Data-Driven Discovery of Sustainable Clinker Replacements

Building a Broader Portfolio of Lower-Carbon Concrete Solutions

MIT CSHub and Olivetti Group Research Brief
Soroush Mahjoubi*, Hessam AzariJafari, Randolph Kirchain, and Elsa Olivetti
*Corresponding Author, mahjoubi@mit.edu

Why find and screen clinker replacements?

Concrete is the world's most used manufactured material. Because it is produced on such a large scale, concrete is also associated with significant greenhouse gas (GHG) emissions. The majority of these emissions originate from the production of **clinker**, the primary constituent of cement, driven by CO_2 emissions from heating limestone and high-temperature (>1300 °C), usually fossil-fired kilns.

As a result, one of the immediate implementable levers to cut concrete's carbon intensity is replacing clinker with lower-carbon **reactive clinker replacements (RCRs)**. This kind of substitution is not new to the concrete industry. Fly ash from coal combustion and blast furnace slag from ironmaking have been utilized for this purpose for over a century, replacing 20–40% of clinker while maintaining strength, durability, and workability. However, global supplies of both materials are already used intensively. Expanding the use of RCRs to meet the industry's own climate goals will require developing a broader portfolio of candidates.

Step one: identify RCR candidates

The first step in finding promising lower-carbon RCRs was to sift through the millions of known compounds to identify materials that held some promise for replacing clinker. Reviewing the scientific literature is a critical step in any research project; however, a truly comprehensive search required tools to mine the literature at a scale previously not possible. Fine-tuned large language models scoured over 10,000 academic papers to extract chemical profiles and material classifications relevant to performance. This process revealed more than 14,000 prospective RCRs, categorized into nineteen types, ranging from ashes to construction and demolition waste to natural pozzolans (Figure 1a on the next page). The mined data provided the foundation for predicting which materials could serve as effective RCRs. Figure 1b on the next page presents box-and-whisker plots of the chemical compositions extracted from the literature for each material type. The results highlight substantial variability across materials, which directly affects their reactivity potential. Materials such as silica fume show narrow compositional ranges, indicating consistent chemistry. In contrast, biomass ashes, construction-and-demolition (C&D) wastes, and natural pozzolans exhibit wide distributions, reflecting heterogeneous sources and processing conditions that can lead to large variations in reactivity.

Step two: Estimate reactivity

A central obstacle in finding new RCRs is understanding how they will react when combined with water in a concrete mixture. **Reactivity** is typically measured by laboratory tests that monitor **heat release**, **calcium hydroxide consumption**, and the **amount of water bound in hydration products** [1]. Each test can take weeks to perform, making large-scale screening impractical. To overcome this, the team developed a machine learning model trained on experimental data from more than 300 materials. The model predicted all three reactivity metrics based on chemical composition, particle size, specific gravity, and amorphous content. In validation tests, predictions closely matched experiments.

Key Takeaways:

- An AI-enabled framework screened over 10,000 academic papers to identify more than 14,000 prospective reactive clinker replacements (RCRs).
- A machine learning model evaluated the suitability of candidate materials based on their reactivity metrics.
- Materials like construction by-products, biomass ashes, and various slags emerged as viable candidates.
- Over one million rock samples from around the world were screened for candidate materials; potential natural RCRs can be found in many different geographies.

Cement-based products like concrete and mortar are essential for resilient, high-performance buildings and infrastructure. However, as pictured here, producing cement typically requires intensely heating raw materials in kilns by combusting fuels; this process is responsible for the majority of emissions from cement production. In efforts to reduce emissions, demand for traditional reactive clinker replacements (RCRs) is outpacing supply. This research brief outlines how and why identifying alternative RCRs can help meet climate goals while expanding a circular economy.

Image Source: Adobe Stock.

