Supplementary Notes

1. Identification of AlphaFold-related publications

This section provides additional detail on the construction of the AlphaFold-related corpus, complementing the overview in Methods (M1–M1b). To ensure comprehensive coverage of AlphaFold-related research while minimizing false negatives in metadata-based retrieval, we employed a two-stage identification strategy within the 2019–2025 life-science baseline in OpenAlex.

First, we applied a keyword-based search across title, abstract, venue, and concept fields using a curated set of AlphaFold-related terms (Supplementary Table S1). While this approach captures explicitly labeled AlphaFold studies, it may miss publications that meaningfully use or evaluate AlphaFold without consistently referencing it in searchable metadata.

To address this limitation, we implemented a citation-based expansion step, adding all baseline publications that cite at least one of four foundational AlphaFold outputs: the original AlphaFold method, AlphaFold2, the AlphaFold Protein Structure Database, and AlphaFold 3 (Supplementary Table S2). This step is motivated by the observation that methodological dependencies in computational biology are often reflected through citation rather than explicit textual mention, particularly in downstream applications. Publications identified through either channel were merged and deduplicated using OpenAlex work identifiers.

2. Robustness Analyses

2.1 Network centrality

To examine whether our findings on unequal positioning in the AlphaFold collaboration network depend on a specific centrality metric, we compared four commonly used country-level measures: weighted degree, betweenness, eigenvector, and closeness centrality (Supplementary Fig. S1). All measures were calculated on the undirected, weighted co-authorship network, and their distributions were compared across high-, middle- and low-R&D-strength groups.

A consistent pattern emerges across all four measures. High-R&D-strength countries tend to occupy the most central positions in the network, middle-strength countries generally fall in between, and low-strength countries remain at the periphery. This monotonic stratification appears regardless of how centrality is defined, suggesting that countries with stronger pre-existing R&D capacity are systematically better within the global AlphaFold collaboration system.

These differences are also statistically robust. Kruskal-Wallis tests reject the null hypothesis of equal centrality distributions across R&D-strength groups for weighted degree, betweenness, eigenvector and closeness centrality (all $p < 10^{-7}$; Supplementary Fig. S1). Median weighted degree, eigenvector and closeness centrality decrease stepwise from high- to middle- to low-strength countries. Betweenness centrality shows the same overall tendency, although several high-strength hub countries exhibit particularly large values.

Overall, the results indicate that the core-periphery structure described in the main text is not driven by a single network metric. High-R&D-strength countries are consistently more central in the global AlphaFold collaboration network, both in terms of the scale of their collaborations and their structural position within the network.

2.2 AI4S capacity

To assess whether our findings about cross-country AI4S capacity depend on the specific construction of the composite index, we compare the equal-weight composite used in the main analysis with a PCA-based alternative. The PCA-based index is defined as the first principal component extracted from a principal component analysis of the three standardized AI4S dimensions – adoption speed, production output, and network influence – estimated on the common 77-country sample.

As shown in Supplementary Fig. S2a, the equal-weight composite and the PCA-based index are almost perfectly aligned. Country-level scores lie tightly along the 45-degree line with Pearson's $r = 0.999$ and Spearman's $rho = 0.999$. The first principal component also explains 84% of the total variance across the three dimensions, indicating that the equal-weight composite closely captures the dominant underlying structure in the data.

We further examine whether countries occupy similar relative positions when AI4S capacity discretized into deciles. Supplementary Fig. S2b reports a decile-by-decile across-tabulation of country rankings under the two indices. The exact decile match rate is 97.4%, and the median absolute difference in decile rank is 0, indicating that countries – and therefore the high-, middle-, and low-capacity tiers highlighted in the main text – are classified almost identically under the equal-weight and PCA-based approaches.

These results suggest that the main empirical patterns reported in the paper are not sensitive to whether AI4S is measured using a simple equal-weight composite or a statistically derived PCA-based index.

2.3 Threshold effects of network influence

In the main text, we show that countries with stronger network positions tend to move faster and further in AI4S. Supplementary Fig. S3 examines whether this relationship is linear, or whether returns to network influence accelerate beyond a specific threshold.

