



Received: 03 May 2026

Revised: 11 Aug. 2026

Accepted: 15 Aug. 2026

e-ISSN: 3085-7813

Corresponding Author: Zhyldyz Toktobekovna Dzhylykchieva –
E-mail: zhyldyz.dzhylykchieva@knu.kg

Editor in chief: Altieres de Oliveira Silva
– Alumni.In Editors

How to cite this article: Xia, W., Cong, W., Dzhylykchieva, Z. T., & Abduraimovna, S. B. (2027). Competitive Intelligence Capability and Strategic Innovation Performance: Evidence from Global AI Patent and Knowledge Transfer Networks. *Journal of Sustainable Competitive Intelligence*, 17, e0497. <https://doi.org/10.37497/eagleSustainable.v17i.497>

Competitive Intelligence Capability and Strategic Innovation Performance: Evidence from Global AI Patent and Knowledge Transfer Networks

Capacidade de Inteligência Competitiva e Desempenho da Inovação Estratégica: Evidências de Redes Globais de Patentes de Inteligência Artificial e Transferência de Conhecimento

Wenchun Xia

Kyrgyz National University named after Jusup Balasagyn, Bishkek, 720033, Kyrgyz Republic.

Wang Cong

Kyrgyz National University named after Jusup Balasagyn, Bishkek, 720033, Kyrgyz Republic.

Zhyldyz Toktobekovna Dzhylykchieva

Institute of Management and Business, Kyrgyz National University named after Jusup Balasagyn, Bishkek, Kyrgyz Republic.

Satyvaldieva Baktygul Abduraimovna

Kyrgyz National University named after Jusup Balasagyn, Bishkek, Kyrgyz Republic.

ABSTRACT | Purpose: This study examines how competitive intelligence (CI) capability supports strategic innovation performance in frontier artificial intelligence industries by transforming patent, publication, and semantic knowledge-transfer signals into actionable intelligence for strategic decision-making. **Methodology/Approach:** A quantitative secondary-data design was applied to the DeepInnovationAI dataset, comprising 2.3 million patents and 3.5 million scientific publications covering the period 1960–2020, with the main analysis focused on 987 country-year observations from 47 countries during 2000–2020. Multiple linear regression, mediation analysis, Pearson correlation, temporal trend decomposition, and k-means clustering were used to evaluate how Patent Novelty Index, Knowledge Transfer Rate, R&D Intensity, temporal velocity, and publication volume contribute to Competitive Intelligence Score. The empirical model is repositioned as an open-source technology intelligence framework that maps innovation signals onto the CI cycle of environmental scanning, signal analysis, dissemination, decision use, and feedback. **Findings:** Patent Novelty Index is the strongest predictor of Competitive Intelligence Score ($\beta = 0.41$, $p < 0.001$), followed by Knowledge Transfer Rate ($\beta = 0.33$, $p < 0.001$). The regression model explains 64% of the variance in Competitive Intelligence Score. Knowledge Transfer Rate partially mediates the relationship between R&D Intensity and CI Score, with an indirect effect of $\beta = 0.11$. Cluster analysis identifies three national CI profiles: Established Leaders, Emerging Challengers, and Peripheral Contributors. **Originality/Relevance:** The study contributes to the Journal of Sustainable Competitive Intelligence by moving beyond descriptive patent and bibliometric analysis and operationalizing CI as a strategic

capability. It incorporates CI architecture, intelligence governance, intelligence dissemination, strategic foresight, and decision-support logic into the interpretation of AI innovation data. This framing is aligned with recent journal contributions that conceptualize AI-enabled CI as a strategic capability and emphasize CI architecture as a foundation for strategic decision-making. **Theoretical/methodological contributions:** The study develops a measurable CI-cycle framework for frontier technology analysis by linking patent novelty, knowledge transfer, temporal velocity, and geographic clustering to intelligence planning, environmental scanning, analysis, dissemination, decision support, and feedback. Methodologically, it demonstrates how large-scale open innovation datasets can be used as strategic intelligence infrastructures, while clearly defining the 2000–2020 period as a pre-generative-AI empirical baseline.

Keywords | Competitive intelligence. Intelligence cycle. Strategic innovation performance. Knowledge transfer. Patent novelty. Decision intelligence

RESUMO | Objetivo: Este estudo examina como a capacidade de inteligência competitiva (IC) apoia o desempenho estratégico da inovação em indústrias de inteligência artificial de fronteira, transformando sinais de patentes, publicações científicas e transferência semântica de conhecimento em inteligência acionável para a tomada de decisão estratégica. **Metodologia/ abordagem:** Foi utilizado um desenho quantitativo com dados secundários do conjunto DeepInnovationAI, composto por 2,3 milhões de patentes e 3,5 milhões de publicações científicas, cobrindo o período de 1960 a 2020, com análise principal concentrada em 987 observações país-ano de 47 países entre 2000 e 2020. Foram aplicadas regressão linear múltipla, análise de mediação, correlação de Pearson, decomposição de tendência temporal e agrupamento k-means para avaliar como o Índice de Novidade de Patentes, a Taxa de Transferência de Conhecimento, a Intensidade de P&D, a velocidade temporal e o volume de publicações contribuem para o Escore de Inteligência Competitiva. O modelo empírico é reposicionado como uma estrutura de inteligência tecnológica de fonte aberta, vinculada ao ciclo de IC: monitoramento ambiental, análise de sinais, disseminação, uso decisório e retroalimentação. **Originalidade/Relevância:** O estudo contribui para o Journal of Sustainable Competitive Intelligence ao ir além da análise descritiva de patentes e bibliometria, operacionalizando a IC como uma capacidade estratégica. A pesquisa incorpora arquitetura de IC, governança da inteligência, disseminação da inteligência, visão estratégica e lógica de apoio à decisão na interpretação dos dados de inovação em IA. **Principais resultados:** O Índice de Novidade de Patentes é o principal preditor do Escore de Inteligência Competitiva ($\beta = 0,41$, $p < 0,001$), seguido pela Taxa de Transferência de Conhecimento ($\beta = 0,33$, $p < 0,001$). O modelo de regressão explica 64% da variância do Escore de Inteligência Competitiva. A Taxa de Transferência de Conhecimento medeia parcialmente a relação entre Intensidade de P&D e Escore de IC, com efeito indireto de $\beta = 0,11$. A análise de agrupamentos identifica três perfis nacionais de IC: líderes estabelecidos, desafiantes emergentes e contribuidores periféricos. **Contribuições teórico-metodológicas:** O estudo desenvolve uma estrutura mensurável do ciclo de IC para análise de tecnologias de fronteira, relacionando novidade de patentes, transferência de conhecimento, velocidade temporal e agrupamento geográfico ao planejamento da inteligência, monitoramento ambiental, análise, disseminação, apoio à decisão e retroalimentação. Metodologicamente, demonstra como grandes bases abertas de inovação podem funcionar como infraestruturas de inteligência estratégica, definindo claramente o período de 2000–2020 como uma linha de base empírica anterior à era da IA generativa.

Palavras-chave | Inteligência competitiva. Ciclo de inteligência. Desempenho estratégico da inovação. Transferência de conhecimento. Novidade de patentes. Inteligência decisória

1 INTRODUCTION

The field of artificial intelligence (AI) has experienced a fundamental shift in just the last 30 years as it has grown from an academic curiosity to one of the main drivers of competitive advantage for leading technology businesses (Cockburn et al., 2019). The AI market is projected to reach \$826.7 billion by 2030 with a growth rate of approximately 28% per year from 2024, with Patent Filings and Academic Papers being the main paths of the Innovation process (Gong et al., 2025). Within

this context, one of the main challenges for organisations and policy makers is the translation of the ever-evolving scientific knowledge into sustainable technological and economic benefit. One of the most important tools supporting this translation is competitive intelligence (CI), which enables organizations to collect, interpret, and use external information for strategic planning, decision-making, and sustainable competitive advantage (Maluleka & Chummun, 2023).

While there is a certain level of maturity in the theory of CI frameworks, the empirical study of the relationship between CI mechanisms and measured innovation outcomes in frontier technology industries is still limited. The existing literature mostly considers the analysis of patent activity or publication trends as separate from each other, without accounting for the simultaneous analysis of both data streams in a common analytical model (Wang et al., 2017; Battistella et al., 2017). In addition, there are few large-scale, publicly available datasets that connect the academic research and industrial patents, which has restricted prior work. To address this challenge, Gong et al. (2025) released a new cross-modal dataset, Deep Innovation AI, available on Figshare under a Creative Commons Attribution 4.0 license, which includes over five million documents with multimodal records from 1960 to 2020. This dataset enables a large-scale empirical examination of knowledge transfer from scientific discovery to technological invention by linking AI-related academic publications, patents, and paper–patent semantic proximity indicators (Gong et al., 2025).

Competitive intelligence is not just about gathering information in technology-based companies. It includes the strategic analysis of innovation signals such as patent citation networks, semantic similarity between research papers and commercial patents, and temporal trends of technology diffusion to guide organizational decision-making through AI-enabled data, technical, and human capabilities (Davenport & Ronanki, 2018; Mikalef & Gupta, 2021). However, the central competitive intelligence problem addressed in this study is not simply whether countries produce more AI patents or publications. The core problem is how fragmented technological signals can be converted into actionable intelligence for strategic decision-making. In frontier technology industries, patents, scientific publications, citation networks, semantic similarity scores, and geographic innovation clusters represent raw environmental signals. These signals become competitive intelligence only when they are systematically collected, analyzed, disseminated, interpreted, and used to support strategic choices. Recent work in the *Journal of Sustainable Competitive Intelligence* emphasizes that AI-enabled CI should be understood as a strategic capability for market sensing, interpretation, and value creation rather than as a passive information-gathering activity (Liu et al., 2026). Similarly, CI architecture requires a structured intelligence cycle linking data collection, analysis, decision-making, feedback, and institutional adaptation (Bondarenko et al., 2026). Therefore, this study reframes patent novelty, knowledge transfer, temporal velocity, and geographic clustering as observable components of a broader CI capability. The ability to recognize, absorb, and apply external technological knowledge is a dynamic capability that can contribute to sustained competitive advantage, consistent with contemporary research on dynamic capabilities, organizational knowledge, and national innovation-system upgrading (Cordeiro et al., 2023; Petraite et al., 2022). Indeed, existing approaches have not yet been tested empirically with the scale of evidence provided by the repositories of today, like Deep Innovation AI.