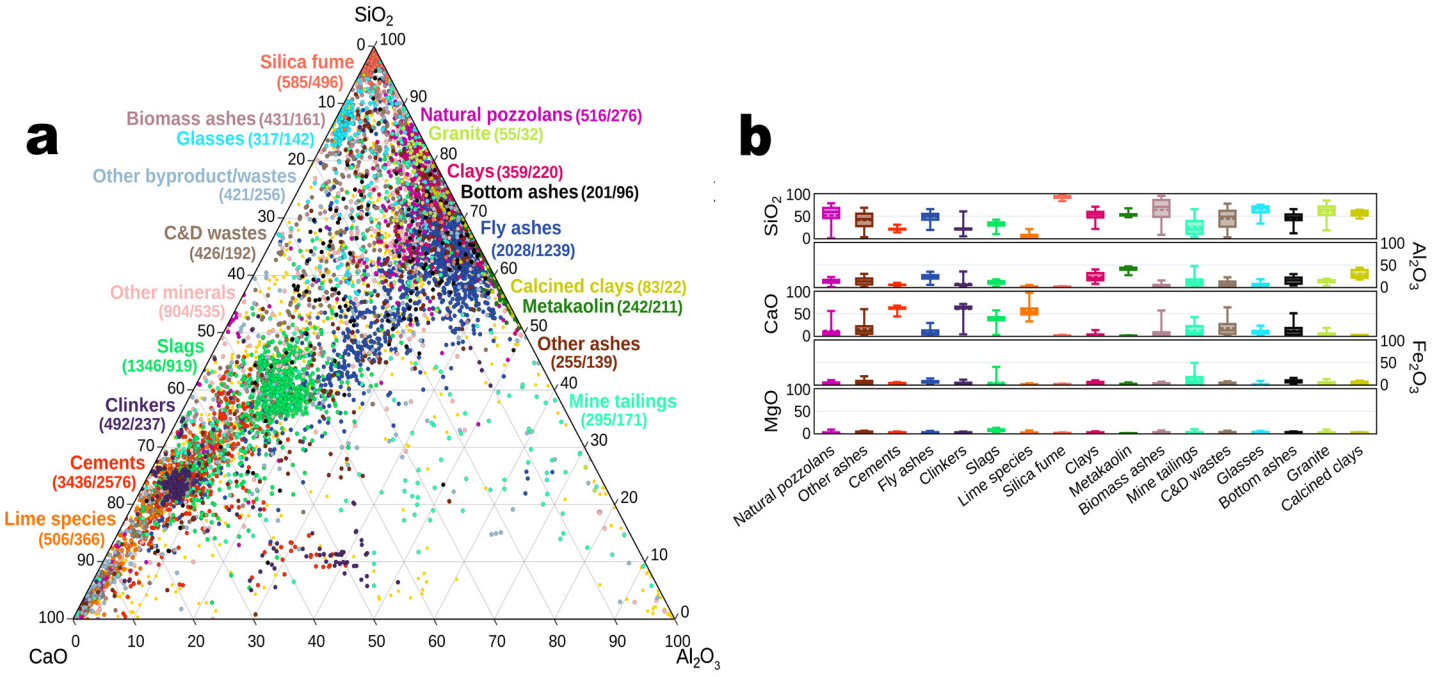

Figure 1: ***a)*** *CaO-Al_2O_3-SiO_2 ternary diagram of the extracted materials (CaO-Al_2O_3-SiO_2>80%wt) in the present work, colored by material types. Below/beside each category label, the numbers represent the ratio of samples to unique papers, indicated as samples/papers. Yellow markers indicate materials with an unspecified category.* ***b)*** *Box-and-whisker plot of oxide compositions of material types; upper and lower whiskers represent the 95% range, while the boxes indicate the interquartile range (IQR), with the mean and median shown as straight and dashed lines; metakaolin, clays and calcined clay, while types of natural pozzolan, were represented separately in the plots. [2]*

A coefficient of determination [R^2] of over 0.85 was observed across all three metrics. Figure 2 on the right illustrates the reactivity space of all analyzed materials, with highly reactive hydraulic materials positioned in the upper left and less reactive pozzolanic materials in the lower right. A large share of natural materials fall within the reactive domain. The figure highlights substantial reactivity variation among the top rock types that were statistically more reactive and contributed the highest number of reactive samples. Most reactive samples lie in the pozzolanic region, characterized by $Ca(OH)_2$ consumption greater than 50 g per 100 g. In total, ~46,700 samples were classified as pozzolanic and ~3,800 as hydraulic. Overall, the results show that the potential for high reactivity differs markedly across rock types.