To investigate this, we estimate piecewise linear models that allow for a structural break at an endogenous breakpoint and compare them with a single-slope linear specification and a flexible LOWESS smooth. For adoption speed (Supplementary Fig. S3a, left), the threshold model captures a clear non-linearity: below the estimated breakpoint, adoption increases only gradually with network influence, whereas above a score of approximately 0.72, the slope increases sharply and closely tracks the non-parametric trend. This indicates that early gains in network position translate into modest improvements in adoption, while higher levels of centrality are associated with substantially faster uptake.

A similar pattern is observed for production scale and impact. In Supplementary Fig. S3b (left), the estimated breakpoint occurs at a lower value of approximately 0.51, but the piecewise specification again captures a pronounced upturn among the most central countries that a linear model fails to represent.

To assess sensitivity to the choice of breakpoint, the right-hand panels report how the estimated post-threshold slope varies as the breakpoint is shifted within ±1 standard deviation of the preferred estimate. The slope remains consistently higher than the pre-threshold gradient and is relatively stable around the optimal breakpoint, indicating that the observed non-linearity is not driven by a specific threshold choice.

Therefore, these results suggest non-linear "take-off" dynamics in AI4S: once countries reach sufficiently central positions in global collaboration networks, further gains in influence are associated with disproportionately large increases in both adoption speed and production outcomes.

2.4 Using World Bank income groups

To assess whether our main findings depend on grouping countries by R&D strength, we re-estimate the adoption and output analyses using the World Bank income classification. Supplementary Fig. S4a shows the cumulative share of AlphaFold adopters within each income group over time, relative to the initial release at CASP13 in December 2018. High-income countries adopt earliest and most extensively, followed by upper-middle-income countries, while lower-middle- and low-income countries adopt later and remain substantially behind throughout the observation period.

These differences are reflected in adoption lags (Supplementary Figure. S4b). The median delay is shortest in high-income countries and increases progressively across income groups, reaching approximately 21 months longer in low-income countries, with upper-middle- and lower-middle-income groups occupying intermediate positions.

Supplementary Fig. S4c and S4d examine how these adoption patterns translate into research output. Panel c shows cumulative AlphaFold-related publications by income group, indicating that high-income countries maintain a persistent lead throughout the period despite partial catch-up from upper-middle-income countries following the release of AlphaFold2 and the AlphaFold Protein Structure Database. Panel d further links adoption timing to post-adoption output at the country level, showing a strong negative association between adoption lag and subsequent AlphaFold-related publications: earlier adopters tend to generate substantially more output.

These results indicate that the temporal and production inequalities documented in the main analysis are robust to alternative country groupings and persist when countries are stratified by World Bank income levels.