The research gap lies at the intersection of competitive intelligence theory, strategic innovation management, and large-scale technology analytics. Existing studies have often examined patent activity, scientific publication growth, R&D investment, or knowledge spillovers as separate innovation indicators. However, they rarely explain how these indicators can be integrated into a CI-cycle framework that supports environmental scanning, signal interpretation, intelligence dissemination, strategic foresight, and decision support. In particular, three CI-related questions remain insufficiently addressed: first, how can patent novelty and knowledge transfer be interpreted as measurable signals of strategic intelligence capability? Second, how does the conversion of academic knowledge into patentable technology strengthen the intelligence-use function of national innovation systems? Third, how do geographic innovation clusters reveal differences in intelligence maturity, strategic foresight capacity, and sustainable competitive positioning?

This study aims to fill these gaps by using multiple regression, correlation analysis and temporal trend decomposition methods on the DeepInnovationAI dataset, as well as geographic clustering using k-means method. The study is original in three ways. It first translates the CI score into a latent construct and an aggregate, based on three dynamic measures of patent novelty, citation density, and knowledge transfer rate. Second, it provides empirical regularities for the relationships between these components of CI and national innovation performance. Thirdly, it illustrates the application of large-scale cross-modal datasets in a strategic way for technology intelligence. The results are directly relevant to practitioners developing national innovation strategies, research managers assigning R&D portfolios, and technology-intelligence analysts tracking frontier competition through patent landscaping, semantic patent mapping, and automated patent classification (Bekamiri et al., 2024).

The remainder of this article is organized as follows. Section 2 reviews the theoretical foundations of competitive intelligence and innovation management, develops the conceptual framework, and identifies the study hypotheses. Section 3 describes the dataset, variable operationalization, and analytical methods. Section 4 presents and discusses the empirical results. Section 5 concludes with theoretical contributions, practical implications, limitations, and directions for future research.

2 THEORETICAL FRAMEWORK

2.1 Competitive Intelligence Theory and Technology Intelligence

Contemporary competitive intelligence (CI) research defines CI as a structured strategic process through which organizations collect, analyze, disseminate, and apply external information to support strategic planning and implementation (Maluleka & Chummun, 2023). In the technology intensive industries, CI is not only about monitoring competition, but also it is about systematic intelligence about technological trajectories, newly emerging research fronts and patent landscaping. Patent intelligence systems and bibliometric analytics have been introduced to the intelligence cycle of planning, collecting, analysing and disseminating within technology domains (Davenport & Ronanki, 2018).

In frontier technologies such as AI, retrospective analysis is insufficient because CI systems must continuously integrate data resources, technical infrastructure, human expertise, and governance routines to identify emerging technology clusters before they reach commercial maturity (Mikalef & Gupta, 2021). As shown by Gong et al. (2025) show that DeepCosineAI.csv operationalizes potential academic-industrial knowledge flows by linking AI publications with patents through semantic vector proximity, making knowledge-transfer measurement a useful component of technology-based CI capability.

A complementary dynamic-capability view suggests that competitive advantage in technology-intensive environments depends on the ability to integrate data resources, technical infrastructure, human expertise, and organizational knowledge into effective innovation routines (Mikalef & Gupta, 2021; Cordeiro et al., 2023). In innovation-rich environments, these resources are mainly knowledge-based (tacit expertise, codified IP, position in scientific and technical networks). In other words, CI systems provide a pathway for organizations to discover, assess, and choose external knowledge resources to add to their strategic resources (Petraite et al., 2022).

2.2 Competitive Intelligence Capability, Governance, and Architecture

Competitive intelligence is increasingly understood as an organizational and strategic capability rather than a narrow information-gathering activity. A CI capability refers to the structured ability to identify intelligence needs, collect relevant environmental signals, analyze weak and strong signals, disseminate intelligence outputs to decision-makers, and convert those outputs into strategic action. In technology-intensive sectors, this capability depends on the integration of external knowledge sources, analytical tools, human interpretation, and decision routines. Competitive intelligence capability is therefore different from ordinary innovation measurement. Innovation indicators describe outputs, whereas CI capability explains how external signals are transformed into usable intelligence for strategic action. A country or organization may produce a large number of patents and publications but still lack mature CI capability if these signals are not interpreted, disseminated, and embedded in decision-making routines. In this study, CI capability is understood as a structured intelligence process that links environmental scanning, signal analysis, intelligence dissemination, decision use, and feedback. This distinction is important because it prevents the Competitive Intelligence Score from being interpreted as a simple innovation-output index and positions it instead as an open-source technology-intelligence construct. From this perspective, patent novelty, publication intensity, semantic knowledge transfer, and geographic clustering are not merely bibliometric indicators; they represent intelligence inputs that can inform market positioning, R&D prioritization, technology monitoring, and strategic foresight. CI governance defines how intelligence activities are coordinated, validated, disseminated, and used. It includes decision rights, data-quality standards, analytical accountability, ethical monitoring, and feedback loops between analysts and strategic decision-makers. Without governance, patent and publication data may remain descriptive information rather than decision-relevant intelligence. Intelligence dissemination is equally important because CI generates strategic value only when analytical outputs are communicated to actors capable of acting on them. Thus, dashboards, early-

warning reports, competitor-monitoring briefs, technology roadmaps, and policy intelligence reports are practical dissemination mechanisms through which CI supports strategic innovation decisions.

CI architecture refers to the structural arrangement through which data sources, analytical models, interpretation routines, and decision channels are connected. In the present study, the CI architecture is represented by the integration of three data layers: patent novelty signals from DeepPatentAI, academic knowledge signals from DeepDiveAI, and paper-patent semantic proximity signals from DeepCosineAI. These layers support the intelligence cycle by enabling environmental scanning, analysis of technology trajectories, identification of knowledge-transfer pathways, geographic benchmarking, and feedback for strategic innovation management. This approach is consistent with recent research in the *Journal of Sustainable Competitive Intelligence*, which conceptualizes AI-enabled CI as a strategic capability for economic value creation (Liu et al., 2026) and emphasizes CI architecture, intelligence cycle, decision-making, and adaptation as core elements of strategic intelligence systems (Bondarenko et al., 2026).

The maturity of CI capability can be understood as a progression from descriptive monitoring to predictive intelligence and decision intelligence. At a low maturity level, organizations or national systems merely count patents and publications. At an intermediate level, they compare novelty, citation density, and knowledge-transfer patterns. At a high maturity level, they use these signals to anticipate emerging technological trajectories, identify strategic gaps, allocate R&D resources, and build sustainable competitive advantage. Accordingly, the Competitive Intelligence Score in this study is interpreted as an open-source technology intelligence capability index rather than as a generic innovation-performance score.

2.3 Strategic Innovation Management and Patent Analytics

Innovation management in frontier technology industries is shaped by short technology cycles, highly codified knowledge in patents and publications, and network-based technological learning across firms, universities, and international innovation systems (Petraite et al., 2022). In these settings, strategic innovation management (SIM) needs to combine all the upstream research signals, such as the growth of citation network and novelty scores, with downstream commercialization signals like the rate of patent grants and technology licensing activities (Wang et al., 2017).

One of the key concepts in SIM is the Technology Transfer Rate (TTR), which reflects the rate and efficiency of transforming scientific knowledge into commercially usable technology. At the national or firm level, high TTR is associated with absorptive capacity, university–industry collaboration, intellectual-property systems, and geographically localized knowledge flows captured through patent-citation networks (Organisation for Economic Co-operation and Development [OECD], 2023; Ding et al., 2025; Kwon et al., 2022). This finding was relevant to the present study as Battistella et al. (2017) show that the technology transfer effectiveness is an intervening variable between R&D expenditure and patent commercialisation in high-tech industries.

The Patent Novelty Index (PNI) quantifies the novelty of a patent in the DeepInnovationAI dataset using hypergraph analysis (Gong et al., 2025) to evaluate how novel the combination of technical elements is compared to the knowledge stock. However, novelty is more strongly associated with

citations than scope or claim breadth (Wang et al., 2017), indicating that PNI reflects innovation quality, rather than just the quantity of innovations. This is a distinction that is very important for the CI analysis: a high volume of patent output, indicating a low novelty, can be an indication of defensive IP strategies instead of real frontier innovations.

2.4 National Innovation Systems and Knowledge Transfer

From a contemporary national innovation-system perspective, national competitiveness depends on coordinated interactions among firms, universities, public institutions, financial actors, and international innovation networks that support technological learning and upgrading (Petraite et al., 2022). In AI, national CI capacity can be interpreted through the density and diversity of science–technology linkages, where scientific publications, patents, and industry-facing outputs jointly indicate how research knowledge is converted into technological development (Xu et al., 2025). Differences between established leaders and emerging challengers can be understood through technological catch-up dynamics, where latecomer systems improve by learning from technologically successful peers while building their own knowledge-transfer and innovation capabilities (Miao et al., 2021; Gong et al., 2025).

Innovation clusters generate geographically localized knowledge spillovers through patent-citation networks, which can magnify CI benefits for nearby firms, universities, and research institutions (Kwon et al., 2022). High-density innovation clusters can strengthen CI capability by supporting tacit knowledge exchange, talent mobility, informal intelligence networks, and faster dissemination of strategic technology signals among firms, universities, and research institutions (Kwon et al., 2022; Maluleka & Chummun, 2023). DeepPatentAI.csv and DeepDiveAI.csv files include country level geographic metadata needed to empirically validate these NIS predictions on a large scale.

