nascent technologies stimulate intensified R&D activity while also facilitating the transition to commercialization, a phenomenon that underscores the non-linear nature of technological innovation.

Direct air capture (DAC) is an emerging technology (alongside other climate-techs, such as sustainable aviation fuels and marine energy; TRLs: 4–7), entering the market with first-of-a-kind commercial deployments. Here, we observe the formation of initial industrial structures. In the first period (01.2020–06.2022), the DAC innovation network was highly concentrated around two core technology manufacturers: Climeworks (degree centrality=86) and Carbon Engineering (45; Fig 2b). These firms were primarily supported through finance from venture capital firms, such as Breakthrough Energy (28), and governmental organizations, such as the US Department of Energy (DOE) (28). In the second period (07.2022–12.2024), the network has expanded substantially with the emergence of additional technology manufacturers as central actors (e.g., Heirloom (60), Capture6 (56)). At the same time, DAC manufacturers have increasingly started to form diverse partnerships with other types of organizations, including early adopters, such as Stripe, Milkywire, Shopify, and complementary service providers, such as PuroEarth and Carbon Direct, which facilitate the monetization of DAC through carbon credit

markets. In addition, the US DOE has positioned itself at the very center of the innovation network, by supporting the formation of direct air capture hubs through a USD 3.5 billion bill [33]. As a result, partnerships around first commercial deployments increased by 1,500% between the two time periods. The evolution in the DAC network highlights the critical role of early movers (e.g., Climeworks, Breakthrough Energy, US DOE) in developing and financing core technologies towards commercialization, which facilitates market entry by service providers, adopters as well as competitors.

Green hydrogen is a scaling technology (alongside other climate-techs, such as batteries, heat pumps, and biofuels; TRLs 8–9) that is experiencing increasing uptake in manufacturing and deployment, accompanied by reductions in cost. As a result, its innovation network has expanded substantially over the two periods, with the number of participating organizations and partnerships doubling (Fig. 2c). While partnerships have expanded across all innovation stages – from R&D to operations and maintenance – the strongest relative increase is observed in offtake agreements, reflecting a growing focus of industrial companies on securing future supply in green hydrogen markets. Unlike nascent and emerging technologies, where new types of organizations gain centrality, the green hydrogen innovation network remains structurally comparable at the level of organization types, with shifts occurring primarily at the level of individual organizations. This suggests that by the first period (01.2020–06.2022), the sector had already developed a well-established industrial structure, which was subsequently scaled through a growing number of projects in the second period (07.2022–12.2024). Incumbent manufacturers from adjacent industries, such as Siemens Energy, Air Liquide, and Shell, play a central role in this network. Their increasing engagement shows how established industries intensify investments to integrate hydrogen technologies into their business, as the commercial viability of specific applications becomes clearer. The widespread involvement of incumbents from different industries (e.g., Airbus, Shell, Thyssenkrupp) also underscores the complemen-

tary nature of green hydrogen, with potential applications spanning diverse use cases, including industrial decarbonization, seasonal energy storage, and aviation. More broadly, these findings highlight that, in scaling technologies, the expansion of innovation networks is driven less by the emergence of new organization types and more by the increasing engagement of established industrial players seeking to integrate the technology into their core business.

Overall, we observe a divergence in the geographic proximity of organizations in innovation networks across different technology stages. Partnerships of nascent and emerging technologies have become increasingly domestic over the two periods, while scaling and mature technologies have expanded internationally. An exception to this trend is biochar and bioenergy with carbon capture and storage (BECCS), which have become more international despite their early development stages (likely this deviation in pattern arises from the separation of physical processes and digital value proposition, where carbon credits that were generated in a specific region can be traded anywhere in the world as a digital commodity . The increased nationalization of innovation networks in nascent and emerging technologies reflects a transition from highly international research collaborations to more localized partnerships supporting first-of-a-kind deployments and early adoption, where close geographical proximity fosters trust [34] and knowledge exchange [35, 13]. In contrast, innovation networks in scaling and mature technologies become more international as organizations expand into new geographic markets, thereby increasing the global diffusion of physical technologies (and components), operations, and knowledge [36, 13].

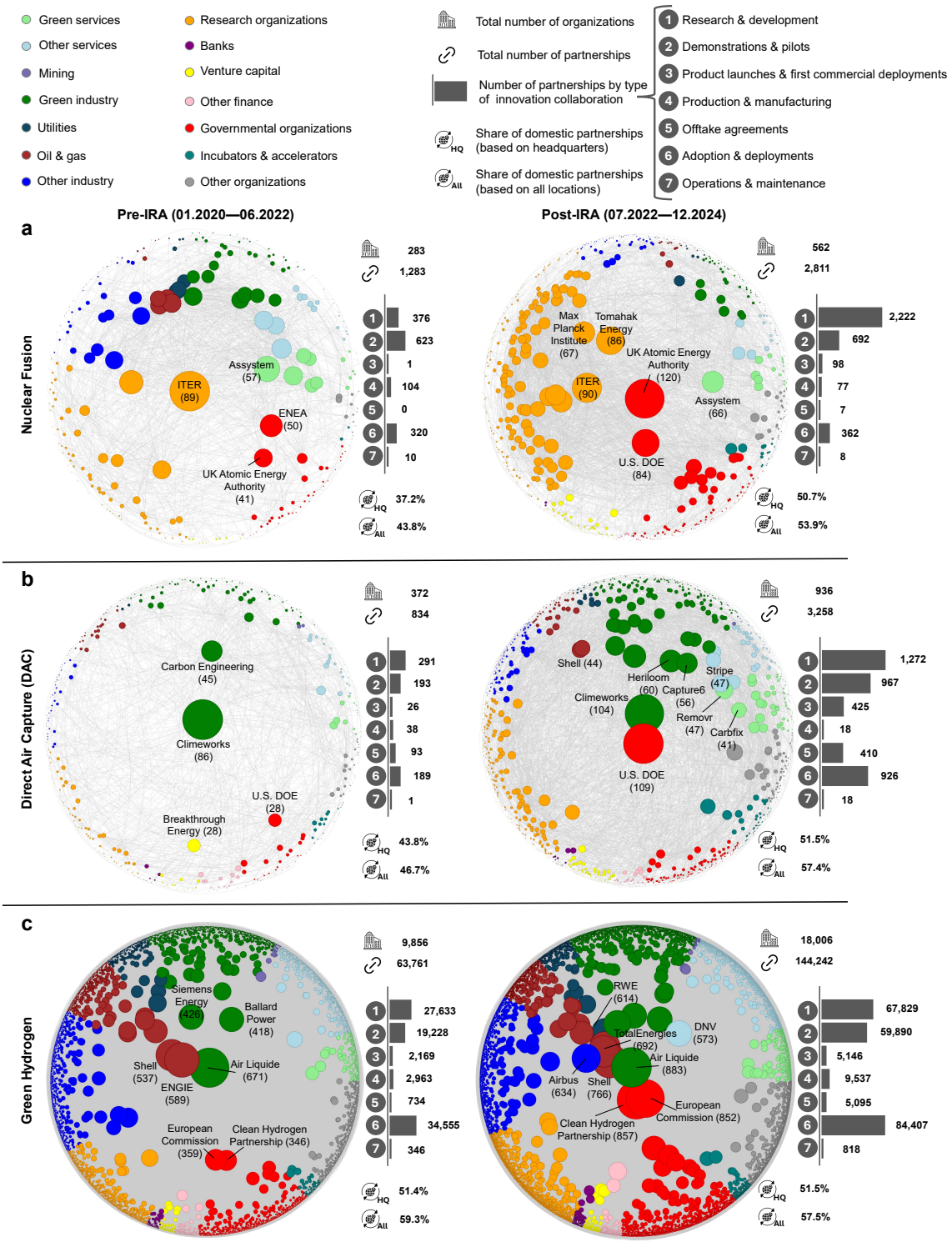


Figure 2: **Changes in the structure of global innovation networks.** Shown are two periods before (Pre-IRA: 01.2020–06.2022) and after (Post-IRA: 07.2022–12.2024) the introduction of the inflation reduction act (IRA) for nuclear fusion (a), direct air capture (DAC) (b), and nuclear fusion (c). As innovation collaborations can involve multiple organizations, we extract the direct edges between individual organizations from collaborations (called "partnerships"). To prevent double counting from multiple announcements, partnerships between two organizations are only counted once per period. The size (radius) of organization nodes represents the degree centrality (i.e., the number of partnerships), with centrality values indicated in brackets for highlighted organizations. Node colors correspond to different groups of organizations. To reduce the number of groups, they represent aggregates of underlying organization types (e.g., universities and research institutes are aggregated to *research organizations*, manufacturers and project developers specialized in climate technologies are aggregated to *green industry*). The horizontal bars illustrate the distribution of partnerships across different types of innovation collaborations. For simplicity, financial collaborations (e.g., equity investments, loans, grants) and other structural collaborations (e.g., mergers and acquisitions, joint ventures, spin-offs) are not shown in the bars. In the network plots, we only illustrate partnerships as a binary indicator (existing/not existing), while in the bars we count partnerships separately for each type of collaboration.

3. Governmental organizations moving towards the center of climate-tech innovation networks

Between the two analyzed periods (January 2020–June 2022 and July 2022–December 2024), governmental organizations have gained substantial centrality in climate-tech innovation networks. Key governmental institutions, such as the US Department of Energy (DOE), the European Commission, and France's Alternative Energies and Atomic Energy Commission (CEA), have positioned themselves at the very center of global innovation networks across many climate technologies (Figure 2). This trend becomes particularly evident for emerging and scaling technologies that require substantial upfront investment for first-of-a-kind deployments and infrastructure build-up (e.g., DAC, geothermal energy, green hydrogen). In contrast, for technologies that integrate into existing technologies, infrastructure, and business models (e.g., sustainable aviation fuels, carbon capture and storage, biofuels) incumbent manufacturers, utilities, and oil and gas majors are more likely to take up central positions. These patterns suggest that key governmental organizations step in as central ecosystem builders where incumbent industries remain hesitant and where investment risks are high.

Our data indicate that increased governmental involvement is associated with the expansion of domestic industrial ecosystems. Across the 27 climate technologies analyzed, an overall increase in centrality of governmental organizations between the two periods consistently aligns with greater centrality of domestic actors. In 19 of these 27 technologies, the country that experienced the largest increase in governmental partnerships also recorded the highest overall growth in domestic partnerships (i.e. collaborations occurring between or with organizations from the respective country). In these cases, a one-degree increase in governmental centrality is associated with a multiple of new domestic partnerships, where factors range between 3.2 (nuclear fission, United States), to 5.8 (hydrogen, France), 7.2 (offshore wind, United Kingdom), 8.5 (battery, United States), and 10.3 (biochar, United States). For instance, in the global

DAC network, the United States (US) DOE positioned itself as the most central organization (Figure 2c). This structural shift has been accompanied by a 427% increase in the centrality of US-based organizations in the global network, culminating in a global innovation network where 31% of all DAC-related partnerships now take place between US entities, and 56% involve at least one US-based actor. These findings reinforce the notion that governmental organizations not only foster technological progress but also strategically enable domestic industrial leadership in emerging climate technologies.

The high centrality of governmental organizations in climate-tech innovation networks also highlights the dependency on public-sector involvement. Between July 2022 and December 2024, 62% of all partnerships depended on governmental organizations, with 28% of partnerships directly involving governmental organizations, while an additional 34% of partnerships arise from multi-actor collaborations that are supported or initiated by governmental organizations (Fig. 3). In the US innovation network, the Department of Energy (DOE) accounts for 10.3% of all climate-tech partnerships depending on the DOE, either directly (3.8%) or indirectly (6.5%). In comparison, in the EU, only 3.9% of climate-tech partnerships depend on the European Commission. High dependence on government involvement does not only appear in emerging and scaling technologies where governmental organizations are positioned at the center of the innovation network, but also in more mature technologies, where governments act as providers of public infrastructure and regulators (e.g., railway, electric nuclear fission, hydro energy).

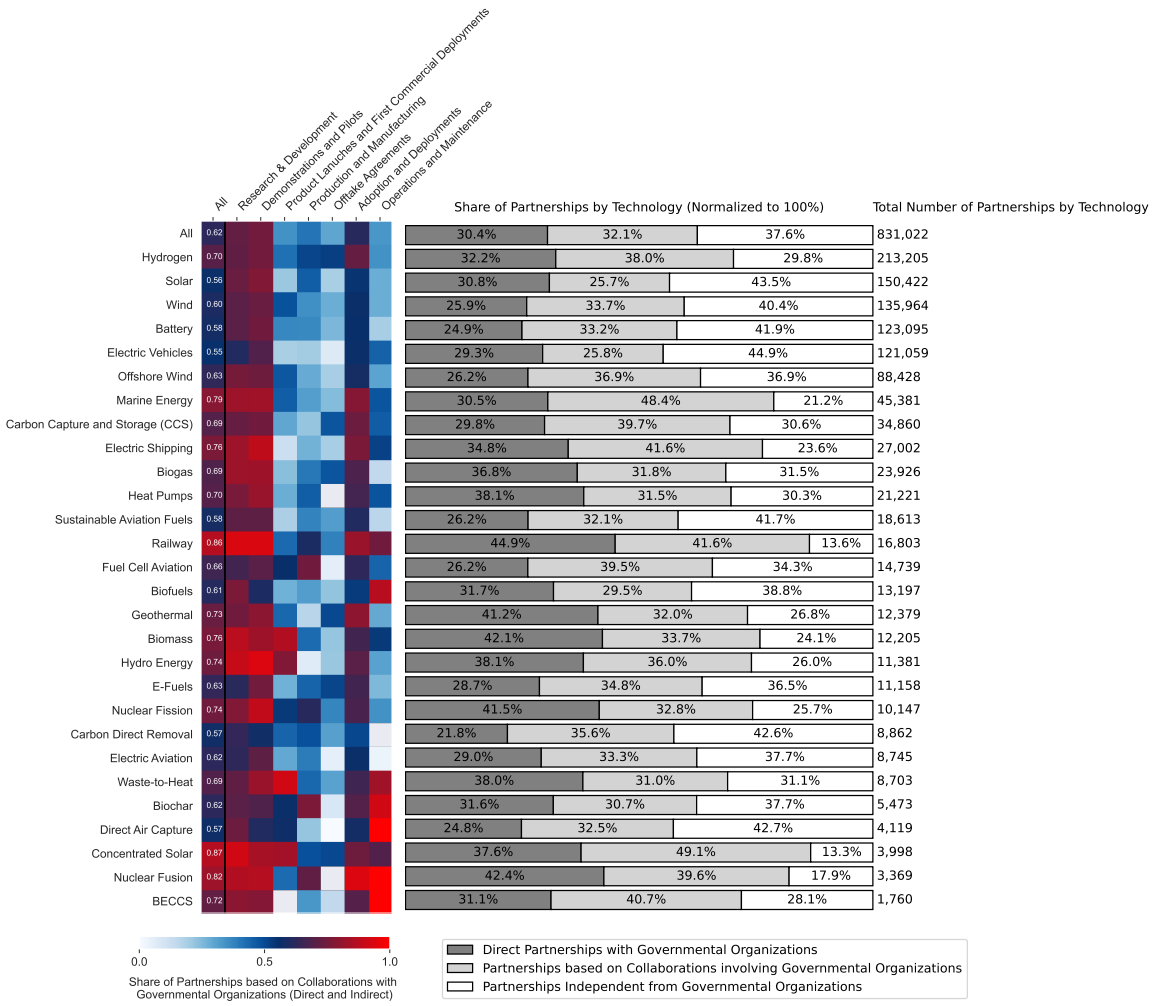


Figure 3: **Indirect and direct involvement of governmental organizations in innovation networks by climate technology and from July 2022 to December 2024**. Direct involvement refers to partnerships that include governmental organizations, whilst indirect involvement refers to partnerships originate from multi-stakeholder collaborations including governmental organization.