Supplementary Figures

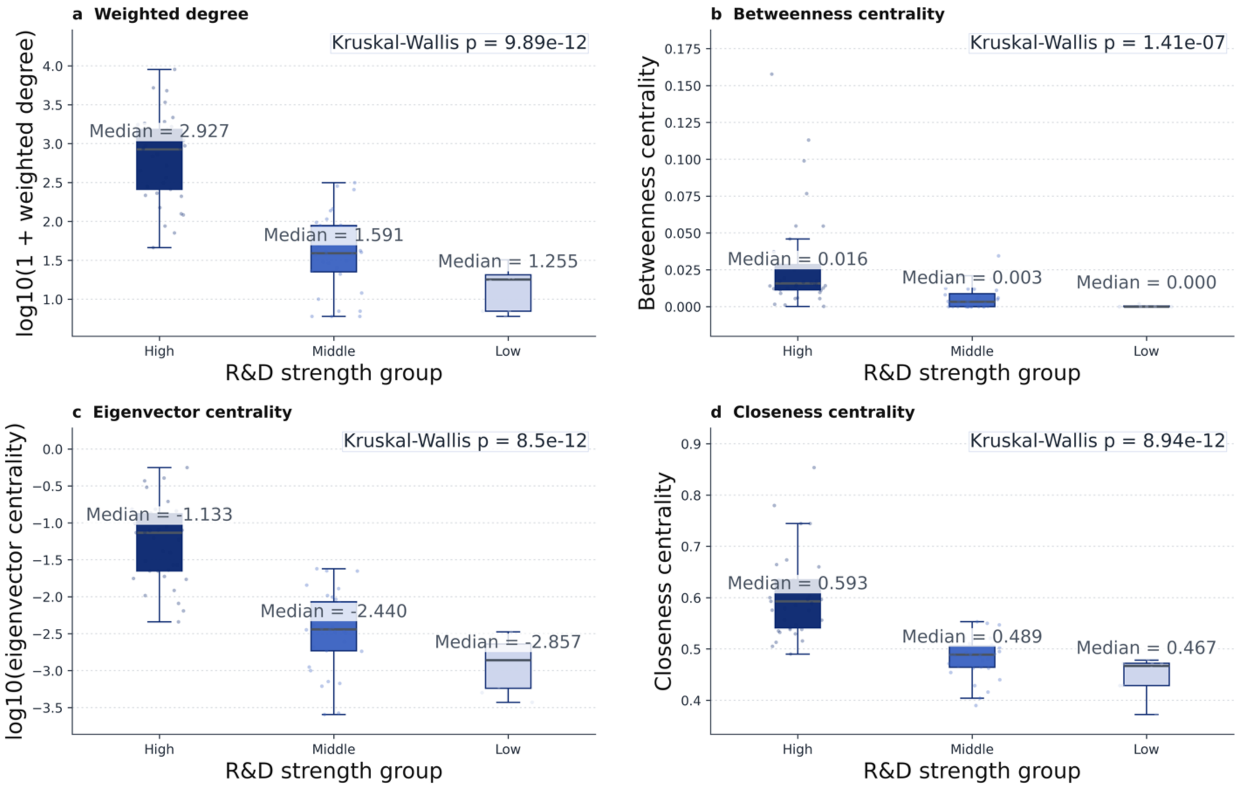

Figure S1 | Alternative network centrality measures by R&D group. Country-level distributions of four centrality measures in the AlphaFold collaboration network, comparing high-, middle- and low-R&D-strength groups. Panels show weighted degree (a), betweenness (b), eigenvector (c) and closeness (d) centrality, with values transformed to log scale where appropriate and presented as boxplots overlaid with individual country points. Median centrality scores decline monotonically from high- to middle- to low-strength groups for all four measures, and Kruskal-Wallis tests reject the null hypothesis of equal distributions across groups in every case ($p < 10^{-7}$). These results indicate that the core-periphery structure highlighted in the main text is robust to the choice of centrality metric.