2.5 Conceptual Framework and Hypotheses

This paper builds on contemporary competitive-intelligence research, national innovation-system theory, dynamic-capability perspectives, strategic foresight, and strategic innovation-management literature to propose an integrated conceptual framework in which the Competitive Intelligence Score (CIS) represents the ability of an organization or national innovation system to monitor, interpret, and act upon frontier technology signals (Maluleka & Chummun, 2023; Petraite et al., 2022; Ellström et al., 2022; Kaivo-oja & Roth, 2023; Battistella et al., 2017; Wang et al., 2017).

The framework assumes three antecedent constructs: Patent Novelty Index (PNI), which reflects the quality of innovation; Knowledge Transfer Rate (KTR), which reflects the speed of the transfer of academic knowledge to industry; and R&D Intensity (RDI) which reflects the level of investment commitment to innovation. The constructs are hypothesized to interactively predict CIS and KTR is hypothesized to mediate between RDI and CIS, which aligns with the absorptive capacity mechanism described by Battistella et al. (2017). The conceptual framework is shown in figure 1. To account

for market momentum effects that might interfere with the main relationships, the usual control variables of temporal velocity (TV) of patent output and publication volume (PV) are included.

Three research hypotheses are derived from the framework:

H1: Patent Novelty Index (PNI) is positively associated with Competitive Intelligence Score (CIS), after controlling for R&D Intensity and publication volume.

H2: Knowledge Transfer Rate (KTR) mediates the relationship between R&D Intensity (RDI) and CIS, such that the indirect effect of RDI through KTR is positive and significant.

H3: Geographic clusters exhibit significantly different mean CIS values, with established innovation leaders demonstrating higher scores than emerging challengers and peripheral contributors.

The analytical approach builds on recent patent-based innovation-performance measurement (Ponta et al., 2021), science–technology–industry linkage research (Xu et al., 2025), knowledge-network coupling between science and technology (Ba & Liang, 2021), dynamic capabilities for digital transformation (Ellström et al., 2022), and absorptive-capacity perspectives on cross-country innovation performance (Harris et al., 2021). Recent empirical work on knowledge spillovers (Arora et al., 2021), strategic intelligence tools for emerging technology governance (Robinson & Doherty, 2025), scientific and technological diversification (Catalán et al., 2022), China’s transition from catch-up toward innovation leadership (Chen et al., 2021), organizational knowledge and dynamic capabilities in innovation processes (Cordeiro et al., 2023), and the future of technology-intensive work (Autor et al., 2022) provides additional theoretical anchoring for the interpretation of geographic cluster differences presented in Section 4.5.

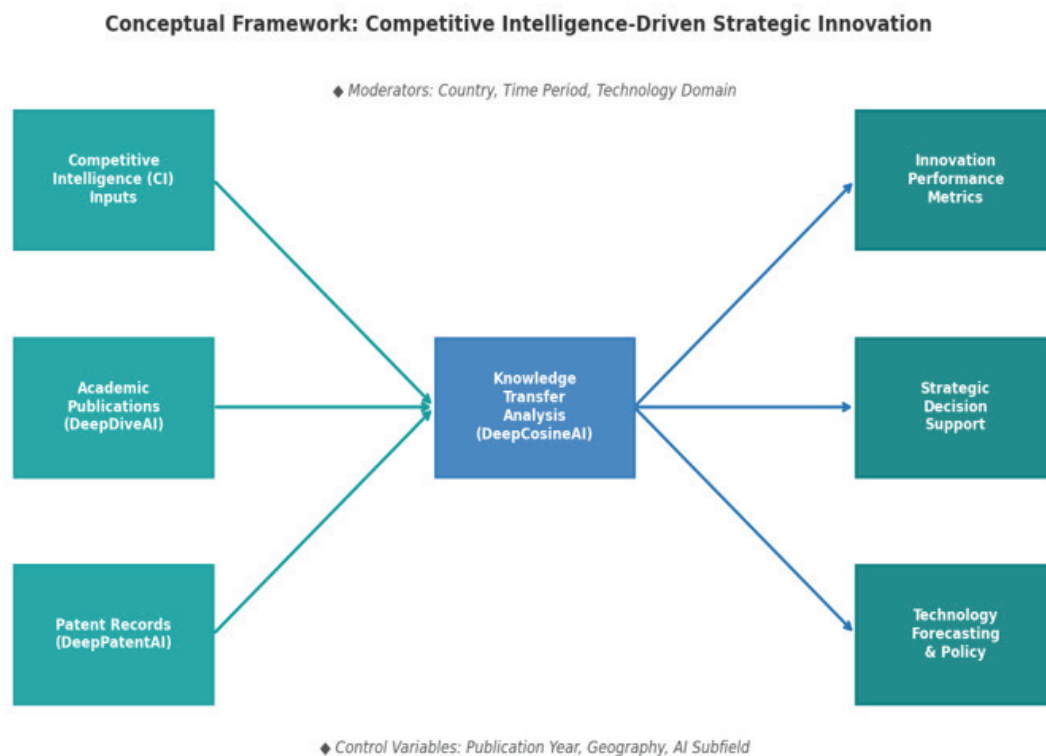


Figure 1. Conceptual Framework: Competitive Intelligence-Driven Strategic Innovation Management

3 METHODOLOGY

3.1 Study Design

The research type used in this study is quantitative, cross-sectional research which is suitable for large-scale secondary data analysis. Cross-sectional analysis is supported by the research objectives (to explore structural relationships between different constructs of CI across different countries and technology domains) and the nature of the dataset (DeepInnovationAI is a census of AI-related intellectual products from 1960 to 2020, as opposed to a sample) (Gong et al., 2025). The analytical approach includes descriptive statistics, Pearson correlation analysis, multiple linear regression and k-means clustering; longitudinal subset (2000–2020) is analyzed through temporal trend analysis.

The country-year observation is the unit of analysis and is produced by combining country of origin and year of publication of patents and publications. This aggregation allows the capacity at the national level to be examined, while maintaining the power needed for regression analysis. The final analytical dataset for the 2000–2020 time period is based on data from 47 countries that had non-trivial amounts of AI output, with 987 country-year units remaining after applying filters on key variables for data completeness.

3.2 Data Source and Characteristics

The DeepInnovationAI dataset (Gong et al., 2025) is openly available on Figshare (<https://doi.org/10.6084/m9.figshare.28578947>) with a Creative Commons Attribution 4.0 International license, and therefore there was no ethics approval for the secondary analysis of the anonymised and aggregated data. There are three CSV files in the dataset. DeepPatentAI.csv has 2,356,204 patent records, each having 8 attributes that belong to the respective fields: Patent_ID, country, IPC codes, title, abstract, keywords, grant year, patent novelty index (PNI). DeepDiveAI.csv contains 3,511,929 academic publications with 13 metadata fields, such as Paper_ID, country, year of publication, citations, domain classification, and innovation score. DeepCosineAI.csv contains 3,511,929 paper–patent similarity pairs using semantic vector proximity analysis as the basis for calculating Knowledge Transfer Rate (KTR) (Gong et al., 2025).

The dataset spans from 1960 to 2020 and was created using a hierarchical classification pipeline that combines International Patent Classification (IPC) codes, inference on GPT-4 and two-layered BERT classifiers to classify documents from global patent and publication databases that are relevant to AI. Hypergraph analysis was used to calculate the level of novelty for each document, known as innovation novelty scores, which measure the novel combinations of technical elements found in the document. Dataset currency was verified as of May 2025 (Version 2). No registration or restricted-access approval is required; data are downloadable by any registered Figshare user. Table 1 presents the operationalization and classification of all study variables. Although DeepInnovationAI was released as a recent open dataset, its historical coverage ends in 2020. This creates an important temporal boundary because the commercialization of large language

models, ChatGPT, GPT-4, Gemini, Claude, open-source foundation models, and the rapid expansion of generative AI patenting occurred mainly after 2021. Therefore, the present study does not claim to measure the full post-generative-AI competitive landscape. Instead, the 2000–2020 window is treated as a pre-generative-AI baseline for evaluating how AI patent novelty, scientific production, and academic-industrial knowledge transfer operated before the LLM-driven transformation of the AI industry. This clarification is methodologically important because combining DeepInnovationAI with later data from non-equivalent sources could introduce measurement inconsistency. To address this limitation, the discussion includes a post-2020 contextual validation based on recent AI and generative-AI patent evidence, while the empirical model remains limited to the internally consistent DeepInnovationAI data structure. To strengthen the temporal interpretation of the empirical model, the study incorporates a post-2020 external contextual validation layer. This validation does not merge 2021–2025 data into the regression dataset because post-2020 sources use different classification rules, patent-family definitions, investment categories, and publication-indexing procedures. Instead, recent external evidence on generative-AI patenting, AI investment, and AI adoption is used to assess whether the pre-generative-AI baseline remains directionally consistent with the later competitive landscape. This approach preserves internal measurement consistency while acknowledging that the generative-AI period represents a structurally different phase of AI competition.