4. Comparing innovation networks between the EU and the US

A comparative analysis between the United States (US) and the European Union (EU) demonstrates substantial structural differences and divergent pathways in government-led ecosystem building. Our data show that climate-tech innovation networks in the EU are more geographically clustered than in the US. While, in the US, 71.3% of innovation partnerships link organizations from different states (based on headquarter locations of organizations), only 39.3% of partnerships in the EU involve different European countries. Relatedly, we observe low levels of geographic specialization, when comparing the number of manufacturers in different climate-techs across EU member countries. The most notable deviations can be found in France, with a comparatively high concentration of manufacturers in nuclear fission, and in Denmark, with a comparatively high concentration of manufacturers in wind energy (Supplementary Figure S4). In contrast, US states show large variations in technological specialization including, for example, high manufacturing concentration in batteries and electric vehicles in Michigan, carbon capture and storage (CCS) in Texas, biomass and hydropower in Pennsylvania, and marine energy in Massachusetts (Supplementary Figure S5). Compared to the EU, the US innovation network appears to build more diverse regional capabilities and exploit these capabilities through cross-state collaborations.

We also observe a higher fragmentation in government-led interventions in the EU compared to the US (Figure 4). US-based innovation networks are highly centralized, with the Department of Energy (DOE) acting as the dominant orchestrator of innovation activities across climate technologies. In contrast, in European innovation networks, a variety of public institutions at national and European levels occupy central positions, including organizations, such as the European Commission. The French Agency for Ecological Transition (ADEME), the French Alternative Energies and Atomic Energy Commission (CEA), and the German Federal Ministry for Economics and Climate Action (BMWK). This fragmentation of governmental

participation in the EU's climate-tech innovation networks could be a critical factor driving geographical clustering within national borders and lower levels of technological specialization between countries.

Finally, our data show distinct patterns of how governmental organizations interact with private actors in climate-tech innovation networks. In the US, the direct involvement of governmental organizations is higher, with 37% of partnerships directly involving governmental organizations compared to 28% in the EU. However, in the EU, the indirect dependence on governmental support is considerably higher, with 34% of partnerships between non-government actors relying on government support (e.g., through funding or public procurement) compared to only 18% in the US. This difference likely originates from a multi-stakeholder approach practiced by EU institutions. For example, the main orchestrator in the EU – i.e., the European Commission – generally supports larger consortia that average 5.65 actors, while projects supported by the US DOE average only 2.66 actors. DRAW CONCLUSION (e.g., EU institution as coordinators of innovation fostering collaboration among other institutions, US more as direct partner financing innovation within institutions)

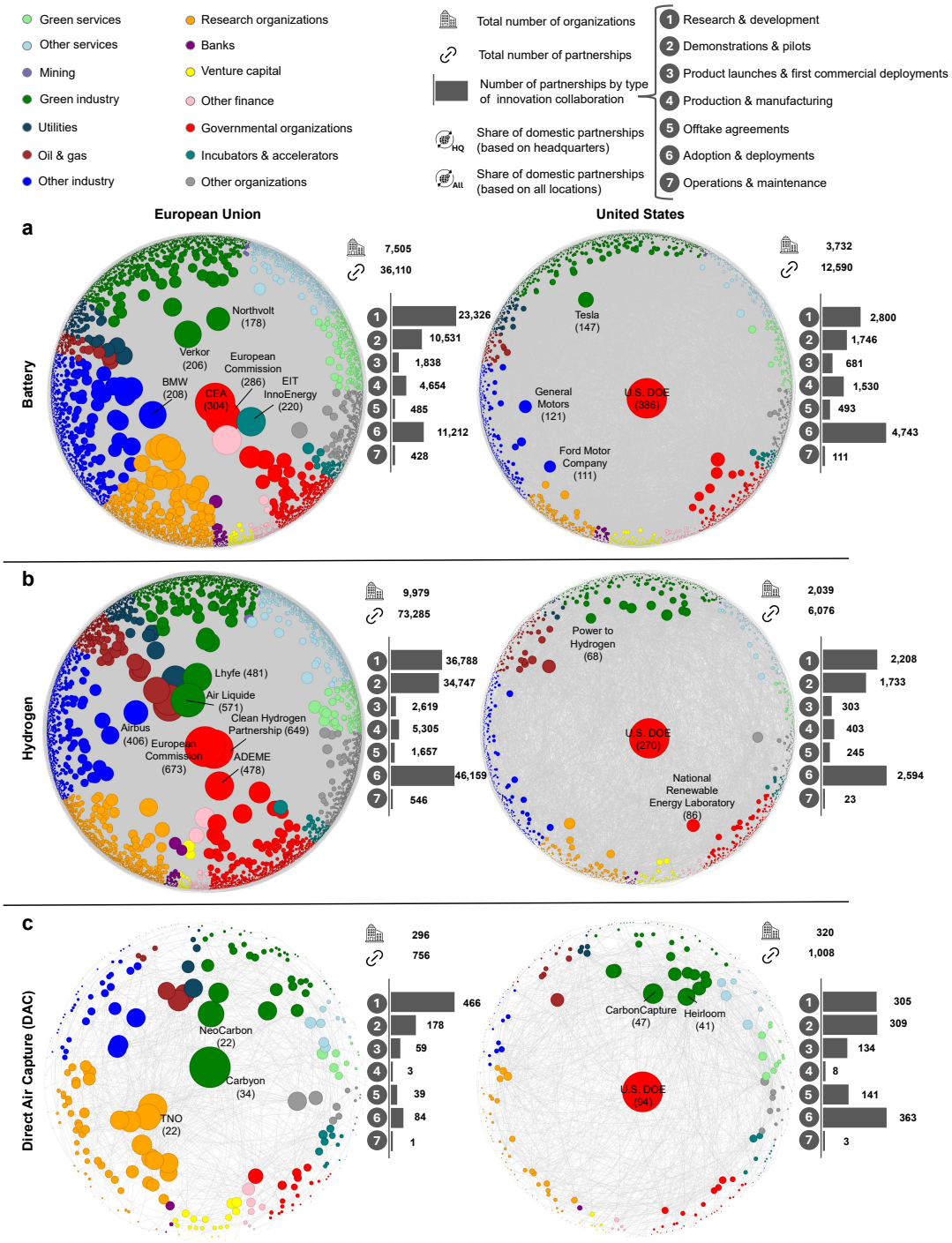


Figure 4: **Comparison of innovation networks between the US and the EU for the period after the inception of the inflation reduction act (Post-IRA: 07.2022–12.2024) and for three exemplary technologies: batteries (a), green hydrogen (b), and direct air capture (DAC) (c).** The size (radius) of organization nodes represents the degree centrality (number of edges), which is also indicated in brackets for highlighted organizations. Organizations are attributed to different groups of organizations (e.g., "Green services", "Research organizations"), represented by the colors. Some of these groups represent aggregates of underlying organization types (e.g., universities and research institutes are aggregated to research organizations). Horizontal bars represent the number of links that can be attributed to different types of innovation collaborations. For simplicity, financial collaborations (e.g., equity investments, loans, grants) and other types of collaborations (e.g., mergers and acquisitions, joint ventures, spin-offs) are not shown in the bars. To be included in the domestic network, each partnership needs to involve at least one organization with a headquarter location in the region, whereas the other organization only needs to have any type of location in the region.

5. Regional innovation clusters in green hydrogen

Here, we analyze regional innovation networks for green hydrogen, featuring dynamically evolving industry [37]. Establishing regional green hydrogen innovation networks, often referred to as "hydrogen hubs" (H2Hubs) or "hydrogen valleys", has become a central pillar in national green industrial strategies. For instance, Germany's National Hydrogen Strategy, launched in 2020, allocated €7 billion to scaling up domestic hydrogen partnerships [38], while the U.S. National Clean Hydrogen Strategy and Roadmap, introduced in 2022, dedicated $8 billion from the Infrastructure Investment and Jobs Act to establish regional H2Hubs focusing on production, storage, and end-use applications [39]. By fostering synergies among diverse actors that are located in close proximity, H2Hubs are expected to accelerate the development of domestic hydrogen industries by creating dense networks that facilitate the co-creation of hydrogen projects.

To explore the dynamics of emerging H2Hubs, we focus on the hydrogen innovation clusters in Berlin and Houston, which, based on the number of local hydrogen organizations, are the largest hydrogen hubs in the EU and the US, respectively. We define the clusters by filtering all hydrogen-related collaborations formed between organizations located within a 50 km radius of each city center.

Despite similar size (in terms of the number of organizations), the two networks show distinct structures and activities, which reflect their specific regional characteristics. In Berlin, the hydrogen innovation network is primarily centered around governmental organizations (Figure 5a). As such, Berlin as Germany's capital benefits from its proximity to federal institutions (e.g., the BMWI), accompanied by a strong presence of research institutes (e.g., the Fraunhofer Institutes). This setup seems to have fostered an innovation network primarily focusing on technology push, indicated by comparatively large shares in R&D collaborations and demonstration projects, where financial support is mainly provided through grants (Figure 5c).

In contrast, Houston's hydrogen cluster is characterized by more market-driven features. The innovation network is primarily centered around oil and gas firms (Figure 5b) and shows a higher share of partnerships focusing on commercialization, offtake agreements, and deployments as compared to Berlin. As such, the innovation network seems to capitalize on the city's extensive industrial ecosystem. Houston's innovation network also features a substantially higher share in equity-based financial partnerships compared to Berlin. While the U.S. Department of Energy (DOE), which oversees the federal support for regional H2Hubs, is not located in Houston, grant-based financial partnerships within Houston's innovation network still remain substantial.

Comparing the periods 01.2020–06.2022 and 07.2022–12.2024, both innovation networks have grown in terms of the number of organizations and partnerships (Figure 5). Interestingly, in Berlin, the share of regional partnerships of hydrogen firms (i.e., partnerships with other organizations located in Berlin) has increased (from 19.3% to 21.9%), while, in Houston, it has decreased (from 16.4% to 14.2%). As such, Berlin's hydrogen innovation network has become denser.

The key challenge for Berlin's hydrogen network is the gradual transition from public sector-driven innovation to private sector-led commercialization and adoption. Already between the two periods, we see that the innovation network has transitioned from government-centric to a more diversified structure that now includes manufacturers, utilities, and research utilities as central actors. As the technological maturity of electrolyzers increases, both reducing dependency on government funding and attracting private investment will be critical to sustaining the growth of the cluster. In Houston, on the other hand, the main question is whether oil and gas firms will continue to invest in the development and adoption of green hydrogen, especially if political incentives shift under the Trump administration. Decreasing demand may steer the relocation of international manufacturers, such as Topsoe, potentially affecting the city's status

as a leading hydrogen hub.

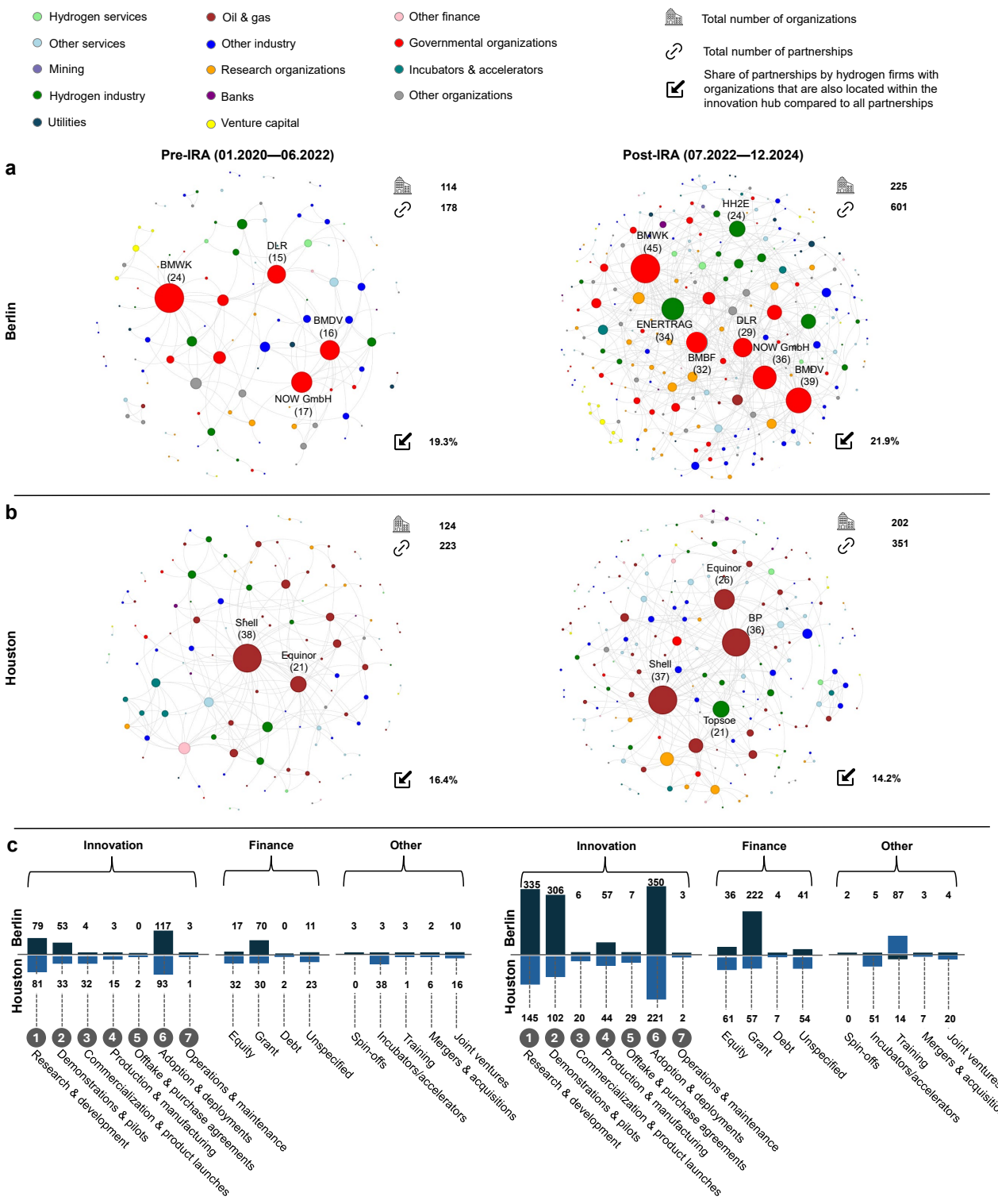


Figure 5: **Comparison of the hydrogen innovation clusters in Berlin and Houston for the periods 01.2020–06.2022 and 07.2022–12.2024.** For comparability, we included all organizations within a 50 km radius around the city centers. **a,b,** Comparison of the network graphs of collaborating organizations. **c,** Comparison of the share of innovation activities across different types of collaborations.

6. Discussion

Through our machine learning approach, we have compiled a global dataset comprising 166,459 organizations linked through 442,250 collaborations ($\tilde{1}$.4 million direct partnerships) that contribute to climate-tech innovation. To the best of our knowledge, our dataset enables the first large-scale analysis of innovation networks across different climate technologies, geographies, organizational types, and stages of the innovation process over time. The breadth and granularity of our dataset allow for novel insights into the structural evolution of climate-tech innovation and the role of key actors.

Our findings highlight the crucial role of governmental organizations in supporting domestic industries in climate-tech. In 19 of the 27 analyzed climate technologies, the largest increases in domestic partnerships occurred in countries where centrality of governmental organizations increased the most. Especially in emerging and scaling technologies that require substantial upfront investment for first-of-a-kind deployments and infrastructure build-up (e.g., direct air capture, geothermal energy, green hydrogen), key governmental organizations, such as the European Commission and the US Department of Energy, now represent central actors in global innovation networks and thereby foster critical partnerships between research, industry, investors, and adopters. Underpinning activities reflect the dual role of governmental organizations: both as financiers of innovation activities (e.g., through funding and public procurement) and as active participants, often represented by governmentally owned organizations, such as governmental utilities or accelerator programs. By shaping innovation networks, governmental organizations act as catalysts of industry and market formation, accelerating the development and scaling of climate technologies.

Despite the benefits of governmental involvement, the high degree of network dependency on governmental organizations raises concerns about the resilience of climate-tech innovation systems in the face of shifting political and fiscal priorities. Our data show that 62% of all part-

nerships either directly involve governmental actors or indirectly depend on their support. In the US, the centrality of the Department of Energy (DOE) stands out, with 10% of all climate-tech partnerships in the US depending on the DOE, either directly (3.8%) or indirectly (6.5%). Beyond individual projects, precipitate reductions of governmental support could trigger systemic disruptions in climate-tech innovation networks. In this context, the decline of the German solar industry following the phase-out of the Renewable Energy Sources Act (EEG) in 2013 serves as a cautionary tale [13]. To ensure the long-term stability of innovation networks, policy makers should develop long-term strategies for the gradual transition from active ecosystem builders in early-stage industries (e.g., forming partnerships through public procurement and auctions), to facilitators during market adoption (e.g., de-risking investments, incentivization adoption), and ultimately to a regulatory function in mature markets (e.g., ensuring competition).

