

Tracking Direct Air Capture Innovation through LLM-Driven Patent Analysis

Celine Körner^{1*}, Jan-Maris Göpfert^{1,2}, Jann Michael Weinand¹, Jochen Linßen¹, Mak Đukan¹

1 Forschungszentrum Jülich, Institute of Climate and Energy Systems – Jülich Systems Analysis, 52425 Jülich, Germany

2 RWTH Aachen University, Chair for Fuel Cells, Faculty of Mechanical Engineering, 52062 Aachen, Germany

*Corresponding Author: c.koerner@fz-juelich.de

ABSTRACT

Direct air capture (DAC) is a central pillar of pathways to net-zero and net-negative emissions. However, DAC has not converged on dominant designs yet. Furthermore, it remains uncertain which technical approaches in capture materials, unit engineering, or system integration will prevail. This uncertainty complicates decision-making for investors, policymakers, and technology developers. Although patents provide the most comprehensive public record of this technological evolution, their rapidly growing volume precludes manual expert analysis, while keyword- and classification-code-based analyses are insufficient for mapping patents to DAC's diverse technical approaches or capturing the technical context that distinguishes them.

Therefore, we developed a large language model (LLM) pipeline that analyzes patent families and characterizes the inventions they describe. Patents are retrieved from The Lens (<https://about.lens.org/>) using a high-recall query, which yields a broad set of patent families. An LLM then filters these patent families by relevance into core, peripheral, and non-DAC families. Guided by a human-validated, self-extending taxonomy, each family is mapped to technical concepts and assigned a primary invention focus, allowing inventive activity to be tracked. Applied to 2,326 DAC patent families published between 2010 and early 2026, the pipeline maps detailed concepts across different DAC technology aspects: capture material, unit and contactor engineering, CO₂ regeneration method, DAC system integration with low-carbon energy supply as well as linkage to CO₂ storage and utilization. This enables systematic tracing of DAC's technological evolution and scale-up over time.

Our analysis yields four key findings: First, DAC patenting accelerated exponentially through 2025, with inventive activity showing no sign of leveling off. Second, the focus of innovation appears to be shifting from individual

components to integrated units and systems. The share of patents on materials such as sorbents fell from more than half between 2015 and 2019 to roughly one-third by 2025, while unit engineering and system integration now account for the majority of patents. This shift may signal a pivot toward deployment. Third, distinct actor dynamics are evident. Leading startups pursue divergent technological pathways, with the most innovation concentrated in solid-sorbent DAC, whereas universities primarily advance experimental materials and processes at lab scale. The automotive and oil & gas industries lead patenting activity among large corporations, but they focus on refining technologies already established in the field, notably solid adsorption with temperature-swing regeneration, rather than exploring new materials or system designs. Fourth, patenting momentum and field deployment diverge. Two-thirds of patent families aim to use the captured CO₂ (mainly for the production of synthetic fuels), whereas operating DAC plants deliver CO₂ almost exclusively to geological storage.

The study demonstrates how LLMs can convert unstructured patent text into theory-grounded innovation indicators at full-corpus scale. It also surfaces which DAC technologies are positioned to scale fastest and integrate into future power systems with high renewable shares, supporting technology planning, investment, and system-integration decisions in the energy transition.

Keywords: direct air capture, carbon dioxide removal, patent analysis, large language models, technological innovation