Step three: Map out priority replacement candidates

With this predictive tool in hand, researchers could screen both industrial by-products and natural resources at a scale that had never been possible before. Construction and demolition wastes, incineration ash from biomass, and various types of slags emerged as particularly promising by-products. Together, they could replace up to 40% of global clinker production. Such a shift would reduce annual greenhouse gas emissions by roughly 1.2 billion tons, equivalent to removing 260 million cars from U.S. roads. These materials not only offer a lower-carbon alternative to clinker but also create opportunities for transforming waste streams into valuable construction inputs.

The second frontier explored in this work was natural rock formations. Over one million rock samples from around the world were screened using global geochemical data. Twenty volcanic

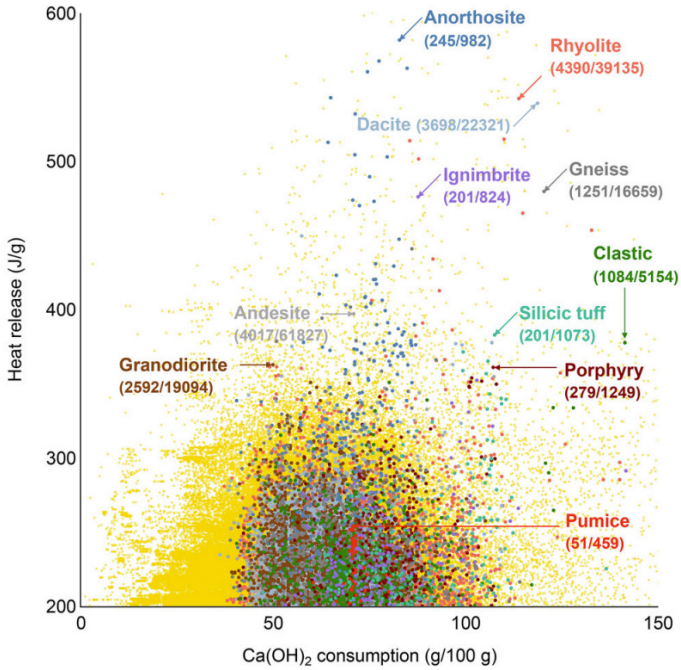

Figure 2: *Reactivity variations of selected rock types, chosen based on highest abundance or reactive-to-total ratio, colored by rock types. Below each label, the numbers represent the ratio of moderate-to-high reactive materials to the total number of rock samples of that type, indicated as reactive samples/total samples. Each arrow points to the sample with the highest degree of reactivity. Yellow markers denote rock samples whose rock types were not selected. [2]*

and sedimentary rock types showed consistent promise, including ignimbrites, silicic tuffs, pumice, shales, and rhyolites. Many of these rocks are found to be prevalent in tectonically active regions such as the Andes, the Great Rift Valley, and the Appalachians [2]. This means that natural RCRs could serve as locally available substitutes in areas where industrial by-products are scarce, such as along the U.S. West Coast. Natural rocks can serve as viable RCRs, especially in regions with limited access to industrial by-products or waste-derived RCRs. Figures 3a and 3b show that these naturally occurring RCRs are not randomly distributed but form distinct geological clusters. Significant concentrations appear across northern and central Europe (Baltic Shield, Scafell Pike), Asia (Japan, Taurus and Zagros Mountains, Indian cratons, Tibetan Plateau), Africa (Great Rift Valley), Oceania (Great Dividing Range and ancient cratons), and South America (Andes). In the U.S., the Rocky Mountains and Cascade Volcanic Arc in the West, as well as the Appalachian along the East Coast, seem to have particularly promising availability of reactive rocks (Figures 3c, 3d).