Comparing innovation networks in the EU and the US, we observe differences in the degree of cross-regional connectivity and industrial specialization. In the U.S., 71.3% of innovation partnerships span multiple states, whereas, in the EU, only 39.3% of partnerships cross national borders. Furthermore, EU member states exhibit lower degrees of regional specialization in climate-tech manufacturing, as measured by the distribution of manufacturers across different climate technologies. Regional specialization has been shown to facilitate industrial development by enabling technological spillovers, shared infrastructure, and the development of specialized labor markets [40]. If well-connected across regions, comparative advantages through specialization can generate trade benefits [41] and lead to improvements in productivity, employment, and regional economic development [42]. The lower levels of connectivity and specialization observed in EU innovation networks may, in part, be attributed to the fragmentation of governmental interventions. While the US DOE functions as a central coordinating entity at the federal level, we observe a variety of national and EU-level organizations at the center of EU innovation networks. Greater coordination at the EU level, targeting cross-border

partnerships while fostering regional specialization, could enable a more effective innovation network and enhance industrial competitiveness.

Our machine learning approach for tracking innovation networks provides substantial advantages compared to existing approaches, in terms of scope, timeliness, and costs. To the best of our knowledge, we created the first dataset that spans various types of organizations, technologies, and collaborations in innovation. In contrast, existing datasets used in innovation research are typically either based on small samples (e.g., [11, 18]) or confined to distinct aspects of innovation. For example, the Crunchbase database [43] primarily focuses investment and funding of private firms. Data by BloombergNEF [29], on the other hand, primarily focus on large-scale deployment finance for mature climate technologies, such as wind and solar, and provide only aggregated estimates on emerging technologies and early-stage innovation activities. Other secondary data sources used in innovation research, such as patent [18, 44] or trade data [21, 22], are also limited to specific stages of the innovation process. In addition, our machine learning approach allows data collection in near-real time and with continuous updates, which thereby avoids extensive time lags between innovation activity and data collection of existing datasets. Further, automating the information retrieval process via machine learning enables considerable reductions in work and costs.

Our approach for monitoring innovation is nevertheless subject to limitations. First, our coverage of organizations may vary across countries and is likely lower in Japan, Korea, and China (we discuss data representation in Supplementary Discussion S1). Still, Linkedin is the largest existing professional social media platform worldwide with high penetration rates in the majority of countries, which also includes various lower- and middle-income countries [45, 46]. Future research could extend applications of our machine learning approach to other, country-specific professional social media platforms. Second, despite high incentives for organizations to self-report collaborations and partnerships on LinkedIn (e.g., for attracting job applicants,

investors, partners, and customers [47, 48, 45]), the degree of information revealed may vary across different types of organizations, based on the sensitivity of information and the general social media engagement of organizations. Third, our data may be biased to organizations that were still alive during the start of our data collection. As such, we did not capture organizations that exited the market and deleted their LinkedIn profile before the first data collection. Frequent updates in our data retrieval will allow us to eliminate this bias in the future and also help us identify market exits and failures of organizations. Fourth, for some technologies, our categorization aggregates different sub-technologies or related infrastructure, such as, for example, charging infrastructure for electric vehicles or different processes for producing hydrogen. Future extensions could offer more detailed categorizations of technologies and collaborations and extract deployed capacities and financial numbers from social media announcements.

In the future, our machine learning approach can be used to continuously track the progress and evolution of climate-tech innovation networks across various technologies, types of organizations, stages of innovation, and regions. The resulting data can be used to complement other, existing indicators for tracking innovation, such as patent data for inventions or trade data for late-stage diffusion. Especially for the middle stages of the innovation process, our machine learning approach can fill existing evidence gaps with real-time data and, therefore, show how technologies evolve from research applications to demonstrations and commercialization. Frequent updates can inform targeted policies and help evaluate the effectiveness of policy interventions over time. This can, for example, help to assess whether policies stimulated increased commercialization activities or produced adverse effects, such as the relocation of firms. A better understanding of innovation dynamics in climate-tech will be fundamental to accelerate the decarbonization of global economies.

Finally, our machine learning approach can be transferred to other domains and help advance the state of technology and innovation in other industries, such as pharmaceuticals, chemicals,

and information and communication technologies.

Methods

6.1. Data

Our global innovation network of climate technologies is based on publicly available information retrieved from the social media platform LinkedIn. LinkedIn is the largest professional social media platform with over 900 million global users worldwide [45, 46]. All data we use is self-reported by organizations through their profile information and posts. As organizations use LinkedIn to attract and retain important stakeholders, such as employees, business partners, and investors, the provided information is generally considered accurate and comprehensive [47, 48, 45].

We retrieved the data through an iterative process, performing three iterations. We started with an initial set of 13,235 organizations provided by the climate technology organizations database from BloombergNEF [29], which has a global coverage of climate technology organizations. The database includes the organizations' website addresses from which we retrieved corresponding LinkedIn accounts as initial input to our machine learning approach. In subsequent iterations, the LinkedIn accounts of organizations were then identified by analyzing the involvement of new organizations in climate-tech collaborations (see Section "Iterative process").

To retrieve data on organizations, we directly accessed the profiles of organizations through the web. In particular, we retrieved two types of data: (1) metadata of organizations and (2) post data, as follows.

(1) Metadata include (i) the locations of organizations, (ii) account data that is automatically compiled by LinkedIn, and (iii) organizational attributes that are self-reported by the organization. Locations are provided on address level (including country, city, postal code, street, and street number). We converted all addresses into geolocations in the form of longitude and latitude coordinates using the Python geopy package [49]. Notably, 36.1% of the organizations in

our datasets have multiple locations, but organizations have to indicate which one of the locations is considered their the headquarter. For 12,842 organizations (7.7% of all organizations), no location was provided. The account data from LinkedIn indicate the number of employees based on LinkedIn members working at the organization and the number of followers on Linkedin. Self-reported data have the following attributes: the founding year; the website address; a description of the organization; the number of employees (as a range); the industry of the organization; a list of hashtags that are relevant to the organization; and a list of "specialities", which comprises technologies, services, and industries related to the organizations' businesses. Overall we collected and analyzed the metadata of 166,459 organizations.

(2) Post data refer to all LinkedIn posts by the organization. Of note, we retrieved all LinkedIn posts of organizations from the first post of the company until the date of the data retrieval. Post data include the text of the post, the date of the post, all other referenced organizations with LinkedIn accounts, and whether the post is a repost or an original post. To avoid double counting, we excluded reposts from our analysis. For each organization, we retrieved all historical posts from the date when the organization joined LinkedIn to our retrieval date. To account for different languages, we classified the language of each post using the fasttext package of Python [50] and translated all non-English text into English using the Google Translate API [51]. Overall we collected and analyzed 26 million posts. The data was collected in 2023/24 and spans collaborations from January 2017 to December 2024.

6.2. Network graph

We created a bipartite network [52] of organizations and climate-tech innovation collaborations [52]. We define climate-tech innovation collaborations as individual projects (e.g., demonstration projects, deployments) and other types of interactions (e.g., equity investments, mergers and acquisitions) through which multiple organizations contribute to climate-tech innovation. We retrieved the collaborations from LinkedIn posts of organizations where we ana-

lyze whether the posts announce a collaboration between two or more organizations (see Section "Machine learning approach"). Then, our bipartite network is a graph $G = \{O, C, E\}$, where O and C represent nodes of organizations and collaborations respectively, and where $E = \{(o, c) : o \in O, c \in C\}$ are the links between the two. Both types of nodes, O and C, have further attributes related to organizations (e.g., founding year, organization type, locations) and collaboration projects (e.g.,, technology, collaboration type, post date for collaborations), respectively. As such, organizations are not directly connected in our graph but only through collaboration projects.

Based on the bipartite graph G, we further compute a transformed network $G' = (V', E')$ where V' are the organizations from above and where E' are edges linking two companies that are involved in the same collaboration. In the main paper, we call these direct edges "partnerships". We use the transformed network G' for all network statistics and all plots in the main paper. To avoid double counting, multiple partnerships between the same organizations and with the same characteristics (i.e. climate technology and type of collaboration) are projected onto a single link for a given period. The degree centrality of a specific organization node is then calculated as the sum of its number of partnerships. For plotting the network and computing network characteristics (e.g., clustering coefficient, degree centrality of nodes), we used the Python package NetworkX [53].

6.3. Machine learning approach

Our network was constructed by analyzing organizations and their social media posts with the large language model (LLM) Deepseek-V3 [54], which represents a Mixture-of-Experts language model consisting of 671B parameters that has outperformed other open-source models and achieved comparable performance with closed-source models, such as GPT-4o [55] and Claude-3.5-Sonnet [56].

We used an iterative process of two steps (details are described below): *(1) organization*

analysis and *(2) post analysis*.

(1) Organization analysis: In this step, we analyzed the organization data from LinkedIn. Specifically, we used the description, industry, and specialities self-reported by the organizations to classify the organization type (e.g., manufacturer, venture capital, university) and the sector of organizations. Instead of using the industry category from LinkedIn, we opted for our own set of categories for two reasons: first, the industry category does not provide information on the specific sector of organizations. Second, it only provides a single tag, while most organizations can be attributed to multiple organization types (e.g., utilities and project developers; accelerators and venture capital firms; governmental organizations and banks). Our sector categories are based on the STAN industry list by the Organization of Cooperation and Development [57] including all relevant sectors. Within the energy and climate sectors, we used granular sub-sector categories (e.g., hydrogen, direct air capture, oil, gas). The full list of categories for organization types and sectors are provided in the Supplementary Tables S S4 and S3). To automate the above analysis, we designed two classifiers: (i) an organization type classifier and (ii) a sector classifier, as described in the following.

(i) For organization type classifier, we leveraged in-context learning with few-shot examples containing manual annotations. Specifically, we constructed a prompt containing a short instruction and manual annotations of exemplary organization profiles with organization type labels (Supplementary Table S6). We designed the organization type classifier for multi-class classification, thus allowing each organization to be mapped to multiple organization types (e.g., an organization can be both a manufacturer and project developer). We tested the final performance of the classifier on manually annotated LinkedIn accounts, achieving an out-of-sample micro F1-score of 0.92 (Supplementary Table S4).

(ii) The sector classifier was developed analogously to the organizations classifier, but with a different prompt using the sector labels and different manual annotations of exemplary organi-

zations (Supplementary Tables S5). The sector classifier also allowed for multi-class classification (e.g., mapping an organization simultaneously onto both the oil and hydrogen sector). The sector classifier achieved an out-of-sample micro F1-score of 0.94 (Supplementary Table S3).

(2) Post analysis: In this step, we analyzed whether LinkedIn posts of organizations represent collaborations related to climate-tech innovation.

For this, we this we split the task in two steps. First, we trained and implemented a relevance classifier which only filters posts that are related to climate technologies. The classifier is based on the language model BERT, which has comparatively few parameters and thus lower computational costs [58, 59]. Second, we implement the LLM Deepseek-V3 [54] to extract collaborations from the posts and classify them regarding relevant types of collaboration and climate technologies. By pre-filtering relevant posts with a an efficient language model, we reduce the sample size of posts to be analyzed by the LLM by around 81.5% and thereby substantially saved computational costs [60]. The classifiers are as follows:

(i) The relevance classifier for climate technologies is a binary classifier that classifies whether a post is related to climate technologies. To be classified as relevant, a post must discuss a specific climate technology or primarily relate to decarbonizing a sector or process (e.g., decarbonization of cement). To deal with the large volume of input data (around 26 million posts), we trained a supervised classifier based on the language model BERT [58] on a set of 10,000 manually annotated posts (details on the model architecture and training setup are provided in the Supplementary Discussion S2). The climate-tech relevance classifier achieved an out-of-sample F1-score of 0.97 (Supplementary Table S8).

(ii) The collaboration extraction and classification is based on the Deepseek-V3 model in inference state with a prompt containing the task instruction and few-shot examples for in-context learning (the full prompt is provided in the Supplementary Table S9). It performs the following three tasks:

1. *Collaboration relevance:* the model classifies whether a post announces a collaboration between two or more organizations (which must be mentioned explicitly) and extracts the collaborating organizations. Relevant are only collaborations that are contributing to the innovation process. For example, we did not include joint participation in events such as panels, conferences, or industry fares. If multiple, independent collaborations are announced, the model extracts them separately.
2. *Technology classification:* the model maps each collaboration onto a list of 27 climate technologies (e.g., solar, hydrogen, and direct air capture; the full list + explanation, supplementary).
3. *Collaboration type classification:* the model maps each collaboration onto a list of 17 collaboration types (e.g., R&D collaboration, equity investment, and joint venture).

Both climate technology and collaboration type labels are non-exclusive, allowing collaborations to be mapped onto multiple labels.

For each task, the model's performance was validated ex-post via manual inspection of a random sample of classified posts. The collaboration relevance classification achieved an out-of-sample F1-score of 0.95; the technology classification achieved an out-of-sample micro F1-score of 0.93; and the collaboration type classification achieved an out-of-sample micro F1-score of 0.95 (detailed performances across individual categories are reported in the Supplementary Table S8).

6.4. Iterative process

Based on the organization analysis and the post analysis of the machine learning approach, we ran three iterations of our data collection process where, in each iteration, the network was extended by adding newly identified organizations and collaborations. To identify new organizations, we extracted the involved organizations from each collaboration. If an organization

was not yet included in our database, we added it to our database. For the newly added organizations, we restarted the machine learning approach from above retrieving and analyzing metadata and post data of organizations.

We ran the machine learning approach for three iterations: once with the initial set of organizations from BloombergNEF and once for the second set of newly added climate-tech organizations identified during the organization processing step. In the third iteration, we only retrieved and analyzed the organizational metadata to allow for complete analyses of innovation network participants. Overall, we started with 13,235 organizations from BloombergNEF and found 121,492 new organizations engaging in climate-tech innovation based on 399,318 identified collaborations in climate-tech.

6.5. Validation via i3 data

To assess the completeness of our network, we validated our data against the i3 database by the CleanTech group [61], which has not been included in our data collection process. The i3 database is considered one of the most thorough collections of information on firms and collaborations between firms in the global cleantech sector [11]. We validated our data along two dimensions: (1) completeness of organizations and (2) completeness of partnerships.

(1) Completeness of organizations: To validate the completeness of our LinkedIn organizations, we computed the share of companies from the i3 database that could be matched with organizations from our LinkedIn data (details of the matching approach are provided in the Supplementarey Discussion S3). From the 12,249 companies of the i3 database representing the energy & power and transportation & logistics industries, we could match 8,947 with our LinkedIn organizations (73.0%). Except for Russia (40% of the 34 Russian companies were matched), we found no significant negative deviations from the average (an overview of the share of matched i3 companies by country is provided in the Supplementary Table S10).

Furthermore, we manually evaluated a random sample of 100 companies from the 3,302

companies that were not matched with an organization from our LinkedIn data. Here, we found that only 56% of companies should have been captured by our LinkedIn data, as 44% represent companies that were found dead, changed in name, or acquired, or not showing a topical link to climate-tech. The 56% of companies that should have been captured represent relatively small companies (below 200 employees). More than half of them were missed due because of the source data (missing LinkedIn account or inactivity; details on the results are presented in the Supplementary Discussion S4).

(2) Completeness of partnerships: To validate the completeness of partnerships in our network, we manually analyzed the share of ”relationships” from the i3 database that match with an identical pair of organizations in the partnerships table of our LinkedIn data, both in the period between 2020 and 2024. To include only relevant relationships (i.e., not related to agriculture or recycling) we only included companies from the *Industry Groups* ”energy & power” and ”transportation & logistics”, resulting in 443 relationships in the i3 database.

From the 443 relationships, 350 (79.0%) had a matching pair of organizations in our partnerships table. From the non-matched relationships, 16 relationships involved companies that were not included as organizations in our network data. The distributions in the share of matched relationships across (of companies) and typess of relationships are reported in the Supplementary Table S12 and S11.