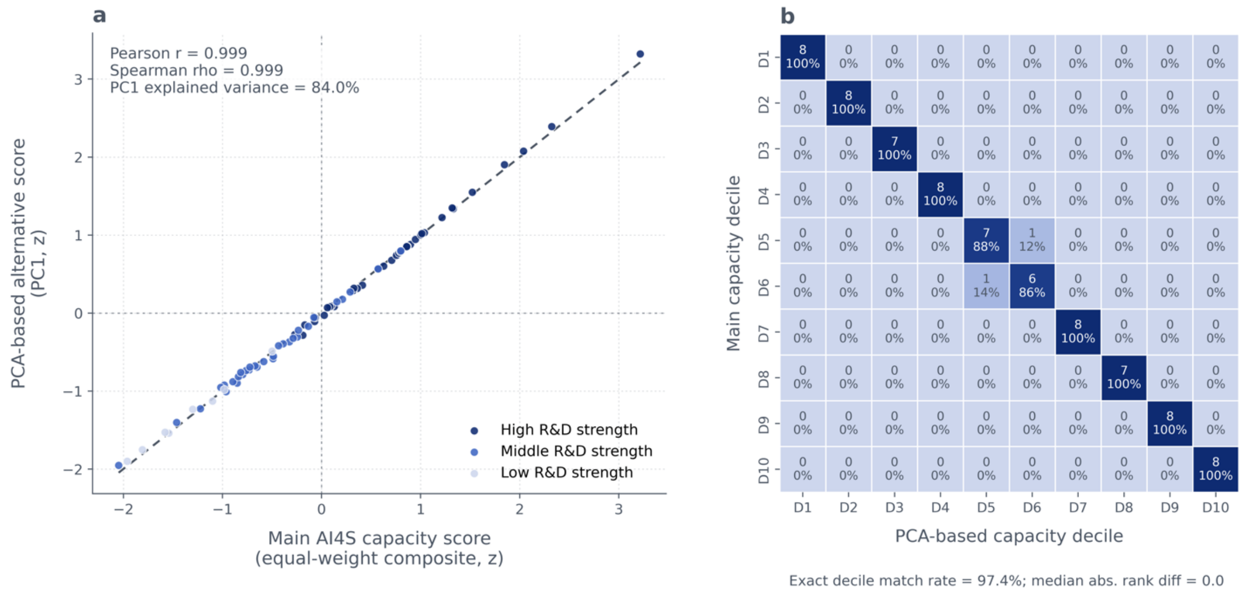

Figure S2 | Robustness of AI4S capacity scores to PCA-based alternatives. **(a)** Country-level scatterplot comparing the composite AI4S capacity score used in the main analysis with the first principal component from a PCA of adoption speed, production output, and network influence, showing a near-linear relationship and high rank-order concordance. The main equal-weight construction and the PCA-based alternative recover very similar country ordering. **(b)** Decile-by-decile comparison of country classifications based on the main AI4S capacity score versus the PCA-based index, indicating that countries— and thus high-, middle-, and low-capacity tiers—are assigned almost identically under both constructions (exact decile match rate 97.4%; median absolute rank difference 0). Both panels are based on the same 77-country intersection sample.