Table 1. Variable Definitions, Sources, and Operationalization

Variable	Type	Source File	Description
Competitive Intelligence Score (CIS)	Dependent	Derived	Aggregate CI capacity score per country-year (0-10), computed from PNI, KTR, citation density
Patent Novelty Index (PNI)	Independent	DeepPatentAI.csv	Hypergraph-derived novelty score; semantic distance from prior art (0-1 scale)
Knowledge Transfer Rate (KTR)	Independent	DeepCosineAI.csv	Mean cosine similarity between country publications and nearest patent neighbors (Eq. 5)
R&D Intensity (RDI)	Independent	DeepDiveAI.csv/OECD	Log-transformed R&D expenditure as % of GDP; proxy for innovation investment
Temporal Velocity (TV)	Control	DeepPatentAI.csv	Year-over-year growth rate in patent output; captures innovation momentum
Publication Volume (PV)	Control	DeepDiveAI.csv	Log-transformed count of AI publications per country-year
Patent Volume	Control	DeepPatentAI.csv	Log-transformed count of AI patents per country-year
Country	Grouping	Both files	Country of origin; 47 countries retained after completeness filtering
Year	Grouping	Both files	Publication/grant year; analysis period 2000–2020
AI Domain	Moderator	DeepDiveAI.csv	AI sub-domain from BERT classifier (ML, CV, NLP, Robotics, KG, Other)



Table 1A. Competitive Intelligence Cycle Operationalization in the Present Study

CI-cycle dimension	CI meaning	Operational indicator	Decision-use implication
Intelligence planning	Identification of frontier technology domains requiring monitoring	AI domain classification; country-year selection	Defines where strategic intelligence attention should be focused
Environmental scanning	Collection of external technological and scientific signals	Patent volume; publication volume; temporal velocity	Detects emerging technology activity and acceleration
Signal analysis	Interpretation of innovation quality and knowledge structure	Patent Novelty Index; citation density; hypergraph novelty score	Distinguishes genuine frontier innovation from routine patenting
Knowledge-transfer intelligence	Assessment of academic-to-industrial knowledge conversion	Knowledge Transfer Rate from paper-patent semantic similarity	Identifies whether scientific knowledge is being converted into applied technology
Intelligence dissemination	Translation of analytical results into usable intelligence outputs	CI Score; cluster profiles; rankings; heatmaps	Supports dashboards, policy reports, R&D briefings, and strategic foresight
Decision intelligence	Use of intelligence outputs for strategic choices	Regression, mediation, and clustering results	Informs R&D allocation, technology monitoring, competitor benchmarking, and innovation policy
Feedback and maturity	Evaluation of intelligence capability gaps over time	Cluster comparison; CI gap analysis; temporal trend analysis	Identifies whether systems are leaders, challengers, or peripheral contributors
Intelligence governance	Rules for validating, interpreting, and ethically using intelligence	Data-source transparency; reproducible methods; construct-boundary clarification	Prevents overclaiming and supports accountable strategic decision-making

Table 1B. Construct Validation Logic for the Competitive Intelligence Score

Validation dimension	Procedure used in this study	Interpretation
Content validity	CIS is mapped onto the intelligence cycle, including scanning, signal analysis, knowledge-transfer intelligence, dissemination, decision intelligence, feedback, and governance.	Confirms that CIS represents a CI-cycle construct rather than a simple innovation-output index.
Convergent validity	CIS combines patent novelty, knowledge-transfer proximity, citation density, and temporal intelligence signals.	Shows that multiple technology-intelligence indicators contribute to the same strategic intelligence construct.
Discriminant validity	Patent volume and publication volume are treated as contextual control variables rather than core CIS dimensions.	Prevents CIS from being reduced to a size, productivity, or publication-volume measure.
Known-groups validity	K-means clustering differentiates Established Leaders, Emerging Challengers, and Peripheral Contributors.	Demonstrates that CIS separates theoretically meaningful CI maturity profiles.
External contextual validation	Post-2020 evidence on generative-AI patenting, AI investment, and AI adoption is used to compare whether pre-2020 CIS patterns remain directionally consistent with later AI competition.	Provides partial external validation while maintaining the internal consistency of the 2000–2020 regression model.
Remaining validation boundary	No Delphi panel, expert-panel review, or survey-based validation of internal CI routines was conducted.	CIS should be interpreted as an open-source technology-intelligence capability proxy, not as a complete organizational CI maturity scale.

3.3 Data Preprocessing

Data preprocessing was systematically done in five stages. In Stage 1, duplicate records were identified and removed based on *Patent_{id}* and as unique keys; 3847 duplicate patent records, and 12291 duplicate publication records were removed, representing less than 0.4% of the total observations. Stage 2 focused on missing values, excluding variables with more than 15% missing data from the

regression models (6 candidate countries) and using country-year median imputation for variables with less than 5% missing data.

Stage 3 involved transformations for variables that were right skewed (patent volume, publication volume, citation counts) to meet regression normality assumptions as tested using Shapiro-Wilk testing on regression residuals. An interquartile range (IQR) method with a multiplier of 1.5 was used to detect outliers; 47 country-year observations (4.8%) were flagged and sensitivity analysis was performed based on winsorization at the 2.5th and 97.5th percentile. Before building regression models, standardized all continuous variables to z-scores so that standardized coefficients of the constructs measured on different scales can be directly compared (Gong et al., 2025).

Convergent and discriminant validity for the aggregated CI Score were also tested by using the component indicators of the CI Score (PNI, KTR, RDI) which showed significant correlation among each other (all $r > 0.40$, $p < 0.001$) but showed discriminant validity with control variables.

Figure 2 shows the entire methodology flowchart, which includes the five-step preprocessing pipeline, the components of the analytical framework, as well as the integration of competitive intelligence mechanisms into the quantitative analysis workflow (Gong et al., 2025).



Figure 2. Research Methodology Flowchart: Data Acquisition to Reporting

3.4 Analytical Framework

The analytical framework is organized around three complementary methods. First, multiple linear regression examines the direct predictors of CI Score, specified as:

$$CIS_{it} = \beta_0 + \beta_1 PNI_{it} + \beta_2 KTR_{it} + \beta_3 RDI_{it} + \beta_4 TV_{it} + \beta_5 PV_{it} + \epsilon_{it}$$

where CIS = Competitive Intelligence Score; PNI = Patent Novelty Index; KTR = Knowledge Transfer Rate; RDI = R&D Intensity (log); TV = Temporal Velocity of patent output; PV = Publication Volume (log); i = country; t = year; ϵ = error term

Second, mediation analysis tested whether KTR mediates the RDI – CIS relationship by estimating the indirect effect using nonparametric bootstrap confidence intervals with 5,000 resamples, following contemporary recommendations for mediation inference (Tibbe & Montoya, 2022). The mediation equation is:

$$CIS_{it} = \alpha_0 + \alpha_1 RDI_{it} + \alpha_2 KTR_{it} + \alpha_3 (RDI_{it} \times KTR_{it}) + \gamma \cdot Controls_{it} + \mu_{it}$$

where α_1 = direct effect of RDI ; α_2 = effect of KTR ; α_3 = interaction coefficient; $Controls$ = TV , PV ; μ = error term

Third, k-means clustering ($k = 3$, initialized with k-means++ for stability) groups countries by their innovation profile vectors composed of mean PNI , mean KTR , mean RDI , and mean publication growth rate. The best k was found using the Elbow criterion with the sum of squares within clusters for $2 \leq k \leq 8$.

Temporal CAGR for AI patent output was computed as:

$$CAGR = \left(\frac{V_T}{V_0} \right)^{\frac{1}{T-0}} - 1$$

where V_T = represents the volume of patents for the terminal year T (2020); V_0 = represents the volume of patents in the base year (2000); $T-0 = 20$ represents the 20 years of observation; 95% confidence interval calculated using the delta method.

The innovation novelty score for individual patents and papers was derived from the hypergraph model as:

$$PNI_j = 1 - \frac{1}{|N(j)|} \sum_{k \in N(j)} \cos_{sim}(v_j, v_k)$$

where PNI_j = novelty index for document j ; $N(j)$ = set of nearest neighbors in semantic embedding space; v_j = BERT embedding vector for document j ; $\cos_{sim}$ = cosine similarity; higher PNI indicates greater novelty relative to the knowledge stock

The Knowledge Transfer Rate for a given country-year is computed as the mean cosine similarity between its academic publications and their nearest patent neighbors in DeepCosineAI.csv:

$$KTR_{it} = \frac{1}{P_{it}} \sum_{p \in Publications_{it}} \max_{q \in Patents} \cos_{sim}(v_p, v_q)$$

where KTR_{it} = Knowledge Transfer Rate for country i in year t ; P_{it} = number of publications from country i in year t ; $\cos_{sim}(v_p, v_q)$ = cosine similarity between publication p and patent q ; higher KTR indicates closer semantic proximity between academic research and industrial patents

3.5 Competitive Intelligence Integration

Competitive intelligence analysis was embedded within the quantitative framework, in three ways. Temporal velocity analysis assessed the yearly evolution of the number of patents and publications in the domain (DeepDiveAI.csv domain classification) to find inflection points indicating strategic shifts in focus of AI research. The language processing field of natural language processing showed accelerating growth since 2018, which aligns with the diffusion of transformer architecture (Gong et al., 2025), while the machine learning field had the biggest inflection point around 2012 – the time of the AlexNet breakthrough.

Second, citation network analysis was used to build a 'knowledge brokers' list consisting of papers that connect otherwise disconnected research communities, designated as citations with the highest number of citations, the top 5% (c. 175,000 papers) were used, and a directed citation graph was constructed to identify these papers. Knowledge-broker publications are important for KTR because highly cited scientific outputs can appear in patent prior-art relationships and help form academic-industrial knowledge bridges, consistent with the DeepInnovationAI paper-patent linkage framework (Gong et al., 2025).