References

1. D. Popp, Environmental policy and innovation: a decade of research, National Bureau of Economic Research (2019).
2. IEA, Net zero by 2050, 2021. URL: https://www.iea.org/reports/net-zero-by-2050.
3. IEA, Energy technology perspectives 2020: Special report on clean energy innovation (2020).
4. J. Bistline, G. Blanford, M. Brown, D. Burtraw, M. Domeshek, J. Farbes, A. Fawcett, A. Hamilton, J. Jenkins, R. Jones, et al., Emissions and energy impacts of the Inflation Reduction Act, Science 380 (2023) 1324–1327.
5. European Commission, The Green Deal Industrial Plan, 2024. URL: https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal/green-deal-industrial-plan_en.
6. M. P. Hekkert, R. A. A. Suurs, S. O. Negro, S. Kuhlmann, R. E. Smits, Functions of innovation systems: A new approach for analysing technological change, Technological Forecasting and Social Change 74 (2007) 413–432.
7. K. S. Gallagher, A. Grübler, L. Kuhl, G. Nemet, C. Wilson, The energy technology innovation system, Annual Review of Environment and Resources 37 (2012) 137–162.
8. S. Haakonsson, Changing geography of clean tech, Nature Energy 5 (2020) 735–736.
9. C. Lechner, C. Leyronas, The competitive advantage of cluster firms: the priority of regional network position over extra-regional networks–a study of a french high-tech cluster, Entrepreneurship & Regional Development 24 (2012) 457–473.
10. H. Overholm, Collectively created opportunities in emerging ecosystems: the case of solar

service ventures, Technovation 39 (2015) 14–25.

11. C. Doblinger, K. Surana, L. D. Anadon, Governments as partners: The role of alliances in US cleantech startup innovation, Research Policy 48 (2019) 1458–1475.

12. J. Noailly, V. Shestalova, Knowledge spillovers from renewable energy technologies: Lessons from patent citations, Environmental Innovation and Societal Transitions 22 (2017) 1–14.

13. G. F. Nemet, How solar energy became cheap: A model for low-carbon innovation, Routledge, 2019.

14. A. Elia, M. Kamidelivand, F. Rogan, B. Ó. Gallachóir, Impacts of innovation on renewable energy technology cost reductions, Renewable and Sustainable Energy Reviews 138 (2021) 110488.

15. A. Stephan, L. D. Anadon, V. H. Hoffmann, How has external knowledge contributed to lithium-ion batteries for the energy transition?, iScience 24 (2021) 101995.

16. J. Huenteler, J. Ossenbrink, T. Schmidt, V. H. Hoffmann, How a product's design hierarchy shapes the evolution of technological knowledge—evidence from patent-citation networks in wind power, Research Policy 45 (2016) 1195–1217.

17. I. D. Lacasa, M. H. Shubbak, Drifting towards innovation: The co-evolution of patent networks, policy, and institutions in China's solar photovoltaics industry, Energy Research & Social Science 38 (2018) 87–101.

18. K. Surana, C. Doblinger, L. D. Anadon, N. Hultman, Effects of technology complexity on the emergence and evolution of wind industry manufacturing locations along global value chains, Nature Energy 5 (2020) 811–821.

19. A. Goldstein, C. Doblinger, E. Baker, L. D. Anadón, Patenting and business outcomes for cleantech startups funded by the advanced research projects agency-energy, Nature Energy

5 (2020) 803–810.

20. M. Cardenas Rodriguez, I. Haščič, N. Johnstone, J. Silva, A. Ferey, Renewable energy policies and private sector investment: Evidence from financial microdata, Environmental and Resource Economics 62 (2015) 163–188.

21. S. Halleck-Vega, A. Mandel, K. Millock, Accelerating diffusion of climate-friendly technologies: A network perspective, Ecological Economics 152 (2018) 235–245.

22. D. Fadly, F. Fontes, Geographical proximity and renewable energy diffusion: An empirical approach, Energy Policy 129 (2019) 422–435.

23. N. Ameli, S. Kothari, G. Livan, G. Caldarelli, The hidden structure of energy efficiency finance, Available at SSRN 3895862 (2021).

24. R. Gross, R. Hanna, A. Gambhir, P. Heptonstall, J. Speirs, How long does innovation and commercialisation in the energy sectors take? historical case studies of the timescale from invention to widespread commercialisation in energy supply and end use technology, Energy Policy 123 (2018) 682–699.

25. G. F. Nemet, V. Zipperer, M. Kraus, The valley of death, the technology pork barrel, and public support for large demonstration projects, Energy policy 119 (2018) 154–167.

26. N. Wang, K. Akimoto, G. F. Nemet, What went wrong? Learning from three decades of carbon capture, utilization and sequestration (CCUS) pilot and demonstration projects, Energy Policy 158 (2021) 112546.

27. Intergovernmental Panel on Climate Change, Climate change 2022: mitigation of climate change, 2022. URL: https://www.ipcc.ch/report/ar6/wg3/downloads/report/IPCC_AR6_WGIII_FullReport.pdf.

28. LinkedIn, Statistics, 2024. URL: https://news.linkedin.com/about-us#Statistics.

29. BloombergNEF, BloombergNEF, 2023. URL: https://www.bnef.com/.

30. IEA, ETP clean energy technology guide, 2025. URL: https://www.iea.org/data-and-statistics/data-tools/etp-clean-energy-technology-guide.

31. ITER, The project: Building ITER, 2025. URL: https://www.iter.org/building-iter.

32. Commonwealth Fusion Systems, SPARC: Proving commercial fusion energy is possible, 2025. URL: https://cfs.energy/technology/sparc.

33. U.S. Department of Energy, The infrastructure investment and jobs act, 2022. URL: https://www.energy.gov/sites/default/files/2021-12/FECM%20Infrastructure%20Factsheet.pdf.

34. F. Huber, On the role and interrelationship of spatial, social and cognitive proximity: Personal knowledge relationships of r&d workers in the cambridge information technology cluster, Regional Studies 46 (2012) 1169–1182.

35. M. Grubb, J.-C. Hourcade, K. Neuhoff, Planetary economics, Routledge London, 2014.

36. R. Mudambi, Location, control and innovation in knowledge-intensive industries, Journal of economic Geography 8 (2008) 699–725.

37. IRENA, Global hydrogen trade to meet the 1.5°c climate goal, 2022. URL: https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2022/Jul/IRENA_Global_hydrogen_trade_part_1_2022_.pdf.

38. BMWK, The national hydrogen strategy, 2020. URL: https://www.bmwk.de/Redaktion/EN/Publikationen/Energie/the-national-hydrogen-strategy.pdf?__blob=publicationFile.

39. U.S. Department of Energy, U.S. National Cöean Hydrogen Strat-

egy and Roadmap, 2023. URL: https://www.hydrogen.energy.gov/docs/hydrogenprogramlibraries/pdfs/us-national-clean-hydrogen-strategy-roadmap.pdf?sfvrsn=c425b44f_5.

40. G. Duranton, A. J. Venables, Place-based policies for development, National Bureau of Economic Research (2018).

41. T. Kemeny, M. Storper, Is specialization good for regional economic development?, Regional Studies 49 (2015) 1003–1018.

42. H. L. F. de Groot, J. Poot, M. J. Smit, Agglomeration externalities, innovation and regional growth: theoretical perspectives and meta-analysis, in: Handbook of regional growth and development theories, Edward Elgar Publishing, 2009.

43. Crunchbase, Crunchbase data, 2025. URL: https://about.crunchbase.com/.

44. B. Probst, S. Touboul, M. Glachant, A. Dechezleprêtre, Global trends in the invention and diffusion of climate change mitigation technologies., Nature Energy (2021).

45. K. Rajkumar, G. Saint-Jacques, I. Bojinov, E. Brynjolfsson, S. Aral, A causal test of the strength of weak ties, Science 377 (2022) 1304–1310.

46. World Population Review, LinkedIn users by country 2024, 2024. URL: https://worldpopulationreview.com/.

47. S. Pisano, L. Lepore, R. Lamboglia, Corporate disclosure of human capital via LinkedIn and ownership structure: An empirical analysis of European companies, Journal of Intellectual Capital 18 (2017) 102–127.

48. J. Park, I. B. Wood, E. Jing, A. Nematzadeh, S. Ghosh, M. D. Conover, Y.-Y. Ahn, Global labor flow network reveals the hierarchical organization and dynamics of geo-industrial clusters, Nature Communications 10 (2019) 3449.

49. GeoPy, GeoPy documentation, 2023. URL: https://geopy.readthedocs.io/en/stable/index.html.

50. fasttext, fasttext 0.9.2, 2023. URL: https://pypi.org/project/fasttext/.

51. Google Cloud, Cloud Translation API, 2023. URL: https://cloud.google.com/translate/docs/reference/rest/.

52. D. V. Filho, D. R. J. O'Neale, The role of bipartite structure in r&d collaboration networks, Journal of Complex Networks 8 (2020) cnaa016.

53. A. Hagberg, P. Swart, D. S Chult, Exploring network structure, dynamics, and function using NetworkX, 2008.

54. A. Liu, B. Feng, B. Xue, B. Wang, B. Wu, C. Lu, C. Zhao, C. Deng, C. Zhang, C. Ruan, et al., Deepseek-v3 technical report, arXiv preprint arXiv:2412.19437 (2024).

55. OpenAI, API platform, 2025. URL: https://openai.com/product.

56. ANTHROPIC, Claude 3.5 sonnet, 2025. URL: https://www.anthropic.com/news/claude-3-5-sonnet.

57. OECD, Stan industry list, 2023. URL: https://web-archive.oecd.org/2012-06-15/127386-40729523.pdf.

58. J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, Bert: Pre-training of deep bidirectional transformers for language understanding, arXiv preprint arXiv:1810.04805 (2018).

59. D. Kim, C. Park, S. Kim, W. Lee, W. Song, Y. Kim, H. Kim, Y. Kim, H. Lee, J. Kim, et al., SOLAR 10.7B: Scaling large language models with simple yet effective depth up-scaling, arXiv preprint arXiv:2312.15166 (2023).

60. M. A. K. Raiaan, M. S. H. Mukta, K. Fatema, N. M. Fahad, S. Sakib, M. M. J. Mim, J. Ahmad, M. E. Ali, S. Azam, A review on large language models: Architectures, applications, taxonomies, open issues and challenges, IEEE Access (2024).

721 61. Cleantech Group, i3, 2023. URL: https://i3connect.com/.

722 62. LinkedIn, China: Sunset of localized version of linkedin and launch of new injobs app later
723 this year, 2021. URL: https://www.linkedin.com/blog/member/product/
724 china-sunset-of-localized-version-of-linkedin-and-launch-of-new-injobs-

725 63. World Bank, World Bank Open Data, 2024. URL: https://data.worldbank.
726 org/.

727 64. N. Webersinke, M. Kraus, J. A. Bingler, M. Leippold, ClimateBERT: A pretrained lan-
728 guage model for climate-related text, arXiv preprint arXiv:2110.12010 (2021).

729 65. M. Toetzke, A. Stünzi, F. Egli, Consistent and replicable estimation of bilateral climate
730 finance, Nature Climate Change 12 (2022) 897–900.

731 66. A. J. Sietsma, J. D. Ford, J. C. Minx, The next generation of machine learning for tracking
732 adaptation texts, Nature Climate Change 14 (2024) 31–39.

733 67. V. Sanh, L. Debut, J. Chaumond, T. Wolf, DistilBERT, a distilled version of BERT: smaller,
734 faster, cheaper and lighter, arXiv preprint arXiv:1910.01108 (2019).

Data availability

Training data for the machine learning classifiers and aggregate data of the results will be made publicly available upon acceptance. All datasets to train and validate the machine learning classifiers are provided in the Source Data.

Code availability

The code used for the analyses will be made available through a public code repository upon acceptance. All code and datasets to train and validate the machine learning classifiers are provided in the Source Data. All analyses were conducted using Python 3.10.

Acknowledgements

The authors thank Francesco Re, Devjyoti Chakrabarti, Zsofia Rozenberszki, and Felipe Andrés Badilla Vidal for research assistance and Johannes Brehm, Florian Egli, Alejandro Nuñez-Jimenez, Tobias Schimanski, Anna Stünzi, and Hergen Wolf for helpful comments on earlier drafts of the paper. M.T. acknowledges funding from the Swiss National Science Foundation (SNSF) as part of grant 186932, from the Swiss Federal Office of Energy's "SWEET" program as part of the PATHFNDR consortium, as well as financial support from the TUM Global Postdoc Fellowship and the Friedrich Schiedel Fellowship.

Author contributions

M.T. conceptualized the idea, developed the methods, and performed the data analysis. All authors contributed to the interpretation of the results and the writing of the paper.

Competing interests

The authors declare no competing interests.

760

Supplementary Information for

Machine learning can help to track dynamics in climate technology innovation

Malte Toetzke[a,*], Benedict Probst[b,a,e], Stefan Feuerriegel[c,d], Laura Diaz Anadón[e,f], Volker Hoffmann[a]

[a]ETH Zurich, Zurich, Switzerland
[b]bMax Planck Institute for Innovation and Competition, Munich, Germany
[c]LMU Munich, Munich, Germany
[d]Munich Center for Machine Learning, Munich, Germany
[e]University of Cambridge, Cambridge, United Kingdom
[f]Harvard University, Cambridge, United States

[*]Correspondence to: mtoetzke@ethz.ch

This PDF file includes:

Supplementary Discussions S1–3

Supplementary Figures S1–5

Supplementary Tables S1–12

Supplementary Discussions

S1: Discussion of data coverage across countries

Our analyses are based on the assumption that organizations, primarily firms, are well represented on LinkedIn. As LinkedIn is a professional social media platform, incentives for organizations to be active on LinkedIn are different from those on other social media platforms, such as Twitter or Facebook. Besides engaging with the public and advertising products and services, organizations on LinkedIn report information to attract and retain important stakeholders, such as employees, business partners, and investors, within and outside the country of operation [47, 48, 45]. Reportedly, LinkedIn comprises over 67 million accounts representing organizations [28].

With more than 900 million global users, LinkedIn is considered the largest professional social media platform worldwide [45, 46]. Countries with the highest absolute numbers of users are the US (200 million), India (99 million), and China (60 million; Supplementary Table S1). However, since 2021, LinkedIn deploys a separate service in China, called InJobs [62]. For this reason, we do not consider our data on China as comprehensive as it is likely biased toward Chinese firms with an international presence and may omit organizations with a primary focus on domestic markets.

Overall, the number of users (per capita) is correlated with gross domestic product (per capita) across countries (Supplementary Figure S1). This suggests that the likelihood of individuals to join LinkedIn based on income is fairly similar across countries. The underlying intuition is that, with growing income, individuals focus on building professional networks, which LinkedIn facilitates through its platform [45]. From this, we conclude that, in countries that align with this pattern, LinkedIn likely provides similar incentives for organizations to join the network and disclose information to attract talent and partners.

Examples of countries with over-proportionally high user rates compared to GDP per capita

are the United Arab Emirates, Brazil, Canada, and Singapore. Countries with relatively low user rates are Korea and Japan. As such, organizations in Korea and Japan are also likely underrepresented in our data.

Finally, countries for which we have no organization in our data are Eritrea, Komoros, Nauru, Palau, Saint Vincent and the Grenadines, Tuvalu, and the Vatican, which is not surprising as these countries are relatively small and may not have developed industries in climate-tech.

Table S1: Top 30 countries by total number of LinkedIn users from [46] based on data from 2023.