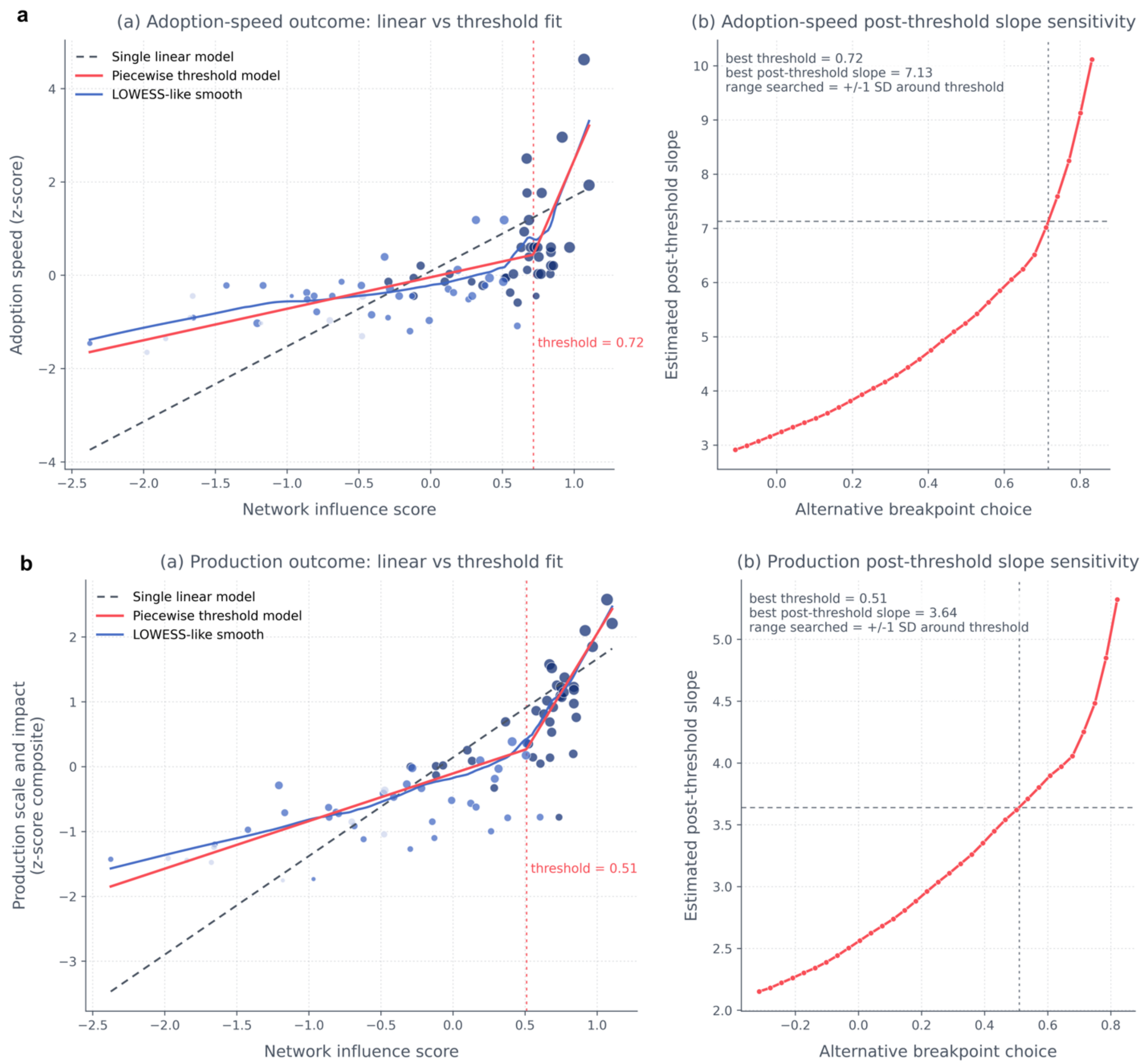

Figure S3 | Threshold model fit and robustness around the breakpoint. (a) Adoption-speed outcome: relationship between network influence and adoption speed, with a single linear model, a piecewise linear threshold model, and a LOWESS-like smooth overlaid on country-level scatterplots (left) and estimates of the post-threshold slope as the breakpoint is shifted within ±1 standard deviation of the estimated threshold (right). **(b)** Production outcome: analogous fits and post-threshold slope-sensitivity curves for production scale and impact, showing that in both cases the threshold specification better captures the sharp upturn at high influence levels and that the conclusion of much steeper returns above the threshold is robust to reasonable variation in the breakpoint choice.