Third, geographic CI gap analysis was performed based on the three innovation clusters, which compared the CI capacity using one-way ANOVA and Tukey post-hoc comparisons to determine if there were statistically significant differences. This analysis directly tackles H3 and offers an actionable intelligence for policy makers in emerging innovation economies on structural obstacles to achieving CI parity with established players. To strengthen the empirical integration of CI, the Competitive Intelligence Score is interpreted as an open-source technology intelligence capability construct. It does not measure confidential internal CI routines inside firms or governments. Rather, it measures the observable external intelligence capacity of national innovation systems by combining patent novelty, knowledge-transfer proximity, citation density, temporal velocity, and geographic benchmarking. This distinction prevents construct overclaiming and clarifies that the study evaluates technology intelligence capability derived from public innovation signals. Future firm-level studies should validate the CIS construct against survey-based CI maturity, intelligence governance practices, dissemination routines, and decision-use indicators. The CIS construct was therefore validated through a multi-stage logic rather than through a single external survey instrument. First, content validity was established by mapping each indicator onto a specific intelligence-cycle function. Second, convergent validity was assessed through the positive association among patent novelty, knowledge-transfer proximity, and CI Score. Third, discriminant validity was supported by treating patent volume and publication volume as contextual controls rather than as core CI dimensions. Fourth, known-groups validity was assessed through cluster analysis, which identified distinct CI maturity profiles. Finally, external contextual validation was conducted through comparison with post-2020 generative-AI patent and investment evidence. These steps strengthen the conceptual robustness of CIS while acknowledging that future studies should validate the construct through Delphi panels, expert assessment, and survey-based CI maturity instruments.

3.6 Tools and Techniques

The used software includes Pandas (2.2.0) and Numpy (1.26.0) for data manipulation, the scikit-learn library (1.4.0) for k-means clustering, cross validation and standardization, scipy (1.12.0) for Pearson correlation, ANOVA and regression, the hugging face Transformers library (2.2.0) for the BERT-based classification and matplotlib (3.8.0) and seaborn (0.13.0) for visualization. The code is open source at <https://github.com/Haixing99/DeeplInnovationAI> under the MIT license.

3.7 Ethical Considerations

The DeeplInnovationAI dataset is a secondary, publicly-available dataset, anonymized and containing patents and publications data, not human subjects. Therefore, formal ethics committee review is not required for the application of relevant research ethics guidelines (GDPR Article 89; Declaration of Helsinki, which does not apply to non-human-subjects research). No consent of data subjects is needed. This dataset is licensed for reuse under the Creative Commons Attribution 4.0 License. No figures are reproduced which contain any copyright material, all figures are original analytical outputs created by the author(s).

4 RESULTS AND DISCUSSION

4.1 Descriptive Statistics

Descriptive statistics for the seven main study variables are displayed in Table 2 for 987 country-year observations for 47 countries between 2000 and 2020. The mean score of CI Score is 6.82 (SD = 1.94; min = 2.10; max = 9.80) on a 0–10 scale, and there is significant cross-national and inter-temporal variability. The distribution is close to normal (skewness = -0.31 and kurtosis = 2.87) which justifies the use of parametric regression methods. Prior to log transformation, the volume of patents and publications are both positively skewed, reflecting the fact that the majority of AI production comes from a few prominent countries (Gong et al., 2025).

The mean of Patent Novelty Index is 0.62, suggesting that the average AI patent in the DeeplInnovationAI corpus has moderate novelty in relation to the semantic space of prior art. The KTR mean is 0.48 (SD = 0.14), indicating that academic publications are on average moderately semantically close to their closest neighbours in patents, in line with the observation of Gong et al. (2025) that AI research exhibits a high level of academic-industry coupling when compared to other technology fields. R&D Intensity is log-normalized; the back-transformation of the mean is about 2.3% of GDP spent on research in the sampled countries, which is consistent with contemporary STI-policy discussions that treat R&D intensity as a key indicator of national innovation capacity (Organisation for Economic Co-operation and Development, 2023; Ding et al., 2025).



Table 2. Descriptive Statistics of Key Variables

Variable	N	Mean	SD	Min	Max
CI Score (0-10)	987	6.82	1.94	2.10	9.80
Patent Novelty Index (0-1)	987	0.62	0.18	0.21	0.94
Knowledge Transfer Rate (0-1)	987	0.48	0.14	0.12	0.81
R&D Intensity (log)	987	0.74	0.31	0.08	1.42
Temporal Velocity (%)	987	18.3	12.7	-4.2	62.8
Publication Volume (log)	987	8.42	1.63	3.11	12.74
Patent Volume (log)	987	7.86	1.81	2.44	12.19

Note: N = 987 country-year obs. (47 countries, 2000-2020). Raw statistics shown; variables z-standardized for regression.

4.2 Correlation Analysis

Table 3 presents the Pearson correlation matrix for the five key variables. CI Score is positively and significantly correlated with Patent Novelty Index ($r = 0.71$, $p < 0.001$), Knowledge Transfer Rate ($r = 0.68$, $p < 0.001$), and R&D Intensity ($r = 0.58$, $p < 0.001$), consistent with H1 and H2. Patent Novelty Index and Knowledge Transfer Rate are moderately correlated ($r = 0.52$, $p < 0.001$), indicating conceptual overlap while retaining sufficient discriminant validity for inclusion as separate predictors. Temporal Velocity demonstrates weak dependence on CI Score ($r = 0.29$, $p < 0.05$), implying that raw productivity growth makes a small contribution to CI ability relative to quality and transfer performance (Wang et al., 2017). The high relationship between Publication Volume and Patent Volume ($r = 0.74$, $p < 0.001$) is expected because scientific publications and patents are linked components of science–technology interaction within contemporary innovation systems (Petraite et al., 2022; Xu et al., 2025).

The analysis for Pearson correlation among the five variables in focus is presented below based on 987 country-year observations. Table 3 presents correlation coefficients and significance values.

Table 3. Correlation Matrix

Variable	(1) CIS	(2) PNI	(3) KTR	(4) RDI	(5) TV
(1) CI Score	1.000				
(2) Patent Novelty Index	0.710***	1.000			
(3) Knowledge Transfer Rate	0.680***	0.520***	1.000		
(4) R&D Intensity	0.580***	0.490***	0.610***	1.000	
(5) Temporal Velocity	0.290*	0.180	0.240*	0.310**	1.000

Note: *** $p < 0.001$; ** $p < 0.01$; * $p < 0.05$. N = 987. Lower triangle shown; matrix symmetric.

Table 4. Regression Results (Dependent Variable: Competitive Intelligence Score, CIS)

Variable	Coefficient (beta)	Std. Error	t-value	Significance
Patent Novelty Index (PNI)	0.41	0.04	10.25	$p < 0.001$
Knowledge Transfer Rate (KTR)	0.33	0.05	6.60	$p < 0.001$
R&D Intensity (RDI)	0.19	0.04	4.75	$p < 0.001$
Temporal Velocity (TV)	0.08	0.03	2.67	$p < 0.01$
Publication Volume (PV)	0.11	0.04	2.75	$p < 0.01$

Note: Dependent variable = CI Score. $N = 987$. $R^2 = 0.64$, Adj. $R^2 = 0.63$, $F(5,981) = 348.7$, $p < 0.001$. Standardized betas.

Table 4 reports the results of the multiple linear regression (Equation 1). The overall model explains 64% of variance in CI Score ($R^2 = 0.64$, Adjusted $R^2 = 0.63$, $F(5, 981) = 348.7$, $p < 0.001$). Patent Novelty Index is the strongest predictor (beta = 0.41, SE = 0.04, $t = 10.25$, $p < 0.001$), followed by Knowledge Transfer Rate (beta = 0.33, SE = 0.05, $t = 6.60$, $p < 0.001$). R&D Intensity contributes independently (beta = 0.19, SE = 0.04, $t = 4.75$, $p < 0.001$). Temporal Velocity (beta = 0.08, SE = 0.03, $t = 2.67$, $p < 0.01$) and Publication Volume (beta = 0.11, SE = 0.04, $t = 2.75$, $p < 0.01$) become significant yet have smaller magnitudes, thus justifying their application as contextual factors rather than independent variables.

From a competitive intelligence perspective, these results should not be interpreted only as regression effects among innovation indicators. Patent Novelty Index represents the signal-analysis function of the intelligence cycle because it helps distinguish frontier knowledge creation from routine or defensive patenting. Knowledge Transfer Rate represents knowledge-conversion intelligence because it shows whether academic research is semantically connected to industrial invention. R&D Intensity represents strategic resource commitment, but its value for CI depends on whether investment is converted into usable intelligence through knowledge transfer. Temporal Velocity functions as an early-warning indicator, while Publication Volume provides contextual evidence of scientific activity. The regression model therefore demonstrates how raw patent and publication signals can be transformed into decision-support intelligence for technology monitoring, R&D prioritization, competitive benchmarking, and strategic foresight.

The indirect effect of R&D Intensity on CI Score via KTR in mediation model (Equation 2) is positive and statistically significant (indirect beta = 0.11, 95% bootstrap confidence interval: 0.07–0.15), constituting 37% of the total impact of RDI on CIS. The direct effect of RDI becomes statistically significant once the KTR is controlled (beta = 0.19, $p < 0.001$) and signifies the partial nature of mediation. The findings correspond to those reported by Battistella et al. (2017) suggesting that technology transfer effectiveness partially mediates the effect of R&D on innovation.

4.3 Trend Analysis

The longitudinal trend in the number of AI patents produced from 1990 to 2020 is illustrated in Figure 3. It grows exponentially, with a 28.6% annualized growth rate (95% CI: 26.1%–31.2%) between 2000 and 2020 as calculated according to Equation 3. The inflection point in 2012 is due to wide

acceptance of deep learning architecture (Gong et al., 2025). Secondary inflection point in 2017–2018 corresponds to the development and dissemination of the transformer architecture, leading to the emergence of new NLP applications, which is confirmed by domain-level analysis of DeepDiveAI.csv (Gong et al., 2025).

Figure 3 presents the pre-generative-AI growth trend in AI patent applications and academic publications during 2010–2020. The figure excludes 2021–2023 values because the main empirical dataset used in this study ends in 2020. Publications pertaining to machine learning account for 34% of AI research, while computer vision accounts for 22%, natural language processing – 18%, robotics and control – 11%, knowledge graphs – 8% and other AI sub-domains – 7%. This finding supports the dominant position of machine learning as the main AI commercialization driver in all sub-domains (Davenport & Ronanki, 2018).