Country	**LinkedIn users (in 1'000s)**
United States	200,000
India	99,000
China	60,000
Brazil	59,000
United Kingdom	35,000
France	26,000
Indonesia	23,000
Canada	21,000
Mexico	19,000
Italy	17,000
Spain	17,000
Germany	15,000
Australia	14,000
Philippines	13,000
Turkey	13,000
Colombia	12,000
South Africa	11,000
Argentina	11,000
Netherlands	10,000
Pakistan	9,300
Egypt	7,900
Peru	7,700
Nigeria	7,500
Saudi Arabia	7,300
Malaysia	7,000
Chile	6,900
United Arab Emirates	6,600
Bangladesh	5,900
Poland	5,500
Vietnam	5,200

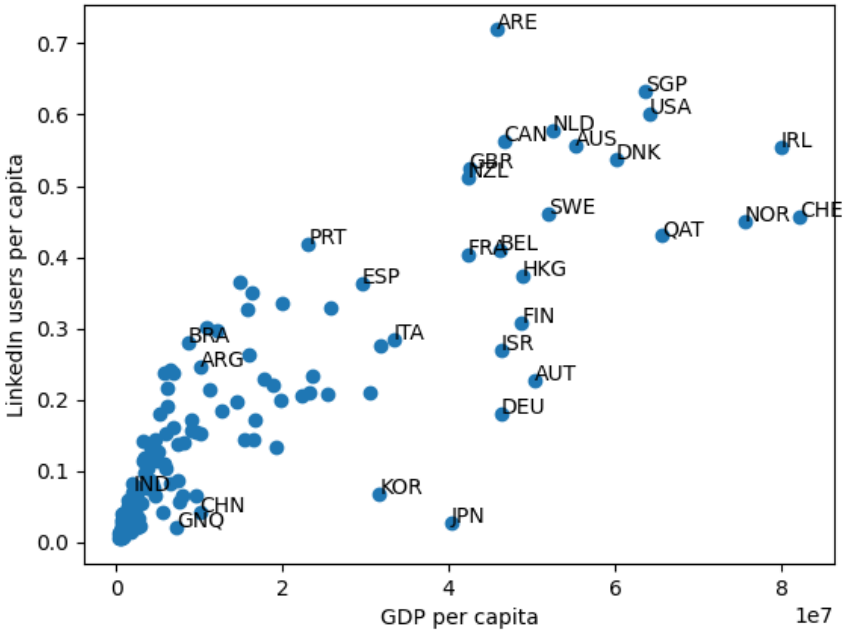


Fig. S1: Comparison between number of LinkedIn users per capita and gross domestic product (GDP) per capita across countries based on data from [46, 63].

S2: Implementation details of the relevance classifier for climate technologies

The relevance classifier for climate technologies is a sequence classifier based on supervised learning. As the underlying model, we used ClimateBERT [64], which is a domain-specific language model that has proven to perform well in the context of climate and energy [64, 65, 66]. ClimateBERT is based on a distilled variant of BERT [58, 67] and has been fine-tuned on 1.6 million additional climate-related texts. We trained the model 50 epochs using a cross-entropy loss and the AdamW optimizer. Furthermore, we used a batch size of 64 and a learning rate of 1×10^{-6}. To avoid overfitting, the model with the highest validation score during training was selected.

As input for training, we randomly sampled 10,000 LinkedIn posts from our initial set of organizations. After excluding reposts, which lead to multiple appearances of the same text, we manually annotated 9,645 texts regarding their relevance for climate technologies and decarbonization, of which 2,482 of the texts were annotated as relevant. We used a train/ validation/test split of 0.85/0.075/0.075. As such, 724 annotated texts where used for testing the final performance of the classifier (Supplementary Table S8). The training, validation, and test data are provided in the Source Data.

S3: Implementation details of the validation via the i3 data regarding the completeness of organizations

The i3 database contains a table with 39,144 companies (as of August 2024). These companies represent diverse industries, such as "agriculture and food", "energy & power", "waste & recycling". To achieve a sample of companies engaged in climate-tech in accordance with our focus, we filtered the variable *Industry Groups* for the categories "energy & power" and "transportation & logistics", resulting in 12,249 companies. We combined two matching techniques: first, we matched companies from the i3 database with our organizations via the indicated websites, by using the stems of the web-address (e.g., converting *http://www.machadomeyer.com.br* to machadomeyer; *https://www.co-opbank.co.ke/* to co-opbank). Second, we matched the remaining companies via their names, dropping legal entities at the end of company names (e.g., *Topspin Partners LP* to Topspin Partners, *China Sunergy Nanjing Co Ltd* to China Sunergy Nanjing).

S4: Details of the results from the manual validation of 100 i3 companies that were not matched to our network data

From the 100 manually evaluated companies, We found that only 56% were truly missed by our LinkedIn data (27 were found dead; 1 had changed its name; 1 was acquired; 1 could be matched to our data manually; 14 firms were found to have no link to climate-tech). From these 56 companies, 32 (57.1%) were missed because of the source data from LinkedIn (17 companies had no LinkedIn account and the LinkedIn account of 15 companies was found inactive). Furthermore, all companies were found relatively small based on the number of employees (all companies had below 200 employees and only 7 above 50 employees).

Supplementary Figures

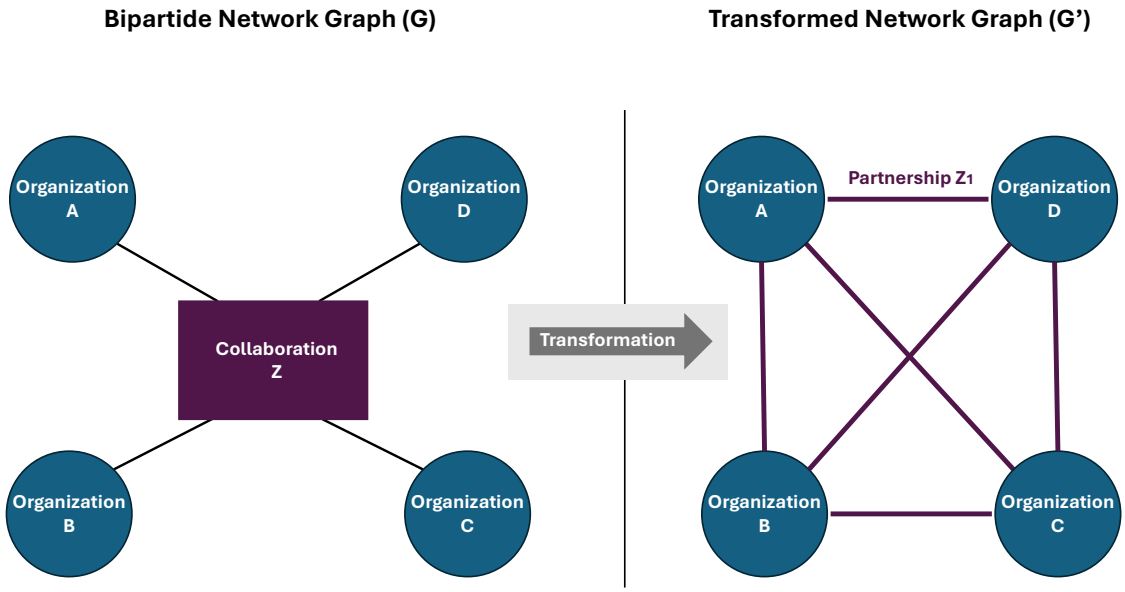


Fig. S2: Transformation from bipartide network graph G to transformed graph G'. In the first graph, organizations can only be linked to collaborations. In the second graph, we project these collaborations onto direct edges which we call "partnership". As illustrated, one collaboration can map onto multiple partnerships.

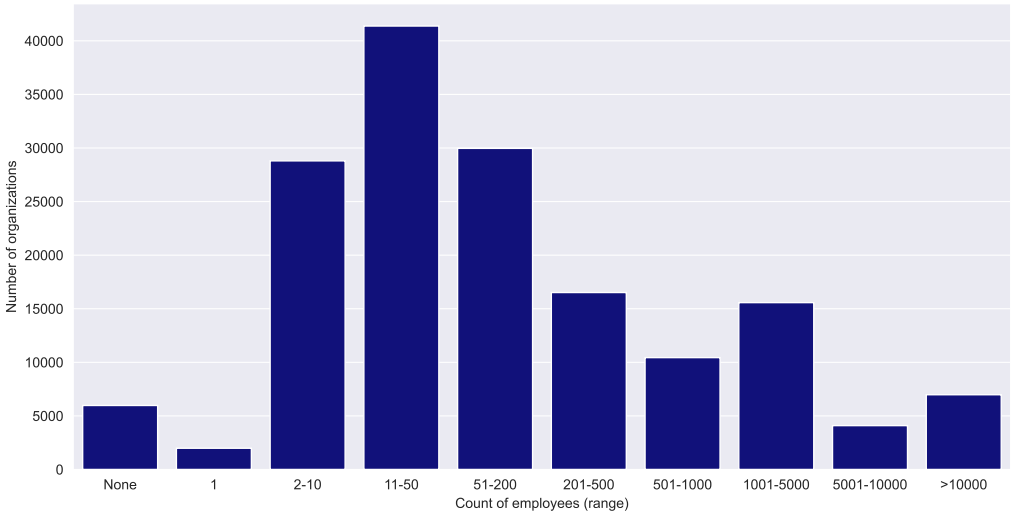


Fig. S3: Frequency of organizations across different counts of employees. These numbers are based on the ranges self-reported by the organization accounts on LinkedIn. Based on reported employment from LinkedIn profiles of members, all captured organizations jointly employ 128,296,940 people. Given that not all employees have a Linkedin account, this is likely a lower bound of the actual number.

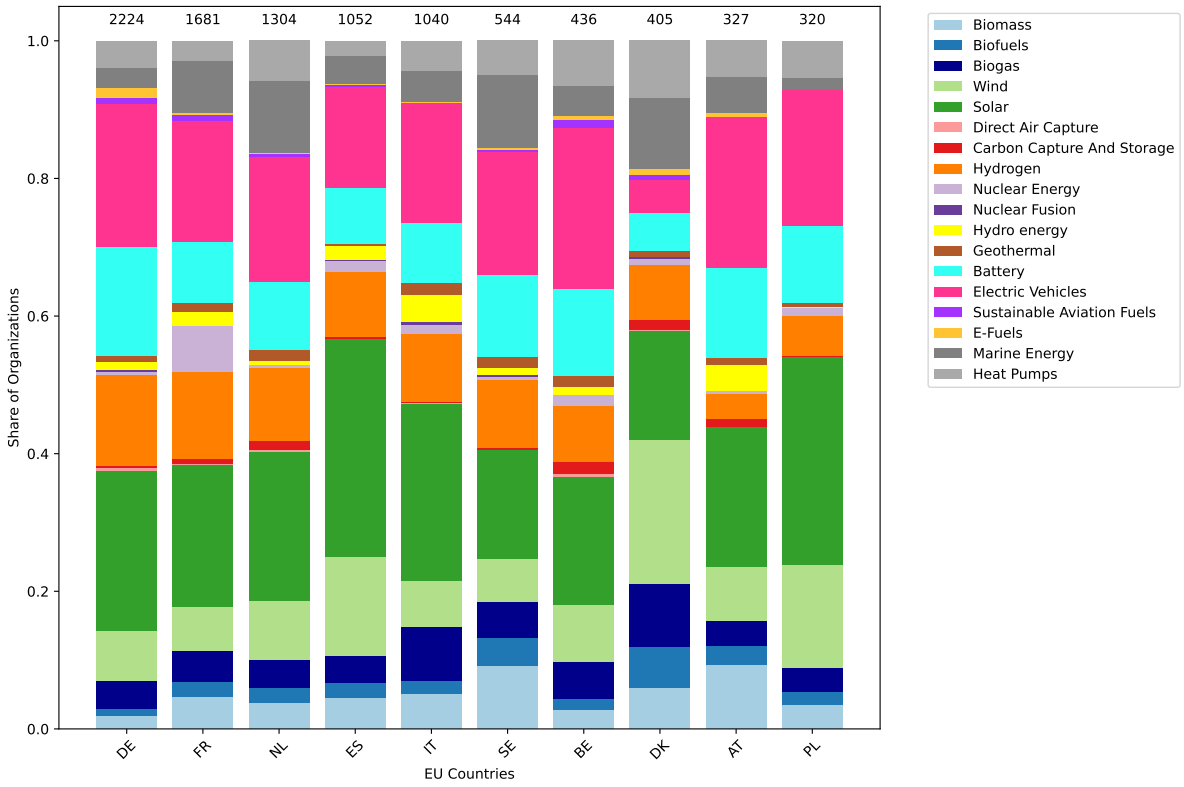


Fig. S4: Relative concentration of organizations in climate technologies by EU countries. Order is based on the absolute number of organizations specialized in climate technologies, which is indicated on top of the bars.

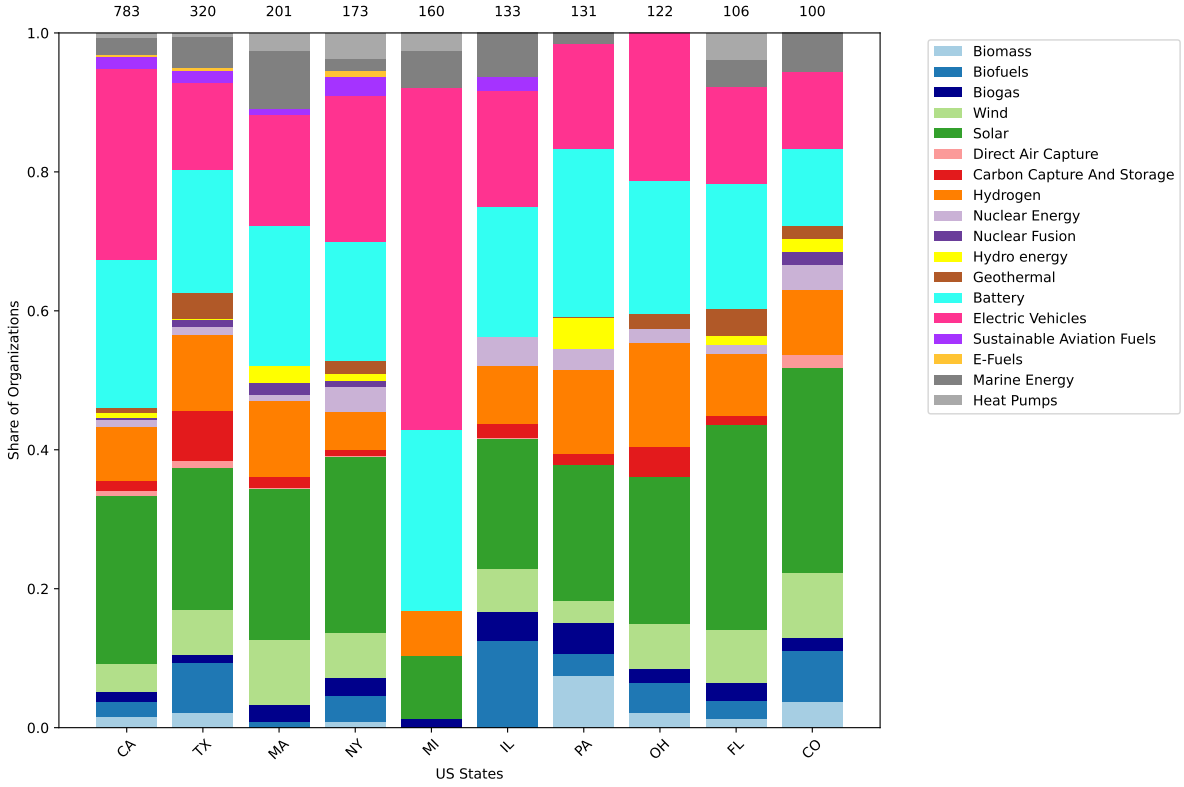


Fig. S5: Relative concentration of organizations in climate technologies by US states. Order is based on the absolute number of organizations specialized in climate technologies, which is indicated on top of the bars.

Supplementary Tables

Table S2: **Types of climate technologies.** Note that each post can be assigned to multiple technologies.