Summary of Key Messages:

- A century of concrete research has identified many candidate materials to partially replace clinker.
- Many candidate materials are likely to have sufficient reactivity to be commercially viable.
- Potential sources of promising RCRs are located throughout the world.

Significance and next steps

The information revealed through this research could reshape the cement industry's material supply chain. In the future, cement and concrete producers may have access to a broad and diversified pool of raw materials. This diversification is essential to achieving the industry's target of reducing the global clinker-to-cement ratio from 76 percent today to 52 percent in 2050 [3]. Moreover, the methodology itself – combining artificial intelligence, large-scale data mining and predictive modeling – offers a powerful template in accelerating discovery in other material-intensive sectors. A graphical overview of the methodology is included as Figure 4 (see next page), with examples of identified RCRs in Figure 5 (see next page). Such analyses can be built rapidly on the toolkit developed for this research by first identifying information (here, composition and reactivity) and the scope of application (here, concrete).

There remain essential steps before many of these RCRs can be deployed at scale. Fresh-state workability, hardened performance, long-term durability, and environmental impacts such as leaching must be tested and validated. In some cases, activation methods may be needed to unlock the full potential of less reactive materials. But the foundation is clear: A vast landscape of low-carbon clinker replacements exists, and it is ready to be systematically uncovered and assessed. By marrying computational power with the pressing challenge of decarbonization, this work points toward a more sustainable future for the world's most important building material.