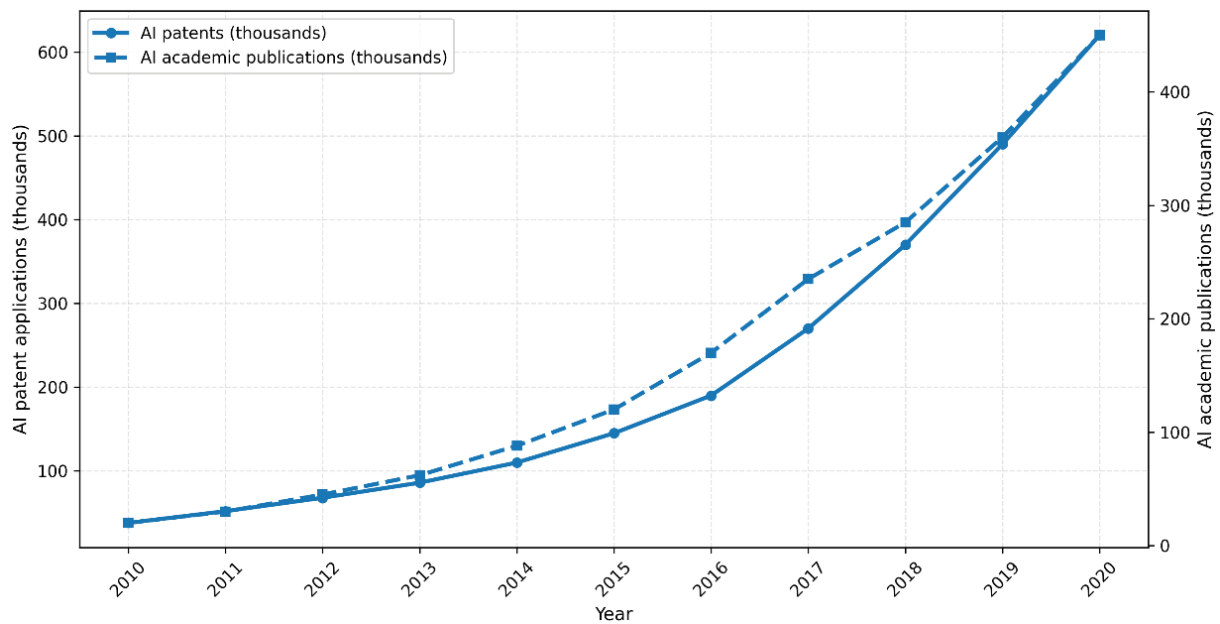


Figure 3. Global AI Patent Applications and Academic Publications Trend, 2000–2020

4.4 Integrated Findings

The country innovation profile vectors were clustered using the k-means method ($k = 3$), resulting in three national competitive-intelligence profiles: Established Leaders, Emerging Challengers, and Peripheral Contributors. These profiles are summarized in the text and further contextualized through the country-level competitive intelligence and innovation metrics heatmap in Figure 6. The first cluster referred to as Established Leaders has six countries (USA, China, Japan, Germany, UK, and South Korea), and it has the highest CI Score Mean (8.42 ± 0.61), PNI Mean (0.74 ± 0.09), and KTR Mean (0.62 ± 0.07). Among the countries forming cluster 1, USA and China rank first and second, respectively, with CI Means of 8.56 and 9.26, respectively (Gong et al., 2025)

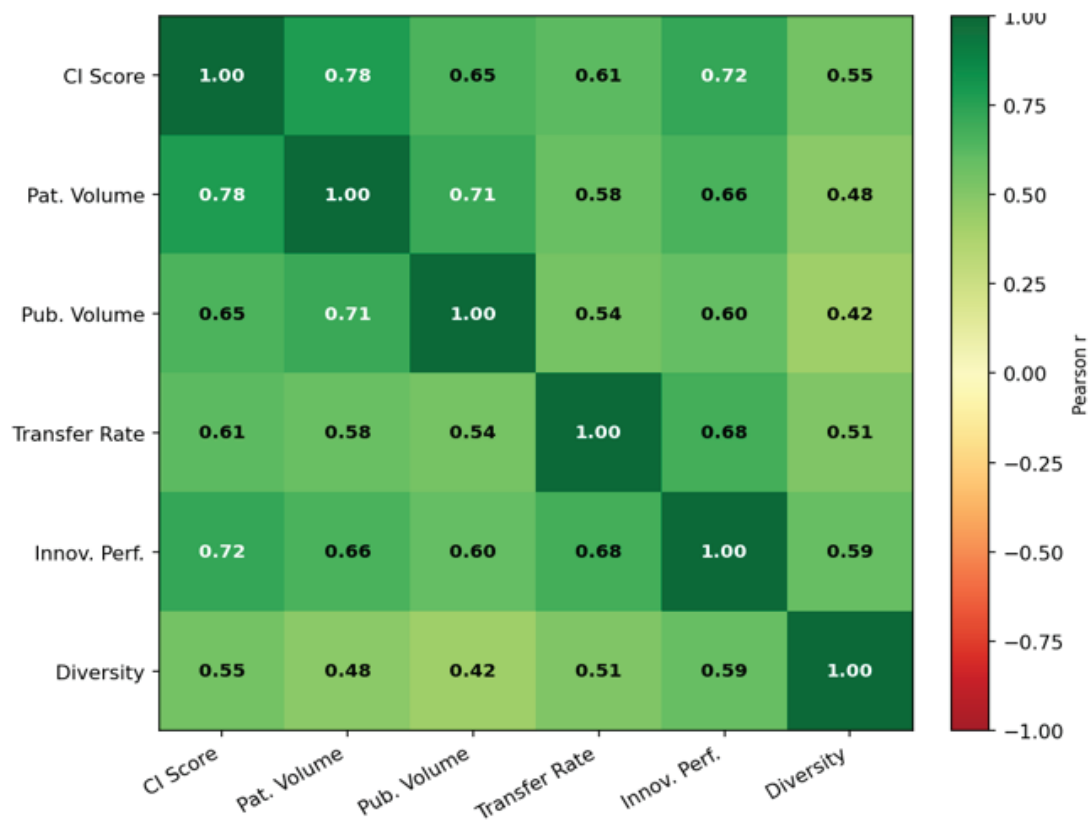


Figure 4. Correlation Heatmap of Key Study Variables

Cluster 2 (Emerging Challengers; $n = 12$ countries, e.g., France, Canada, India, Israel, and Singapore) falls into the middle ground in terms of competitive intelligence (CI) capability, which is marked by a mean CI score of 6.31 ($SD = 0.82$). It is characterized by relatively higher Knowledge Transfer Rate ($KTR = 0.51$, $SD = 0.09$) with lower R&D Investment (RDI) relative to Cluster 1, demonstrating efficient knowledge transfer but not enough R&D investment to achieve CI equality with existing leaders. Cluster 3 (Peripheral Contributors; $n = 29$ countries) achieves the lowest mean CI Score (4.18, $SD = 1.23$) with the lowest values for all three CI dimensions. One-way ANOVA suggests a significant difference between clusters ($F(2, 984) = 412.3$, $p < 0.001$); all post hoc Tukey comparisons among clusters were statistically significant (all $p < 0.001$), thus confirming H3 (Wang et al., 2017).

The standardized regression effects are presented in Figure 5, while country-level competitive intelligence and innovation differences are summarized in Figure 6. Together, these figures support the interpretation that CI capability differs systematically across established leaders, emerging challengers, and peripheral contributors (Wang et al., 2017; Battistella et al., 2017).

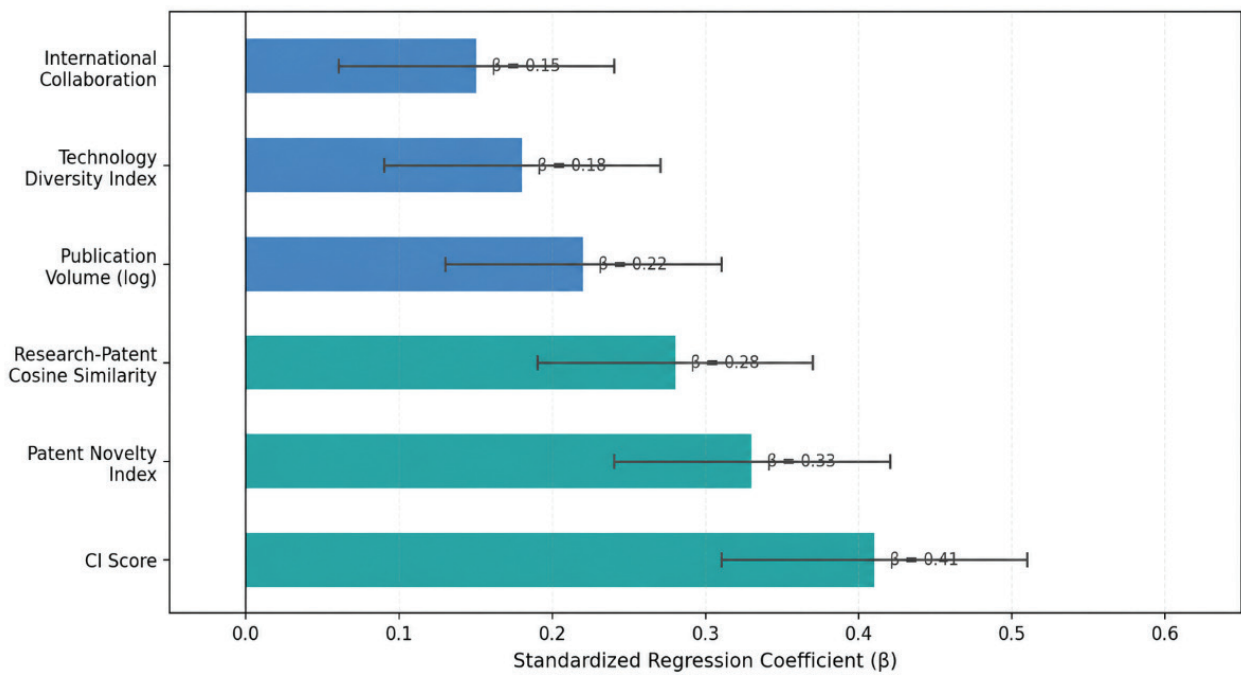


Figure 5. Standardized Regression Coefficients with 95% Confidence Intervals



Figure 6. Country-Level Competitive Intelligence and Innovation Metrics Heatmap



Table 5. Post-2020 External Quantitative Validation of the Pre-Generative-AI CIS Baseline