Technology	Description	Exemplary posts		
Solar	All types of solar photovoltaic (PV) applications from small residential roof-top systes to utility-scale applications.	we are very excited to announce our investment in pvmodule coating pioneer phytonics, alongside high-tech gründerfonds,karlsruher institut für technologie (kit)/ kit-gründerschmiedeand business angles. phytonics ' novel coating enables up to 10 % higher annual electrical energy yields through higher efficiency in capturing and converting sunlight, while at the same time making the modulesglare-free. (...)	announcement: we today bring you news of our partnership with aberdeen city council on a smart social housing project, supported by the uk government's social housing decarbonisation fund (shdf) demonstrator. utilising sms's newly launched solopower(a financed solar pv and battery storage solution), the project will additionally trial the electrification of heat and fabric upgrades as part of a whole-home approach to reducing co2 and fuelpovertyacross an initial 100 homes.	our strong cooperation with fortis energy continues. we have supplied over 100 mw panels to fortis enerji, which we have provided over 100 mw of panels so far, and this time, we have shipped 2,432 panels for the 1,097 kwp ges project in uzundere. thank you
Electric Vehicles	All non-hybrid types of battery electric vehicles, such as electric cars, buses, and trucks. Corresponding collaborations can also refer to the necessary infrastructure (e.g., charging stations) and related business models (e.g., shared e-mobility).	enertech capital is pleased to announce today, our first investment from the enertech global strategic mobility fund in leading electric vehicle ("ev") battery intelligence company recurrent."we are pleased to have enertech join the syndicate of investors in our latest round." (...)	congratulations to arcus asset company horizon energy infrastructure on their new partnership with good energy, creating support for uk business customers looking to install ev charging infrastructure!	vdl nedcar has been instructed to take the very first models of the mini cabrio electric into production. the mini cooper se convertible is the world's first convertible with a 100% electric powertrain. the production of the model is limited to 999 pieces and starts in april 2023.
Hydrogen	All technologies related to the production, transport, storage, and usage of hydrogen across different sectors.	totalenergies has opened a public hydrogen station on the territory of vdl in veldhoven. "vdl groep plays a prominent role in the field of sustainable energy, for both stationary and mobile applications."said pieter van der leegte. (...)	alstom in partnership with the government of quebec, chemin de fer charlevoix, train de charlevoix , harnois énergiesand htechas announced a passenger service demonstration project for its coradia ilint hydrogen-powered train.	we've signed a collaboration agreement with liebherr group in order to optimise hydrogen fuel cells alstom's hydrogen strategy is part of its ambition to innovate by pioneering smarter and greener mobility solutions for all. this agreement is aimed at optimising hydrogen systems, including improving the reliability and durability of fuel cells, increasing their power density and reducing the cost of such solutions.
Wind	Onshore and offshore wind turbines as well as the manufacturing of turbine blades.	we are excited to announce that we've entered into a power purchase agreement with Orsted, supporting the construction of the sunflower wind project, which will match 100% of our power consumption across our entire operational footprint with renewable energy credits.(...)	Our supported company act blade continues to inspire us with its winds of change! after a six-month test in scotland, they explain how their greener, leaner high-tech textile rotor blade s fly through trials in the uk's biggest onshore wind farm. (...)	aerones , supported by eit innoenergy, is revolutionising the maintenance of wind turbines. recent data from inspections carried out by the latvian company, on more than 1,500 blades in different environmental conditions across the world, show that approximately 23% of the blades have broken or damaged spcrs.

Batteries	All types of battery applications for energy storage, including mobility and grid storage applications.	we are pleased to announce that the im3ny new york lithium ion battery plant is now 51% complete see our latest asxannouncement https://lnkd.in/gb_vhxir	the finnish cleantech company tracegrow offers a sustainable solution for recycling valuable materials back into circulation. its technology converts industrial side streams and recycled batteries into valuable products for agriculture and other industries without the need for energy-intense mining operations and reducing its environmental impact. nefco - the nordic green bank's fast-track loan will help tracegrow implement its new growth strategy and allow it to hire additional personnel to strengthen its global reach.	kore power has acquired a 4.22% stake in zevx inc., an innovative phoenix, az area startup focusing fleet electrification and ev conversion kits. this investment expands the company's presence in the e-mobility space after entering into a supply agreement to provide kore battery modules for zevx 's electric powertrain kits through 2030. read more: https://lnkd.in/ggbjshdx
Biogas	Production, transport,storage, and usage of renewable gases based on raw materials, such as agricultiral waste, municipal waste, and plant material.	eit - european institute of innovation and technology innoenergy supported the start-up orchestra which produces economically viable technology for biogas upgrading using selective co2 separation for removal and purification.	a territorial methanisation site in haute-savoie, france, was inaugurated yesterday. for the first time ever, @ ifp energies nouvelles and @ arol energy demonstrated a new biogas purification technology called ae-amine.it will use treated biogas to produce 110 nm3/h of biomethane providing 8000 mwh of energy each year, equal to the annual heating needs of 1 200 people!	we are turning organic waste into renewable natural gas with our partners stormfisher. learn how we're creating rngto heat homes and more at https://lnkd.in/e_43kisc.
Carbon Capture and Storage	All applications capturing CO_2 directly from a point source (e.g., an energy plant). For Bioenergy with Carbon Capture and Storage (BECCS) applications both CCS and the bioenergy source are classified as relevant.	pilot energy limited, and canadian carbon capture & removal solutions providers vante have signed a memorandum of understanding (mou) to collaborate on offering a one-stop-shop solution for carbon capture, transportation, and storage to industrial businesses with hard-to-avoid co2 emissions. pilot energy and svante will initially target the decarbonisation of $\sim$8-million tonnes per annum of co2 emissions from western australia's kwinana industrial area. find out more:bit.ly/3gwazl8	kudos to the translational energy research centre on advancing carboncapturetechnology. we're excited to collaborate with them to expand options for zero carbon solutions.	we're excited to see the carbon capture unit being built at landmark power holdings and victory hill capital partners llp's site in rhodesia. one step closer to cleaner, more flexible power generation!
Biofuels	Production, transport,storage, and usage of renewable fuels from biomass.	c-green's full-scale oxypower htc plant for converting industrial biosludge to renewable biofuel is now under construction. the facility will recycle 16,000 metric tons of pulp biosludge a year at stora enso's fluting paper mill in heinola, finland. the goal is to reduce annual carbon dioxide emissions by as much as 2,500 tons. http://bit.ly/2miufca	we are proud to provide families & businesses with an affordable and reliable energy choice while advancing innovative low-carbon solutions like green fuels. we are excited to partner with municipality of chatham-kentand waste connections of canada to advance these important projects!	congratulations to pkn orlen s.a. andorlen poudnie s.a. for kicking off the execution of their second-generation cellulos ethanol plant in poland. we look forward to our cooperation to make the plant producing biofuels with our sunliquidtechnology a great success.
Hydro Energy	All types of hydropower plants including conventional dams and pumped storage.	tractebel and partner norconsult as celebrated the successful completion of the 456 mw upper tamakoshi hydroelectric project on september 27 at heritage garden, lalitpur, nepal.	the construction of the expansion of the downstream batard began on the hydroelectric development project of the nachtigal dam, in cameroon	new lai chau hydropower plant completed one year early! ge renewable energy and its consortium partner powerchina zhongnan engineering corporation recently received the final acceptance certificate from vietnam electricity (evn) for the new laichauhydropower plant. this great milestone marks the full and successful completion of the plant, one year ahead of schedule.

Sustainable Aviation Fuels	All types of sustainable aviations fuels including biofuel-based aviation fuels and hydrogen-based aviation fuels.	neste my sustainable aviation fuel is introduced for the first time at gatwick airport through a collaboration between neste, q8aviation, easyjet and gatwick airport ltd.this important milestone marks the first time a departing flight at gatwick has used sustainable aviation fuel (saf) and isalso the first usage by any easyjet service.	aer lingus enters into new fuel sales agreement with gevo for 6.3 million gallons of sustainable aviation fuel per year over five years.	breakthrough energy catalyst grantee lanzajet is building a groundbreaking facility that is expected to produce 9m gallons of sustainable aviation fuel annually. the facility will help accelerate the global transition away from high-emission jet fuels. more: https://nt-z.ro/3ixhczc
Nuclear Fission	All processes related to energy generation from nuclear fission including nuclear material production and waste treatment as well as mall-scale and large-scale reactors.	icymi: our partners deep isolation launched the new deep borehole demonstration center at wmsym2023 for nuclear waste in phoenix, arizona!	we are excited to announce that we have signed a contract with dominion energy to design, manufacture and deliver six replacement steam generators at surry nuclear power station in virginia. the new deal reinforces our long-term partnership with dominion to help support surry's operations through 2053.the nuclearenergyproduced by the power station provides carbonfreeelectricity year-round to 419,000 homes.	how to integrate small modular reactors into low-carbon hybrid energy systems. tractebel leads the industrial users group (iug) of tandem project, a euratom supply agency horizon europe research project, that aims at facilitating the safe & efficient integration of smrs into smart low-carbon hybrid energy systems. learn more: https://tandemproject.eu/
Heat Pumps	All applications of heat pumps including production, installation, and operation.	we are thrilled to announce endurant energy's acquisition of geoptimize inc., a specialist in heating and cooling and ground source heat pump gshp design company.	planning has been submitted for a renewable-led energy centre with water source heat pump technology to help power the mersey heat district heat network at peel liverpool waters. we're pleased to be working withpeel l&pas the project enters this exciting phase.	the first heat pump installations of our customer quattzijn are now successfully implemented. the commander in chief is a good result of our intensive cooperation. in the coming period, production will be increased so that more houses can be provided with this beautiful product.
Biomass	All solid types plant-based biomass for energy generation, such as charcoal and pellets.	each year, the new facility will produce 10,000 tonnes of biocarbon as a coal fuel alternative for ontario's heavy industries. the biocarbon is expected to provide enough gas in thorold for 5,500 ontario homes. natural gas must be part of a low-carbon energy future. wastetoenergyinitiatives such as char technologies ltd. biomass to rng facility will improve agricultural and landfill environmental conditions, produce rng and biocarbon, reduce emissions, and advance ontario's sustainability and energysecurity	a team of foley lawyers recently acted as legal counsel to greenleaf power consolidated llc (greenleaf) in securing $36.7 million in project financing from east west bank through its wholly-owned subsidiary greenleaf biomass holdings llc. the transaction marks one of the largest biomass power project financings in recent north american history.	greenwood sustainable infrastructure, our north american sustainable infrastructure investment company, has announced that its michigan-based clean power business, warden electric company, is undertaking a major upgrade at its flagship biomass power plant that will include refurbishment of its ge turbine, as well as a number of improvements to the boiler and related systems. this is the first phase of a refurbishment programme expected to exceed $6 million in capital expenditure. (...)
Marine Energy	All types of applications to generate electricity from ocean waves, tides, and temperature differences.	a tidal-powered airport might not be far off! the team atemec: european marine energy centre is working through the logistics of turning o2's tidal power into green hydrogen, that can then be used for carbon-free local air services here in orkney.	corpower ocean launches pioneering wave energy project in northern portugal.with a 16 meur investment plan, corpower ocean is establishing an r&d, manufacturing & service centre for wave energy converters in viana do castelo.“ corpower's goal is to successfully introduce certified and warrantied wec products to the market by 2024	swedish marine energy developer minesto has secured a power purchase agreement that will support a prospective tidal energy project in the faroe islands. minesto signed a ppa with the faroese electric utility company elfelag sev - electrical companyfor two planned 100kw systems, featuring the swedish company's subsea kite technology, and an additional 2 mw of capacity for large-scale tidal energy systems.

Geothermal Energy	All types of applications to extract heat from the earth's crust.	eavor is proud to sign an agreement with getechto explore geothermal opportunities across latin america! read the full blog here - https://lnkd.in/gxvfvnge eavor energy #geothermal #renewable #energyforeavor #geothermalenergy	we are thrilled to announce the close of our series b round of funding led by breakthrough energy ventures, along with continued support fromgv, comcast ventures, new enterprise associates (nea), collaborative fund, lennar, and building ventures. this new investment enables us to expand operations, deepen our research & development and bring sustainable geothermal heating to the millions of homes that need it.here's to the decarbonized future ! https://lnkd.in/dz78qay #geothermalenergy	wien energie gmbh and omv join forces to develop deep geothermalenergy in the vienna region. joint venture planned for the design and development of geothermal potential in the vienna basin as well as for the operational management of deep geothermal plants.
E-Fuels	All types of electric fuels based on hydrogen.	hy24 , mirova , cdpq and technip energies are joining forces to invest €200m in hy2gen ag, a pioneer in the production of green e-fuels for mobility, agriculture and industry.societe generale acted as financialadvisorto hy24 and mirova on this project, the largest private green hydrogen-focused capital raise to date and paving the way for the decarbonizationof industrial and transportation sectors.	sasol chemicals provides the clean tech company ineratec with catalyst for the fischer-tropsch reactors. the plant at industriepark hoechst will produce up to 2 500 tons of synthetic fuels with power-to-liquid technology.read more https://lnkd.in/dw6qktzd	next milestone for power to liquids production at the hamburg h&r location. p2x europe gmbh & co kg, the joint venture between h&r group and mabanaft, is gaining momentum: the production and marketing of synthesis-based efuels and waxes from renewable raw materials at the hamburg site is set to start soon. with the innovative demonstration plant by technology partner ineratec gmbh around 200 metric tons of e-fuels per year for mobility, and around 150 metric tons of waxes for applications in the cosmetics, pharmaceuticals and food industries will be produced. https://lnkd.in/ewefbemf
Nuclear Fusion	All types of applications generating energy from nuclear fusion.	we're thrilled to announce that we're collaborating with experts at canadian nuclear laboratories to advance commercial fusion power in canada. learn more: http://bit.ly/3udr8vu	the ga energy group is excited to announce a new collaboration with tokamak energyto advance high temperature superconducting magnet technology for fusion energy and other applications. this collaboration will leverage our world-leading capabilities for manufacturing large-scale magnet systems and te's pioneering expertise in hts magnet technologies.read more: https://bit.ly/42fmpi5 fusionenergy	we're building a world-first tritium research centre to develop fuel technology for fusion energy devices across the globe - and we're pleased to announce that atkins will be designing it
Direct Air Capture	All processing steps related to capturing CO_2 from the atmosphere via direct air capture.	oxy and our subsidiary 1pointfive to begin construction of the world's largest direct air capture plant in the tx permian basin, advancing large-scale carboncaptureto help oxy and others reduce emissions and accelerate the path to netzero. learn more: https://bit.ly/3wtb5ue	we're pleased to announce that our new large-scale direct air capture plant "orca" has successfully achieved independent third-party validation from dnv, the global leader in quality and risk assurance. (...)	we're pleased to share that the u.s. department of energy (doe) officially selected fervo's red rocks direct air capture hub in southwest utah for $2.9 million in federal funding.
Biochar	Production and use of biochar based on pyrolysis to extract CO_2 from the atmosphere.	partnership announcement: we are happy to announce our new partnership with vanmander, the first company to trade biochar-based carbon removal credits in spain and portugal, and our first spanish supply partner!	the biggest and most technologically advanced purpose built biochar manufacturing facility in the us takes shape. our two pyreg gmbh p-1500 machines were delivered last week and supporting infrastructure is on the way. this project has been made possible with financing fromcei (coastal enterprises, inc.), finance authority of maine, maine technology institute, grants from the u.s. forest service.	aston university is to receive £1.8 million to transform the west midlands into a powerhouse of low-carbon product development and commercialisation.the university will be building on its existing research facilities to lead the region's biochar cleantech accelerator as part of the west midlands innovation accelerator.

Table S3: Performance of the sector classifier on test data.

Sector	Precision	Recall	F1-score	Number of true positives
Climate technology sub-sectors				
Battery	1.00	1.00	1.00	6
Biofuels	0.86	0.86	0.86	7
Biogas	0.86	1.00	0.92	6
Biomass	0.90	1.00	0.95	9
Carbon Capture And Storage	1.00	1.00	1.00	8
Carbon Offsets	0.80	1.00	0.89	4
Direct Air Capture	0.60	1.00	0.75	3
Electric Vehicles	0.78	1.00	0.88	7
E-Fuels	1.00	0.83	0.91	6
Geothermal	0.86	1.00	0.92	6
Heat Pumps	0.83	1.00	0.91	5
Hydrogen	0.90	0.90	0.90	10
Hydro energy	0.80	1.00	0.89	4
Marine Energy	0.50	1.00	0.67	3
Nuclear Fission	1.00	1.00	1.00	11
Nuclear Fusion	1.00	1.00	1.00	5
Other Carbon Removal	0.75	1.00	0.86	6
Other Energy Storage	0.80	1.00	0.89	8
Solar	0.93	1.00	0.97	14
Sustainable Aviation Fuels	0.83	1.00	0.91	5
Wind	0.85	1.00	0.92	11
Other Renewables	1.00	0.85	0.92	20
Fossil energy sub-sectors				
Coal	1.00	1.00	1.00	5
Gas	1.00	0.89	0.94	18
Oil	1.00	0.89	0.94	18
Other Fossil Energy	1.00	1.00	1.00	5
Other energy-related sub-sectors				
Electricity Grid	1.00	1.00	1.00	8
Energy General	1.00	0.95	0.97	20
Transport sub-sectors				
Automotive	0.83	1.00	0.91	5
Aviation	0.93	1.00	0.96	13
Railway	1.00	1.00	1.00	5
Shipping	1.00	1.00	1.00	7
Other Transport	0.86	0.92	0.89	13
Other relevant sectors				
Built Environment	1.00	0.89	0.94	9
Chemical Industry	1.00	1.00	1.00	11
Computer and Electronic Products	1.00	0.83	0.91	6
Iron and Steel	1.00	0.88	0.93	8
Military	0.83	1.00	0.91	5
Other Construction	1.00	0.95	0.97	20
Other Manufacturing	0.89	1.00	0.94	16
Raw Materials	0.83	1.00	0.91	5
Software	1.00	0.93	0.97	15
Sector classifier (avg.)	0.92	0.95	0.94	376

Table S4: Performance of the organization type classifier on test data.