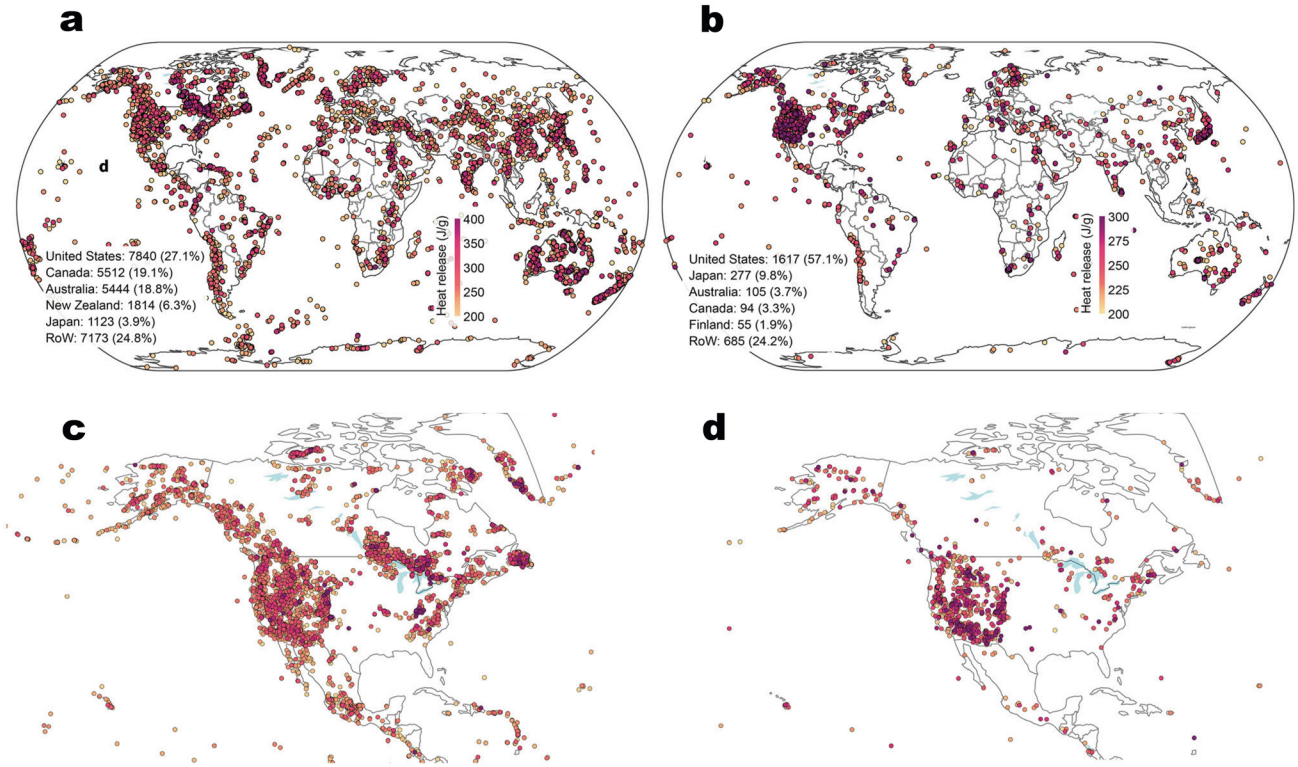

Figure 3: *Geospatial distribution of identified reactive rocks within the globe with **a)** pozzolanic and **b)** hydraulic behavior. The number preceding each country indicates the number of unique locations with cementitious precursors, and the number in parentheses represents the percentage relative to the total identified locations. Geospatial distribution of discovered reactive rocks within the Northern America with **a)** pozzolanic and **d)** hydraulic behavior.*

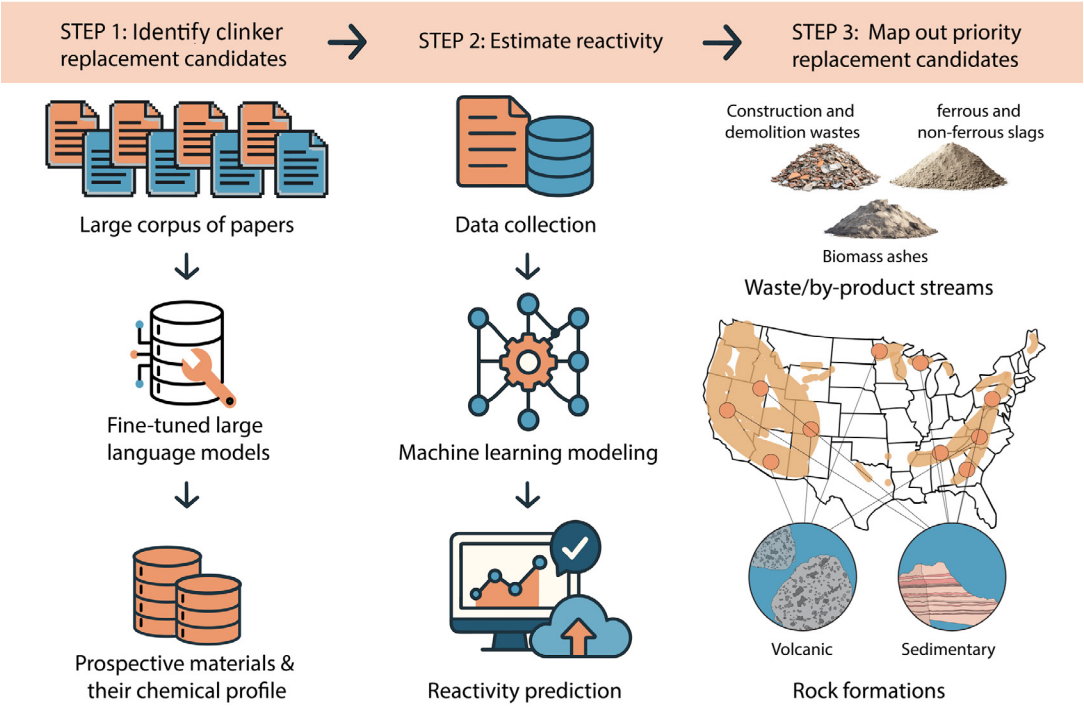

Figure 4: *The methodology behind this research. RCRs are identified by screening the academic literature with AI tools, their reactivity is then estimated with machine learning models, and priority replacement candidates (certain rock formations and by-product streams) are mapped.*