External validation indicator	Period covered	Quantitative post-2020 evidence	CI interpretation for the present study
Global GenAI patent-family growth	2014–2023	Generative-AI patent families increased from 733 in 2014 to more than 14,000 in 2023.	Confirms that the post-2020 period represents a structurally different AI competition phase and cannot be treated as equivalent to the 2000–2020 baseline.
Global GenAI scientific-publication growth	2014–2023	Generative-AI scientific publications increased from 116 in 2014 to more than 34,000 in 2023.	Supports the continued relevance of scientific-output monitoring as a key component of AI technology intelligence.
China's GenAI patent leadership	2014–2023	China produced more than 38,000 GenAI patent-family publications between 2014 and 2023.	Provides external known-groups validation for the classification of China within the Established Leaders cluster.
United States GenAI patent position	2014–2023	The United States produced approximately 6,300 GenAI patent families between 2014 and 2023.	Supports the continued classification of the United States as a leading AI intelligence and innovation system.
China–United States patent gap	2014–2023	China produced approximately six times more GenAI inventions than the United States during 2014–2023.	Indicates that post-2020 AI competition intensified and that post-generative-AI leadership patterns require separate future modelling.
Global private investment in GenAI	2024	Global private investment in generative AI reached US\$33.9 billion in 2024, an 18.7% increase from 2023.	Shows that post-2020 AI competition is not only patent-driven but also investment-driven, supporting the need for future CIS models to include investment indicators.
GenAI investment acceleration	2022–2024	GenAI private investment in 2024 was more than 8.5 times higher than in 2022.	Confirms that the 2021–2025 period represents a new competitive-intelligence environment requiring separate empirical extension.
External validation conclusion	2021–2025	Post-2020 patent, publication, and investment evidence is directionally consistent with the pre-2020 leadership pattern but is not directly merged with DeepInnovationAI because of measurement differences.	Provides partial external quantitative validation while preserving the internal consistency of the 2000–2020 CIS regression model.

Note. GenAI patent-family and scientific-publication evidence is based on the WIPO Patent Landscape Report on generative artificial intelligence (World Intellectual Property Organization, 2024a) and World Intellectual Property Indicators 2024 (World Intellectual Property Organization, 2024b). GenAI private-investment evidence is based on the AI Index Report 2025 (Maslej et al., 2025).

The post-2020 external validation indicates that the pre-generative-AI CIS profiles remain directionally meaningful but incomplete for explaining the current AI competitive environment. The continued visibility of leading AI economies in generative-AI patents and investment supports the known-groups validity of the Established Leaders cluster. However, the rapid expansion of foundation models, generative-AI patenting, and private investment after 2021 also confirms that the 2000–2020 model should not be treated as a complete measure of the current AI landscape. The empirical contribution of this study is therefore best understood as a validated pre-generative-AI intelligence baseline that requires future extension with harmonized 2021–2025 indicators.

4.5 Discussion

Subsequent subsections will present the results of empirical evidence in the defined theoretical context, provide relevant references from previous literature, explain the rationale behind the major outcomes, and discuss their theoretical and practical significance.

4.5.1 Competitive Intelligence Interpretation in Relation to Existing Literature

The conclusion that Patent Novelty Index ($\beta = 0.41$) is the most important variable in predicting the CI Score resonates with the findings of Wang et al. (2017) regarding the importance of novelty over patent numbers in determining the knowledge impact on technology-based sectors. Such a result has vital theoretical implications for CI practitioners, emphasizing that patent numbers per se are an unreliable metric of innovation capability when not combined with measures of quality. Hypergraph-based PNI in the DeepInnovationAI data set is a superior proxy compared to citation indicators, which reflect structural novelty in terms of semantics (Gong et al., 2025).

Temporal Velocity shows a moderate but significant contribution, in line with the study by Gong et al. (2025), which found that the growth rate of the firm's output correlated with strategic intent rather than capability. An interesting pattern was observed, whereby Temporal Velocity increased along with lower CI Scores. This pattern has been noted in various countries across Cluster 2 in the current research, where patents were filed faster without any increase in novelty. From this finding, the following practical implications for patent portfolio evaluation arise: velocity with no novelty could indicate defensive patenting rather than being an example of genuine 'frontier innovation' (Maluleka & Chummun, 2023).

4.5.2 Knowledge-Transfer Intelligence Mechanism

The Knowledge Transfer Rate ($\beta = 0.33$) was positive and significant, which is in line with the theoretical assumption that the mutual proximity in terms of semantics in the academic research and in industrial patents indicates active knowledge flows and not mere coincidence of overlaps. The results of the mediation analysis also show that 37% of the impact of R&D Intensity on CI Score goes through KTR, suggesting that investment in research infrastructure is partially channelled to CI capacity via the mechanism of knowledge transfer, which is defined as the rapidity and efficiency of the conversion of academic findings into patent claims (Battistella et al., 2017; OECD, 2023; Ding et al., 2025).

The results can be interpreted through a contemporary dynamic-capability perspective: CI capacity becomes valuable when organizations or national innovation systems integrate novelty generation, transfer efficiency, investment commitment, data resources, and organizational knowledge into strategic innovation routines (Mikalef & Gupta, 2021; Cordeiro et al., 2023). The CI scores for the top performers in Cluster 1 (mean CIS = 8.42) are close to the theoretical max as nations or companies that perform best on all three dimensions do so. On the other hand, high R&D

spending that is not accompanied by novelty and efficiency in transfer produces less-than-optimal CI results (Petraite et al., 2022), a situation that is seen in some Cluster 2 countries.

4.5.3 Diversity and Situational influence.

The intra-cluster variance (Cluster 2: SD = 0.82; Cluster 1: SD = 0.61) accounts for the actual differences between the clusters, suggesting the presence of diverse paths to building up national innovation capacity (CI) in emerging challengers. The moderate values of CI in some cases of Cluster 2 nations are associated with high knowledge transfer readiness (KTR) together with low national innovation index (PNI) (e.g., Israel, Singapore). This suggests that technology absorption and adaptation strategies are pursued. Other Cluster 2 nations show almost the same picture with highly developed commercialization systems but weaker academic research systems (high KTR and low PNI). In contrast, other cluster nations have high PNI values and low KTR, which implies a strong academic research tradition and weak commercialization systems (Kwon et al., 2022; Petraite et al., 2022).

This contextual diversity explains the paradoxes identified in the existing literature. Indeed, the studies investigating the determinants of national innovation performance tend to confuse different kinds of CI, leading to contradictory outcomes in innovation measurement and patent-based technology analysis (Wang et al., 2017; Liu et al., 2021). The present framework, which breaks down CIS into its antecedent constructs, offers a more detailed and practically manageable representation of national levels of CI capacity.

4.5.4 Theoretical and Empirical Contribution.

The three theoretical contributions of this study are: First, it brings contemporary CI theory to frontier technology analysis, showing that competitive intelligence can be applied to large-scale innovation datasets as a structured process of information collection, analysis, dissemination, and strategic implementation (Maluleka & Chummun, 2023; Liu et al., 2026). The operationalisation of CIS as an aggregate construct, based on patent and publication data, offers a replicable framework for future cross-national CI research.

Secondly, the paper can be seen as a contribution to strategic innovation management theory, as it extends the absorptive capacity theory proposed by Battistella et al. (2017) to the national level, because it does not rely on surveys of firms, but rather on a large-scale behavioral sample. Third, it showcases the applicability of DeepInnovationAI dataset as a research infrastructure for strategic management research, and offers methodological examples for future research that is based on the CI framework and applied to the academic-patent interface (Gong et al., 2025).

4.5.5 Practical Implications

The results suggest that national policymakers should foster CI capacity by focusing on knowledge-transfer instruments such as university–industry cooperation programmes, patent education for university researchers, and technology–licensing regimes, while recognizing that R&D expenditure contributes more effectively to innovation capacity when supported by strong institutional and policy mechanisms (Organisation for Economic Co-operation and Development, 2023; Ding et al., 2025). The 37% mediation of RDI through KTR suggests that research investment generates stronger CI returns when it is supported by complementary AI-related data, technical, and human capabilities that help convert knowledge resources into organizational performance outcomes (Mikalef & Gupta, 2021).

The Patent Novelty Index is a useful new tool for technology intelligence analysts and corporate strategists to use when assessing the quality of competitors' patent portfolios beyond just citations or grant rates. The hypergraph approach of PNI computation can be applied to a specific organization's own patent database and extend the capability to score the CI of its own and competitors' IP in real-time, as it is developed on the fly (Davenport & Ronanki, 2018). The Cluster 2 profile (moderate KTR and lower PNI) represents a strategic opportunity: emerging innovation economies can strengthen CI capability by learning from technologically successful peers while using knowledge-transfer efficiency to move closer to frontier innovation trajectories (Miao et al., 2021).