Organization type	**Precision**	**Recall**	**F1-score**	**Number of true positives**
Accelerator/Incubator	0.80	1.00	0.89	4
Bank	0.83	1.00	0.91	5
Consulting	1.00	0.86	0.92	7
Event Organizer	0.67	1.00	0.80	4
Governmental Organization	0.94	1.00	0.97	16
International Organization	0.80	1.00	0.89	4
Lobby Group	1.00	1.00	1.00	5
Manufacturer	0.88	1.00	0.93	14
Media Firm	0.80	0.80	0.80	5
Mining	0.80	1.00	0.89	4
Non-Governmental Organization (NGO)	0.88	0.88	0.88	8
Other Financing Organization	1.00	0.82	0.90	11
Project Developer	1.00	0.83	0.91	6
Research Institute	1.00	1.00	1.00	5
Service Provider	0.82	1.00	0.90	18
University	0.80	1.00	0.89	4
Utility	1.00	1.00	1.00	5
Venture Capital	1.00	1.00	1.00	6
Organization type classifier (avg.)	0.89	0.95	0.92	131

Table S5: Prompt used for the sector classifier for organizations.

You are a research assistant classifying organization profiles on LinkedIn. Classify the organization into one or multiple of the following sectors:
[Biomass, Biofuels, Biogas, Wind, Solar, Direct Air Capture, Carbon Capture And Storage, Other Carbon Removal, Carbon Offsets, Hydrogen, Nuclear Energy, Nuclear Fusion, Hydro energy, Geothermal, Battery, Other Energy Storage, Electricity-grid, Electric Vehicles, Sustainable Aviation Fuels, E-Fuels, Marine Energy, Heat Pumps, Other Renewables, Gas, Oil, Coal, Other Fossil Energy, Energy General, Railway, Aviation, Shipping, Automotive, Other Transport, Chemical Industry, Built Environment, Other Construction, Computer and Electronic Products, Software, Other ICT, Agriculture, Iron and Steel, Other Manufacturing, Military, Raw Materials, Other Sectors]
Use the label "Other Sectors" if the organization does not specify any sector (e.g. financial services in general) or the sector does not fit to any of the labels (e.g., textiles).

Here are some exemplary classifications:
Text: Fraunhofer-Institut für Solare Energiesysteme ISE; The Fraunhofer Institute for Solar Energy Systems ISE in Freiburg, Germany is the largest solar research institute in Europe. With a staff of about 1400, we are committed to promoting a sustainable, economic, secure and socially just energy supply system based on renewable energy sources. We contribute to this through our main research areas of energy provision, energy distribution, energy storage and energy utilization. Through outstanding research results, successful industrial projects, spin-off companies and global collaborations, we are shaping the sustainable transformation of the energy system. https://www.ise.fraunhofer.de/en/publishing-notes.html; ['Photovoltaik', 'Energiesystemtechnik', 'Solarthermie', 'Gebäudeenergietechnik', 'Wasserstofftechnologien']; ['#energytransition', '#research', '#solar']
Classification: Solar
Text: Esken; We own two core assets; London Southend Airport and Stobart Energy.London Southend Airport is an award-winning airport serving London and the South East. It has ambitions plans to grow and the key ingredients in place to enable that growth. And; in Stobart Energy we have a transformational and maturing business which supplies fuel to biomass plants across the UK so that they can, in turn, create renewable energy.; ['Infrastructure and Support Services', 'Railway Maintenance', 'Biomass Energy', 'Aviation', 'Biomass'];
Classification: Aviation, Biomass, Railway
Text: LanzaTech; Where others see a dire choice, LanzaTech sees a trillion-dollar opportunity.The good news is after 15 years, north of a thousand patents, and millions of hours of pioneering scientific inquiry, LanzaTech has invented a technology big enough to meet the moment. One that transforms pollution into profit, and ensures that humans continue to prosper far into the post-pollution future.The science is state-of-the-art, but the idea is simple. We use nature to heal nature. First we capture carbon emissions. Then we feed them into bioreactors where trillions of carbon-hungry microbes go to work. These tiny dynamos eat the pollution and output valuable raw material commodities. Pure enough to be resold, repurposed, recycled, renewed, re-everythinged — from sustainable jet fuel to yoga pants.It's a commercial scale solution that's ready for market today. In a crowded sector filled with speculation but short on results, our plug-and-play platform is already making our customers money. Turns out science and business makes one helluva team.Waste carbon pollution is humanity's biggest threat. LanzaTech is turning it into an opportunity. Reducing emissions and making money for our customers — today. Let's transform our tired Lose-Lose climate debate into a Win-Win proposition. One that helps companies grow their revenue while helping the planet reach a post-pollution reality. LanzaTech — Welcome to the Post Pollution Future; ['chemical recycling', 'carbon emissions', 'air quality', 'circular economy', 'green chemistry', 'carbon recycling', 'climate action', 'fuels and chemicals']; ['#lanzatech', '#circulareconomy', '#carbonrecycling']
Classification: Other Carbon Removal, Sustainable Aviation Fuels, Chemical Industry
Text: RWE; RWE is leading the way to a green energy world. With an extensive investment and growth strategy, the company will expand its powerful, green generation capacity to 50 gigawatts internationally by 2030. RWE is investing €50 billion gross for this purpose in this decade. The portfolio is based on offshore and onshore wind, solar, hydrogen, batteries, biomass and gas. RWE Supply & Trading provides tailored energy solutions for large customers. RWE has locations in the attractive markets of Europe, North America and the Asia-Pacific region. The company is responsibly phasing out nuclear energy and coal. Government-mandated phaseout roadmaps have been defined for both of these energy sources. RWE employs around 19,000 people worldwide and has a clear target: to get to net zero by 2040. On its way there, the company has set itself ambitious targets for all activities that cause greenhouse gas emissions. The Science Based Targets initiative has confirmed that these emission reduction targets are in line with the Paris Agreement. Very much in the spirit of the company's purpose: Our energy for a sustainable life.; ['Power Generation', ' Renewable Energies', 'Gas Fleet', 'Energy Trading', 'Hydrogen']; ['#teamrwe']
Classification: Energy General
Text: Goldman Sachs; At Goldman Sachs, we believe progress is everyone's business. That's why we commit our people, capital and ideas to help our clients, shareholders and the communities we serve to grow.Founded in 1869, Goldman Sachs is a leading global investment banking, securities and investment management firm. Headquartered in New York, we maintain offices in all major financial centers around the world. More about our company can be found at www.goldmansachs.com. For insights on developments currently shaping markets, industries and the global economy, subscribe to BRIEFINGS, a weekly email from Goldman Sachs. Copy and paste this link into your browser to sign up: http://link.gs.com/Qxf3.com/social-media-disclosures; []; ['#mynamemyheritage', '#makethingspossible']
Classification: Other Sectors
Text: PPC Solar (Paradise Power Company); PPC SOLAR excels in the engineering, procurement and construction (EPC) of energy transition solutions including photovoltaic installation, EV charging infrastructure and energy storage solutions. With over 40-years of local presence PPC has built strong relationships within the region for unsurpassed value. Headquartered in Taos, New Mexico, with a branch office in Albuquerque we serve the Southern, Central and Northern New Mexico Regions along with Southern Colorado.; ['Photovoltaic', 'Solar', 'Renewable Energy', 'Solar Electric', 'PV ', 'EV Charging Stations', 'Energy Storage', 'battery based PV Systems', 'ChargePoint Partner', 'off-grid', 'Residential', 'Commercial', 'Utility', 'Operations& Maintenance', 'electric vehicle charging']; ['#newmexico', '#photovoltaics', '#evcharging']
Classification: Solar, Electric Vehicles, Battery
Text: PT Bukit Asam Tbk; PT Bukit Asam Tbk (IDX: PTBA) is a state-owned enterprise that operates in mining industry, particularly coal mining, located in Tanjung Enim, South Sumatra.
In line with the Company's vision to become a world-class energy company, Bukit Asam continues to be committed to increasing added value for the Company, shareholders, and all stakeholders of the Company.; ['Indonesia-State Owned Coal Mining Company', 'Coal Mining', 'Coal Trading', 'Energy']; ['#energi', '#sustainability', '#transformasidigital']
Classification: Coal
Text: Climeworks empowers people and companies to fight global warming by offering carbon dioxide removal as a service via direct air capture (DAC) technology. Climeworks' DAC facilities run exclusively on clean energy, and our modular CO_2 collectors can be stacked to build machines of any capacity.
Classification: Direct Air Capture
{*Text to Classify*}

Table S6: Prompt used for the organization type classifier for organizations.

You are a research assistant classifying organization profiles on LinkedIn. Classify the organization into one or multiple of the following categories: [Governmental organization, Non Governmental Organization, International Organization, Research Institute, University, AcceleratorIncubator, Bank, Venture Capital, Other Financing Organization, Consulting, Project Developer, Manufacturer, Mining, Utility, Service Provider, Lobby Group, Media Firm, Event Organizer, Other]
Use the label "Other" if the organization does not fit into any of the above categories.

Here are some exemplary classifications:
Text: Fraunhofer-Institut für Solare Energiesysteme ISE; The Fraunhofer Institute for Solar Energy Systems ISE in Freiburg, Germany is the largest solar research institute in Europe. With a staff of about 1400, we are committed to promoting a sustainable, economic, secure and socially just energy supply system based on renewable energy sources. We contribute to this through our main research areas of energy provision, energy distribution, energy storage and energy utilization. Through outstanding research results, successful industrial projects, spin-off companies and global collaborations, we are shaping the sustainable transformation of the energy system: https://www.ise.fraunhofer.de/en/publishing-notes.html; ['Photovoltaik', 'Energiesystemtechnik', 'Solarthermie', 'Gebäudeenergietechnik', 'Wasserstofftechnologien']; ['#energytransition', '#research', '#solar']
Classification: Research Institute
Text: Esken; We own two core assets; London Southend Airport and Stobart Energy.London Southend Airport is an award-winning airport serving London and the South East. It has ambitions plans to grow and the key ingredients in place to enable that growth. And; in Stobart Energy we have a transformational and maturing business which supplies fuel to biomass plants across the UK so that they can, in turn, create renewable energy.; ['Infrastructure and Support Services', 'Railway Maintenance', 'Biomass Energy', 'Aviation', 'Biomass'];
Classification: Service Organization
Text: LanzaTech; Where others see a dire choice, LanzaTech sees a trillion-dollar opportunity.The good news is after 15 years, north of a thousand patents, and millions of hours of pioneering scientific inquiry, LanzaTech has invented a technology big enough to meet the moment. One that transforms pollution into profit, and ensures that humans continue to prosper far into the post-pollution future.The science is state-of-the-art, but the idea is simple. We use nature to heal nature. First we capture carbon emissions. Then we feed them into bioreactors where trillions of carbon-hungry microbes go to work. These tiny dynamos eat the pollution and output valuable raw material commodities. (...). LanzaTech is turning it into an opportunity. Reducing emissions and making money for our customers — today. Let's transform our tired Lose-Lose climate debate into a Win-Win proposition. One that helps companies grow their revenue while helping the planet reach a post-pollution reality. LanzaTech — Welcome to the Post Pollution Future; ['chemical recycling', 'carbon emissions', 'air quality', 'circular economy', 'green chemistry', 'carbon recycling', 'climate action', 'fuels and chemicals']; ['#lanzatech', '#circulareconomy', '#carbonrecycling']
Classification: Manufacturer
Text: RWE; RWE is leading the way to a green energy world. With an extensive investment and growth strategy, the company will expand its powerful, green generation capacity to 50 gigawatts internationally by 2030. RWE is investing €50 billion gross for this purpose in this decade. The portfolio is based on offshore and onshore wind, solar, hydrogen, batteries, biomass and gas. RWE Supply & Trading provides tailored energy solutions for large customers. RWE has locations in the attractive markets of Europe, North America and the Asia-Pacific region. The company is responsibly phasing out nuclear energy and coal. Government-mandated phaseout roadmaps have been defined for both of these energy sources. RWE employs around 19,000 people worldwide and has a clear target: to get to net zero by 2040. On its way there, the company has set itself ambitious targets for all activities that cause greenhouse gas emissions. The Science Based Targets initiative has confirmed that these emission reduction targets are in line with the Paris Agreement. Very much in the spirit of the company's purpose: Our energy for a sustainable life.; ['Power Generation', ' Renewable Energies', 'Gas Fleet', 'Energy Trading', 'Hydrogen']; ['#teamrwe']
Classification: Utility
Text: Goldman Sachs; At Goldman Sachs, we believe progress is everyone's business. That's why we commit our people, capital and ideas to help our clients, shareholders and the communities we serve to grow. Founded in 1869, Goldman Sachs is a leading global investment banking, securities and investment management firm. Headquartered in New York, we maintain offices in all major financial centers around the world. More about our company can be found at www.goldmansachs.com. For insights on developments currently shaping markets, industries and the global economy, subscribe to BRIEFINGS, a weekly email from Goldman Sachs. Copy and paste this link into your browser to sign up: http://link.gs.com/Qxf3.com/social-media-disclosures; []; ['#mynamemyheritage', '#makethingspossible']
Classification: Bank
Text: PPC Solar (Paradise Power Company); PPC SOLAR excels in the engineering, procurement and construction (EPC) of energy transition solutions including photovoltaic installation, EV charging infrastructure and energy storage solutions. With over 40-years of local presence PPC has built strong relationships within the region for unsurpassed value. Headquartered in Taos, New Mexico, with a branch office in Albuquerque we serve the Southern, Central and Northern New Mexico Regions along with Southern Colorado.; ['Photovoltaic', 'Solar', 'Renewable Energy', 'Solar Electric', 'PV ', 'EV Charging Stations', 'Energy Storage', 'battery based PV Systems', 'ChargePoint Partner', 'off-grid', 'Residential', 'Commercial', 'Utility', 'Operations & Maintenance', 'electric vehicle charging']; ['#newmexico', '#photovoltaics', '#evcharging']
Classification: Manufacturer, Project Developer
Text: The ERA Foundation; The ERA Foundation is a non-profit organisation which supports engineering skills development and aims to bridge the gap between engineering research and its commercialisation.; ['Outreach', 'Policy', 'Supporting UK Manufacturing', 'Supporting Engineering in the UK']; ['#engineering', '#stemeducation', '#electech']
Classification: Non Governmental Organization
Text: Parkit Enterprise Inc.; Parkit Enterprise Inc. is engaged in the acquisition, optimization and asset management of income producing industrial properties in Canada and parking facilities across North America. The Company's shares are listed on TSX-V (Symbol: PKT) and on the OTC (Symbol: PKTEF).; ['Industrial Real Estate', 'Parking Assets']; ['#parkitenterprise']
Classification: Other Financing Organization
Text: PT Bukit Asam Tbk; PT Bukit Asam Tbk (IDX: PTBA) is a state-owned enterprise that operates in mining industry, particularly coal mining, located in Tanjung Enim, South Sumatra.
In line with the Company's vision to become a world-class energy company, Bukit Asam continues to be committed to increasing added value for the Company, shareholders, and all stakeholders of the Company.; ['Indonesia-State Owned Coal Mining Company', 'Coal Mining', 'Coal Trading', 'Energy']; ['#energi', '#sustainability', '#transformasidigital']
Classification: Governmental Organization, Mining
Text: The Economist; News and analysis with a global perspective. We're here to help you understand the world around you. To subscribe to The Economist go to: https://bit.ly/3OLFYsJ To make a request about your personal data visit: https://econ.st/2OZnq92; ['News', 'Analysis', 'Global Perspective']; ['#climatechange', '#sustainability', '#environment']
Classification: Media Firm
{*Text to Classify*}

Table S7: Performance of the relevance classifier for climate tech.