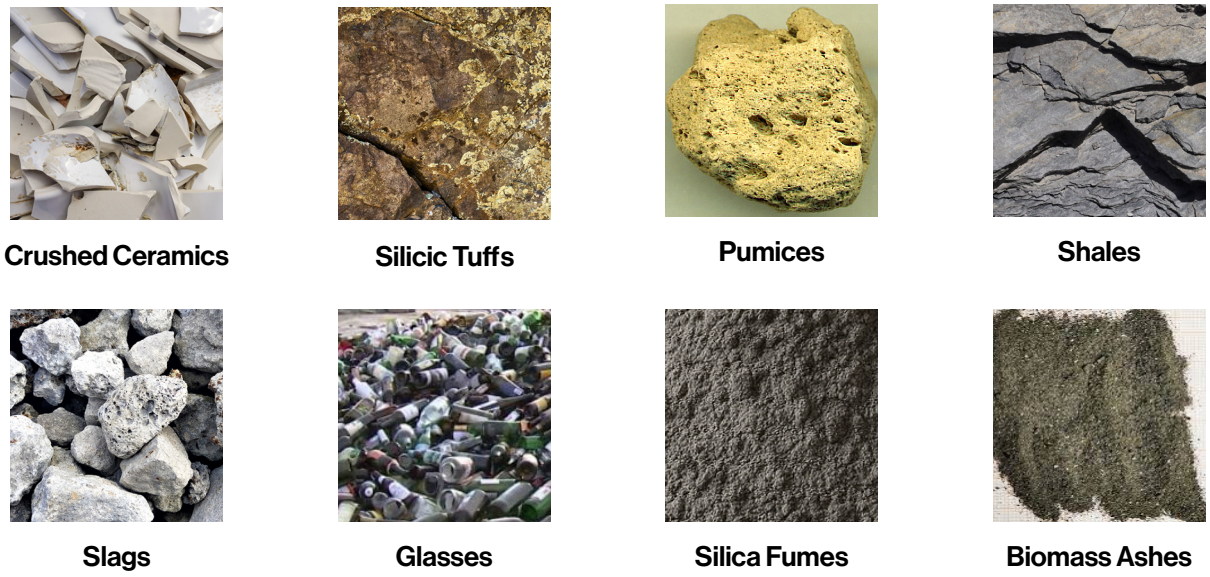

Figure 5: *Examples of RCRs identified and screened by this research.*

Related Links

[i] MIT News, June 2025: "AI stirs up the recipe for concrete in MIT study"

[ii] Nature Communications Chemistry Research Highlight, July 2025: "Towards net zero by data-driven discovery of sustainable cement alternatives"

Citation

Mahjoubi, S., AzariJafari, H., Kirchain, R., and Olivetti, E. (2026). Data-Driven Discovery of Sustainable Clinker Replacements: Building a Broader Portfolio of Lower-Carbon Concrete Solutions. Research Brief, MIT CSHub and Olivetti Group, Volume 2026.

References

[1] ASTM International. (2024). ASTM C1897: Standard Test Methods for Measuring the Reactivity of Supplementary Cementitious Material by Isothermal Calorimetry and Bound Water Measurements. https://doi.org/10.1520/C1897-20.

[2] Mahjoubi, S., Venugopal, V., Manav, I.B., AzariJafari, H., Kirchain, R.E., Olivetti, E.A., 2025. Data-driven material screening of secondary and natural cementitious precursors. Commun. Mater. 6, 99. https://doi.org/10.1038/s43246-025-00820-4

[3] Global Cement and Concrete Association (GCCA), 2021. GCCA 2050 Roadmap to Net Zero Concrete.