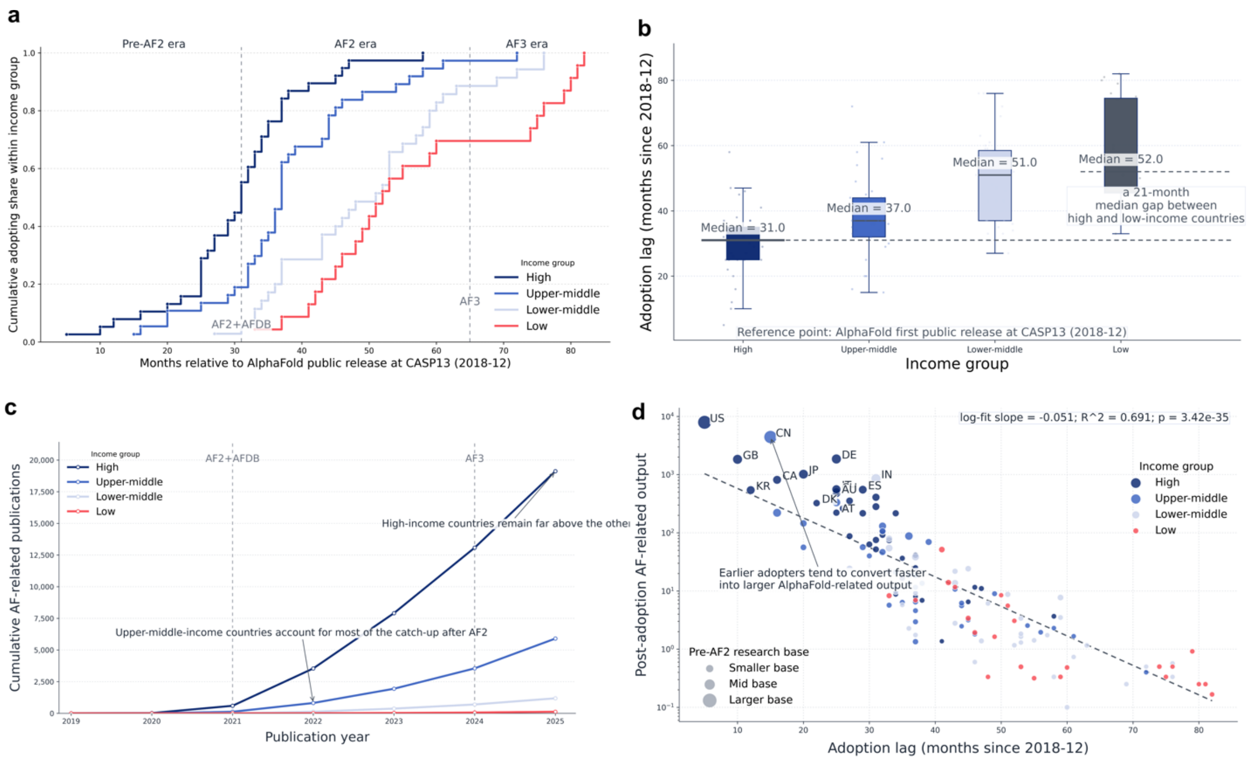

Figure S4 | Robustness using World Bank income groups. (a) Cumulative AlphaFold adoption within each World Bank income group across the pre-AF2, AF2 and AF3 eras. **(b)** Distribution of adoption lags, showing longer delays (with a median lag of 21 months) for lower-income groups. **(c)** Cumulative AlphaFold-related publications by income group over time. **(d)** Relationship between adoption lag and post-adoption AlphaFold-related output, indicating that earlier adopters, particularly high-income countries, convert faster into larger research outputs.

Supplementary Tables

Table 1 | Keyword-based identification of AlphaFold-related publications.

Field searched	Query term/pattern	Rationale/notes
Title, abstract, primary topic, concept labels	"AlphaFold" /"AF"	Core brand name of the system; captures most direct uses.
Title, abstract	"AlphaFold2"/"AlphaFold 2"/"AF2"	Explicit references to the second-generation model.
Title, abstract	"AlphaFold3"/"AlphaFold 3"/"AF3"	Explicit references to the third-generation model.
Title, abstract	"AlphaFold-Multimer"	Multimer variant.
Full text/extended metadata	"AlphaFold Protein Structure Database"/"AlphaFold DB"	References to the protein structure database.
Excluded patterns	"AF" without "AlphaFold"	Too ambiguous, not used as a standalone query term.

Note: Query patterns combine brand-level terms (“AlphaFold”) with model- and component-specific keywords (AF2, AF3, Multimer, and AlphaFold DB) to capture both direct mentions of the system and references to its major releases, while excluding ambiguous uses of “AF” that are unrelated to AlphaFold.

Table 2 | Landmark AlphaFold papers used for citation-based expansion.

ID	AlphaFold component	Full reference (short form)	Year	Journal	Notes
L1	AlphaFold (CASP13)	Senior et al., "Improved protein structure prediction using potentials from deep learning"	2020	Nature	Original AlphaFold CASP13 model.
L2	AlphaFold2	Jumper et al., "Highly accurate protein structure prediction with AlphaFold"	2021	Nature	AF2 model description.
L3	AlphaFold DB	Varadi et al., "AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models"	2022	Nucleic Acids Research	Protein structure database.
L4	AlphaFold3	Abramson et al., "Accurate structure prediction of biomolecular interactions with AlphaFold3"	2024	Nature	AF3 model and interactions

Note: The four landmark papers are used as seed references for citation-based expansion of the AlphaFold corpus, anchoring the first-generation CASP13 system, the AF2 model, the AlphaFold Protein Structure Database, and the AF3 model, respectively, so that downstream analyses cover both model development and database- or application-driven research.