Classifier	**Precision**	**Recall**	**F1-score**	**Number of true positives**
non-relevant	0.99	0.98	0.99	538
relevant	0.95	0.98	0.96	186

Table S8: Performance of the collaboration classifier.

Classifier	Precision	Recall	F1-score	Number of true positives
Collaboration relevance				
non-relevant	0.97	0.97	0.97	69
relevant	0.94	0.94	0.94	31
Average	0.95	0.95	0.95	100
Technology classification				
Battery	1.00	1.00	1.00	9
Bioenergy with Carbon Capture and Story (BECCS)	1.00	1.00	1.00	6
Biogas	0.86	1.00	0.92	6
Biochar	1.00	1.00	1.00	6
Biofuels	1.00	0.88	0.93	8
Carbon Capture And Storage (CCS)	1.00	1.00	1.00	10
Carbon Direct Removal (Other Solutions)	1.00	0.83	0.91	6
Concentrated Solar	1.00	1.00	1.00	5
Direct Air Capture (DAC)	1.00	0.86	0.92	7
Electric Aviation	1.00	1.00	1.00	5
Electric Shipping	0.83	1.00	0.91	5
Electric Vehicles	1.00	0.86	0.92	7
E-Fuels	1.00	1.00	1.00	5
Fuel Cell Aviation	0.83	1.00	0.91	5
Geothermal	1.00	1.00	1.00	9
Green Hydrogen	0.95	0.88	0.91	24
Hydro Energy	1.00	1.00	1.00	5
Marine Energy	0.71	1.00	0.83	5
Nuclear Fission	1.00	1.00	1.00	7
Nuclear Fusion	1.00	1.00	1.00	5
Offshore Wind	1.00	1.00	1.00	6
Railway	0.71	1.00	0.83	5
Solar	0.72	1.00	0.84	13
Sustainable Aviation Fuels (SAFs)	1.00	0.80	0.89	5
Waste to Heat	1.00	0.83	0.91	5
Wind (Onshore)	0.92	0.92	0.92	13
Micro average	0.91	0.95	0.93	200
Macro average	0.93	0.96	0.94	200
Collaboration type classification				
R&D collaborations	1.00	0.89	0.94	9
Demonstrations & pilots	1.00	0.90	0.95	10
Commercialisation & product launches	1.00	1.00	1.00	5
Production & manufacturing	1.00	1.00	1.00	7
Offtake & future agreements	1.00	1.00	1.00	6
Adoption & deployments	0.93	1.00	0.96	13
Operations & maintenance	1.00	0.83	0.91	6
Grants	1.00	0.62	0.77	8
Equity Investments	0.83	1.00	0.91	5
Debt finance	1.00	1.00	1.00	4
Other unspecified finance	1.00	1.00	1.00	5
Spin-offs	1.00	1.00	1.00	5
Mergers & acquisitions	1.00	1.00	1.00	6
Joint ventures	1.00	1.00	1.00	8
Incubators & accelerator programs	1.00	1.00	1.00	5
Certification & approvals	1.00	1.00	1.00	5
Training	1.00	0.88	0.93	8
Micro average	0.97	0.93	0.95	116
Macro average	0.98	0.94	0.95	116

Table S9: Prompt used for the collaboration extraction and classification.

You are a research assistant classifying LinkedIn posts. Follow these steps:

1. **Determine if the post announces an innovation collaboration between the posting organization and other organizations (Yes/No).**

- Collaborations must explicitly involve a cooperation between entities for advancing a technology, OR a financial transaction, OR acquisitions, OR changes in ownership, OR any legally-binding partnership from one of the specified categories.

- Exclude joint events, panel discussions, or PR activities without specific innovation activities.

2. **If Yes, classify:**

- **Technologies:** Choose from the list: {"Biomass", "Biofuels", "Biogas", "Wind", "Offshore_Wind", "Solar", "Concentrated_Solar", "Waste_to_Heat", "Direct_Air_Capture", "Carbon_Capture_And_Storage", "Biochar", "BECCS", "Carbon_Direct_Removal", "Hydrogen" (including Fuel Cells for Green Energy), "Nuclear_Energy", "Nuclear_Fusion", "Hydro_Energy", "Geothermal", "Battery", "Electric_Vehicles", "Sustainable_Aviation_Fuels", "E_Fuels", "Marine_Energy" (tidal and wave electricity only!), "Heat_Pumps", "Railway", "Electric_Shipping", "Electric_Aviation", "Fuel_Cell_Aviation", "Other_Technology"}

- **Collaboration types:** Choose from the list: {"r_and_d_collaborations": "Joint efforts to conduct research and develop new technologies.", "demonstrations_and_pilots": "Testing and demonstrating developed technologies to prove feasibility. Includes the deployment of prototype and testing facilities.", "commercialisation_and_product_launches": "Announcements of first commercial deployments and new product launches.", "production_and_manufacturing": "Setting up or scaling facilities to manufacture technologies or products.", "offtake_agreements_and_futures": "Contracts for future sales or delivery of products.", "adoption_and_deployments": "Concrete deployments, construction of power plants, sales, or adoption of products and technologies in the market. This includes the construction of infrastructure or systems that enable the use of a technology.", "operations_and_maintenance": "Collaborations focused on running and maintaining systems or equipment.", "grants": "Grants must be mentioned explicitly.", "equity_investments": "Financial investments in exchange for ownership stakes in a company or project.", "loans": "Explicitly mentioned debt or loan financing provided for project financing, repaid with interest.", "other_unspecified_finance": "Any financial collaboration that does not fit the categories above.", "spin_offs": "New companies formed to commercialize a specific technology or idea.", "mergers_and_acquisitions": "Business consolidations where companies merge or are acquired.", "joint_ventures": "Only classify as a joint venture if the term 'joint venture' is explicitly mentioned in the announcement!", "incubators_and_accelerator_programs": "Programs supporting startups or early-stage ventures through mentorship and resources.", "certification_and_approvals": "Processes to obtain official recognition, regulatory approval, or standards compliance for technologies or products.", "training": "Training related to a technology provided from one organization to another."}

- **Core technology:** Whether the announcement is about the core technology or corresponding infrastructure, software and digital platforms, raw materials, recycling, or other complementary activities.

3. **If Yes, extract the relevant organizations** for the collaboration mentioned in the post from the list of referenced organizations.

Guidelines for Output:

- Use consistent formatting for responses. Only provide the classification. Do not repeat the post or add explanations.

- If multiple categories apply, include all relevant entries. In the innovation categories (r and d collaborations, demonstrations and pilots, commercialisation and product launches, production and manufacturing, sales agreements and futures), prioritize the most relevant.

- For financial collaborations, use other unspecified finance if the type of finance is unclear.

- Only classify technologies for tidal or wave electricity generation (and not shipping) as Marine Energy.

- Only classify airplanes and aviation projects with Fuel Cell Aviation.

- Do not change the format of the organization names in the referenced organizations list.

Examples:

Post: fraunhofer-gesellschaft: We support Emmvee Group as their solar technology partner in setting up a production facility to produce solar panels at scale; referenced_organizations=['fraunhofer-gesellschaft','Emmvee Group']

Answer: Yes: Technologies = ['Solar'], Collaboration_types = ['production_and_manufacturing'], Core_technology = ['core_technology'], Collaborating_organizations=['fraunhofer-gesellschaft','Emmvee Group']

Post: soato: Bill Nicholson presented a coating technology at CPhI. coating-society brix-technologies.

Answer: No: This is only a product announcement without specifying a collaboration.

Post: hydrogène-de-france: Milestones in Indonesia with u.s. international development finance corporation. Technical assistance to develop 22 hydrogen power plants. $1.5 billion potential investment; referenced_organizations=['hydrogène-de-france', 'us-idfc']

Answer: Yes: Technologies = ['Hydrogen'], Collaboration_types = ['other_unspecified_finance', 'offtake_agreements_and_futures'], Core_technology = ['core_technology'], Collaborating_organizations=['hydrogène-de-france', 'us-idfc']

Post: kengenkenya: KenGen Kenya has partnered with Nairobi Metropolitan Services (NMS) to develop a 45MW Waste to Heat Power Plant financed through a government grant; referenced_organizations=['kengenkenya', 'nairobi-metropolitan-services']

Answer: Yes: Technologies = ['Waste_to_Heat'], Collaboration_types = ['adoption_and_deployments','grants'], Core_technology = ['core_technology'], Collaborating_organizations=['kengenkenya', 'nairobi-metropolitan-services']

Post: vinci-concessions: Since 2021, we are partnering with Airbus to promote the use of hydrogen and accelerate the decarbonization of the aviation sector. In Japan, Kansai Airports and Airbus have recently signed a memorandum of understanding to explore the use of hydrogen at three airports we operate in the country; referenced_organizations=['vinci-concessions', 'airbus', 'kansai-airports']

Answer: Yes: Technologies = ['Hydrogen', 'Fuel_Cell_Aviation'], Collaboration_types = ['demonstrations_and_pilots'], Core_technology = ['core_technology'], Collaborating_organizations=['vinci-concessions', 'airbus', 'kansai-airports']

Post: volvo: Bosch at the World Business Council event highlighting joint efforts.

Answer: No: Participating in joint conferences and events does not count.

Post: faurencia: In partnership with our joint venture Symbio, we showcase tested prototypes for low and zero emission vehicles, including hydrogen and electric vehicles. Hydrogen News Hydrogen Daily; referenced_organizations=['faurencia', 'https://www.linkedin.com/company/symbio/', 'hydrogen-news', 'hydrogen-daily']

Answer: Yes: Technologies = ['Hydrogen', 'Electric_Vehicles'], Collaboration_types = ['joint_ventures', 'demonstrations_and_pilots'], Core_technology = ['core_technology'], Collaborating_organizations=['faurencia', 'https://www.linkedin.com/company/symbio/']

Post: Orbital Marine Power: The world's most powerful tidal turbine is installed, generating power into the UK grid. Made possible by European Commission and Horizon 2020 funding; referenced_organizations=['orbital-marine-power-ltd', 'european-commission', 'horizon-2020']

Answer: Yes: Technologies = ['Marine_Energy'], Collaboration_types = ['adoption_and_deployments', 'grants', 'equity_investments'], Core_technology = ['core_technology'], Collaborating_organizations=['orbital-marine-power-ltd', 'european-commission', 'horizon-2020']

Post: Amarenco: SolarPower Europe launches French version of the Agrisolar Best Practice Guidelines, with the support of French industry! The report explores real-world examples of Agrisolar projects in the EU. At Amarenco, we are extremely proud to be a part of this with SolarPower Europe; referenced_organizations=['SolarPower Europe', 'Amarenco']

Answer: No: This post is about the publication of a report.

Post: AgriVijay: Presenting Multi Utility Biogas Powered Farming Tool power launched in partnership with Sukoon Solutions by AgriVijay! Key benefits: Join the sustainable farming revolution today! Sukoon Solutions AIC-JKLU NITI GOVERNMENT OF INDIA UNICEF Confederation of Indian Industry SIDBI (Small Industries Development Bank of India) Social Alpha Omnivore; referenced_organizations=['AgriVijay', 'sukoon-solutions', 'unicef', 'social-alpha', 'omnivore-partners']

Answer: Yes: Technologies = ['Biogas'], Collaboration_types = ['commercialisation_and_product_launches'], Core_technology = ['core_technology'], Collaborating_organizations=['AgriVijay', 'sukoon-solutions']

Post: siemens-energy: Today, we announce our new partnerships: with ABB we will develop a new generation of electric vehicle charging technologies and with SolarEdge we will construct a new solar farm in the US; referenced_organizations=['siemens-energy', 'abb', 'solaredge']

Answer: Yes: [Technologies = ['Electric_Vehicles'], Collaboration_types = ['r_and_d_collaborations'], Core_technology = ['infrastructure'], Collaborating_organizations=['siemens-energy', 'abb']], [Technologies = ['Solar'], Collaboration_types = ['production_and_manufacturing'], Core_technology = ['core_technology'], Collaborating_organizations=['siemens-energy', 'solaredge']]

Post: grundon-waste-management-ltd: We're delighted to announce we have just invested around £200,000 in two state-of-the-art JCB electric Teletruk forklifts by JBC; referenced_organizations=['grundon-waste-management-ltd', 'jcb', 'pyroban']

Answer: Yes: Technologies = ['Electric_Vehicles'], Collaboration_types = ['adoption_and_deployments'], Core_technology = ['core_technology'], Collaborating_organizations=['grundon-waste-management-ltd', 'jcb', 'pyroban']

{*Text to Classify*}

Table S10: Overview of i3 companies matched with LinkedIn organizations by country. We excluded countries with less than 5 i3 companies.

Country	Share of matched companies	Number of i3 companies
United States	68%	4,237
United Kingdom	77%	1,083
China	82%	784
Germany	70%	666
France	77%	590
Canada	82%	561
India	78%	373
Netherlands	71%	340
Australia	87%	279
Switzerland	69%	267
Sweden	66%	261
Spain	77%	231
Israel	70%	211
Norway	82%	186
Italy	68%	166
Belgium	71%	140
Singapore	82%	132
Finland	72%	127
Denmark	70%	118
Japan	75%	99
Ireland	77%	98
South Africa	70%	88
Korea Republic	70%	81
Nigeria	81%	75
Austria	79%	75
Brazil	77%	65
Turkey	73%	63
Portugal	68%	60
Poland	73%	48
Taiwan	85%	48
Estonia	79%	43
Hong Kong	85%	41
Russia	44%	34
New Zealand	79%	33
Kenya	88%	32
Indonesia	65%	31
United Arab Emirates	71%	24
Thailand	84%	19
Lithuania	59%	17
Greece	75%	16
Malaysia	94%	16
Luxembourg	80%	15
Mexico	93%	15
Bulgaria	60%	15
Czech Republic	60%	15
Vietnam	71%	14
Morocco	57%	14
Slovenia	79%	14
Chile	77%	13
Iceland	83%	12
Latvia	67%	12
Colombia	64%	11
Hungary	80%	10
Slovakia	60%	10
Cambodia	70%	10
Romania	89%	9
Philippines	89%	9
Pakistan	89%	9
Uganda	88%	8
Argentina	50%	8
Rwanda	63%	8
Saudi Arabia	88%	8
Ghana	100%	7
Ukraine	71%	7
Cyprus	50%	6
Tanzania	80%	5
Croatia (Hrvatska)	100%	5

Table S11: Overview of company relationships from the i3 database matched with partnerships in our network, by relationship type. We included relationships between 2020 and 2024 and companies from the industries "energy & power" and "transportation & logistics".

Relationship Type	Share of matched relationships	Number of relationships from i3 (2020–2024)
Technology Partner	77%	131
Development Partner	85%	112
Project Development	82%	101
Joint Venture	74%	35
Customer	87%	30
Channel Partner	57%	21
Licensee	60%	10
Incubation/Acceleration	50%	2

Table S12: Overview of company relationships from the i3 database matched with partnerships in our network, by country. We included relationships between 2020 and 2024 and companies from the industries "energy & power" and "transportation & logistics".

Country	Share of matched relationships	Number of relationships from i3 (2020–2024)
United States	77%	333
Germany	86%	102
United Kingdom	81%	91
France	86%	56
Japan	76%	45
China	62%	39
Canada	79%	29
Sweden	88%	24
Norway	82%	22
Netherlands	69%	16
Australia	75%	16
Denmark	100%	13
Korea Republic	83%	12
Israel	90%	10
Italy	90%	10
India	100%	7
Switzerland	71%	7
Spain	86%	7
Austria	60%	5
Taiwan	100%	5
Greece	67%	3
Estonia	67%	3
Ireland	100%	3
Belgium	100%	2
Singapore	50%	2
Portugal	100%	2
Finland	0%	2
Hong Kong	0%	2
Latvia	50%	2
Vietnam	100%	2
New Zealand	100%	1
Slovakia	0%	1
Malaysia	100%	1
Egypt	100%	1
Lithuania	100%	1
Kenya	0%	1
Indonesia	100%	1