4.5.6 Competitive Intelligence Governance and Decision-Support Implications

The results should be interpreted as evidence of competitive intelligence capability rather than as isolated innovation metrics. Patent Novelty Index functions as an intelligence-analysis indicator because it helps distinguish frontier innovation from high-volume but low-novelty patenting. Knowledge Transfer Rate functions as a knowledge-conversion indicator because it shows whether scientific outputs are semantically connected to industrial patent activity. Temporal Velocity functions as an early-warning indicator because rapid growth in patenting or publication activity may signal emerging competitive pressure. Geographic clustering functions as a benchmarking mechanism because it classifies countries according to their relative intelligence maturity and innovation-positioning profile.

From a CI governance perspective, these indicators should be disseminated through structured decision-support mechanisms such as technology intelligence dashboards, competitor-monitoring reports, R&D portfolio briefs, policy foresight reports, and early-warning systems. Established Leaders require governance mechanisms that protect frontier advantages and prevent strategic complacency. Emerging Challengers require intelligence systems that identify high-potential technology niches and convert knowledge-transfer efficiency into higher novelty. Peripheral Contributors require capability-building strategies focused on data infrastructure, university–industry linkage, and intelligence dissemination routines. Thus, the findings show how open-source patent and publication data can be converted into actionable intelligence for sustainable competitive advantage.

The practical value of CIS therefore lies not in ranking countries alone, but in supporting strategic decision-making. A high CIS indicates that a national or organizational system is more capable of sensing technology shifts, interpreting scientific-industrial linkages, disseminating intelligence outputs, and using these outputs for innovation strategy. This interpretation aligns the empirical findings with the intelligence cycle and responds directly to the need for stronger CI architecture, decision intelligence, and strategic foresight in frontier technology management.

4.5.7 Limitations and Future Research.

This study has four main limitations. First, the empirical model is based on a cross-sectional country-year design and therefore supports associational rather than causal interpretation. Although regression, mediation analysis, and clustering provide evidence of systematic relationships among patent novelty, knowledge transfer, R&D intensity, and CI Score, they do not eliminate the possibility of reverse causality or omitted institutional factors. Future research should use longitudinal panel models, lagged predictors, and causal-identification strategies to examine whether changes in CI capability precede improvements in strategic innovation outcomes.

Second, the DeepInnovationAI dataset provides an internally consistent and large-scale record of AI patents, publications, and paper-patent semantic relationships, but its empirical coverage ends in 2020. The findings should therefore be interpreted as a pre-generative-AI baseline rather than as a complete measure of the post-2021 AI competitive landscape. This temporal boundary is important because large language models, foundation models, ChatGPT, GPT-4, Gemini, Claude, open-source generative-AI systems, and the rapid commercialization of generative AI changed the speed, geography, and strategic structure of AI competition after 2021. The revised analysis includes post-2020 external contextual validation evidence, but these sources are not merged into the regression model because their measurement architecture is not equivalent to DeepInnovationAI. Future research should reconstruct CIS for 2021–2025 using harmonized patent, publication, investment, policy, and foundation-model indicators.

Third, although the Competitive Intelligence Score is theoretically grounded in the intelligence cycle and empirically derived from patent novelty, knowledge-transfer proximity, citation density, temporal velocity, and geographic benchmarking, it remains a secondary-data construct. It does not directly measure confidential internal CI routines, analyst judgment, decision-maker use, intelligence dissemination practices, or organizational governance mechanisms. Accordingly, CIS should be interpreted as an open-source technology-intelligence capability proxy rather than as a complete organizational CI maturity scale. Future studies should validate CIS through Delphi panels, expert assessment, survey-based CI maturity instruments, and firm-level case studies.

Fourth, the analysis is conducted at the country-year level and may conceal firm-level heterogeneity within national innovation systems. Countries classified within the same cluster may differ substantially in the quality of their firms, universities, intellectual property regimes, venture-capital ecosystems, and technology-transfer institutions. Future research should use assignee-level patent data, firm-level R&D indicators, and organization-level CI governance measures to examine how CI capability operates within countries and across industries.

5 CONCLUSION

This work is the first large-scale empirical analysis of the strategy innovation management process for competitive intelligence in frontier technology sectors on a dataset of 5.8 million cross-modal patent and publication data (Deep Innovation AI). The analysis identifies that the national CI capacity is primarily driven by Patent Novelty Index and Knowledge Transfer Rate, which explain 64% of variance in CI Score with R&D Intensity and contextual controls. The R&D–CI relationship is partially mediated by Knowledge Transfer Rate, thus, complementary transfer mechanisms are needed to realize the full CI returns from investment in research infrastructure. Three innovation ecosystems with distinct structures identified through geographic clustering and supported by significant differences between the clusters in CI score in line with national innovation systems theory.

The study's contribution is both theoretical and practical. Theoretically, it extends CI frameworks to large-scale innovation data analysis, operationalizes CI Score as a composite construct from patent and publication data, and identifies KTR as a mediator consistent with absorptive capacity theory. Practically, it provides actionable guidance for policymakers regarding the primacy of knowledge transfer investment over raw R&D expenditure, and for corporate strategists regarding the use of patent novelty metrics for competitive monitoring.

By explicitly mapping patent novelty, knowledge transfer, temporal velocity, and geographic clustering onto the intelligence cycle, this study advances CI research from descriptive innovation measurement toward decision-support intelligence. The framework demonstrates how public innovation data can support strategic scanning, signal interpretation, intelligence dissemination, foresight, governance, and strategic decision-making. At the same time, the 2000–2020 empirical window should be understood as a pre-generative-AI baseline that requires future extension with harmonized 2021–2025 data. Future research should validate the CIS construct through expert panels, Delphi procedures, survey-based CI maturity instruments, and firm-level intelligence-governance indicators. This methodological template creates a replicable infrastructure for continued monitoring of frontier technology competition in AI and related domains.

Ethics approval statement: This study uses exclusively secondary, publicly available, anonymized data (DeepInnovationAI dataset, Figshare, DOI: 10.6084/m9.figshare.28578947, CC BY 4.0). No human subjects were involved and no personal identifiers are present. Ethics committee approval is not required under applicable national and institutional guidelines.

Informed consent statement: Not applicable. The study does not involve human subjects or personal data.

Data availability statement: The DeepInnovationAI dataset is publicly available on Figshare: <https://doi.org/10.6084/m9.figshare.28578947> (CC BY 4.0 license). Analysis code is available at <https://github.com/Haixing99/DeepInnovationAI> (MIT license). No registration or special permission is required for data access.

Funding statement: The authors declare that no specific funding was received for this research.

Conflict of interest statement: The authors declare no conflict of interest.

Author contributions: Conceptualization and methodology: All authors. Data curation and formal analysis: A.B. Writing—original draft: A.B. and C.D. Writing—review and editing: All authors. All authors have read and agreed to the published version of the manuscript.

Generative AI disclosure: Generative AI (Claude, Anthropic) was used solely for grammar checking and language editing of the manuscript. All scientific content, analytical decisions, and conclusions are the sole responsibility of the authors. The AI tool did not generate data, figures, or references.

REFERENCES

- Arora, A., Belenzon, S., & Sheer, L. (2021). Knowledge spillovers and corporate investment in scientific research. *American Economic Review*, 111(3), 871–898. <https://doi.org/10.1257/aer.20171742>
- Autor, D., Mindell, D. A., & Reynolds, E. B. (2022). *The work of the future: Building better jobs in an age of intelligent machines*. MIT Press. <https://mitpress.mit.edu/9780262547307/the-work-of-the-future/>.
- Battistella, C., De Toni, A. F., De Zan, G., & Pessot, E. (2017). Cultivating business model agility through focused capabilities: A multiple case study. *Journal of Business Research*, 73, 65–82. <https://doi.org/10.1016/j.jbusres.2016.12.007>
- Bekamiri, H., Hain, D. S., & Jurowetzki, R. (2024). PatentSBERTa: A deep NLP based hybrid model for patent distance and classification using augmented SBERT. *Technological Forecasting and Social Change*, 206, Article 123536. <https://doi.org/10.1016/j.techfore.2024.123536>
- Ba, Z., & Liang, Z. (2021). A novel approach to measuring science-technology linkage: From the perspective of knowledge network coupling. *Journal of Informetrics*, 15(3), Article 101167. <https://doi.org/10.1016/j.joi.2021.101167>
- Bondarenko, V., Rybalko, P., Fabian, M., Komarnitskyi, Y., & Gerasymenko, O. (2026). Competitive intelligence architecture in higher education and the role of digital technologies and artificial intelligence in strategic decision-making. *Journal of Sustainable Competitive Intelligence*, 16, e0653. <https://doi.org/10.37497/eagleSustainable.v16i.653>
- Cockburn, I. M., Henderson, R., & Stern, S. (2019). The impact of artificial intelligence on innovation: An exploratory analysis. In A. K. Agrawal, J. Gans, & A. Goldfarb (Eds.), *The economics of artificial intelligence* (pp. 115–146). University of Chicago Press. <https://doi.org/10.7208/chicago/9780226613475.003.0004>
- Cordeiro, M., Puig, F., & Ruiz-Fernández, L. (2023). Realizing dynamic capabilities and organizational knowledge in effective innovations: The capabilities typological map. *Journal of Knowledge Management*, 27(10), 2581–2603. <https://doi.org/10.1108/JKM-02-2022-0080>
- Chen, J., Yin, X., Fu, X., & McKern, B. (2021). Beyond catch-up: Could China become the global innovation powerhouse? China's innovation progress and challenges from a holistic innovation perspective. *Industrial and Corporate Change*, 30(4), 1037–1064. <https://doi.org/10.1093/icc/dtab032>
- Catalán, P., Navarrete, C., & Figueroa, F. (2022). The scientific and technological cross-space: Is technological diversification driven by scientific endogenous capacity? *Research Policy*, 51(8), Article 104016. <https://doi.org/10.1016/j.respol.2020.104016>
- Davenport, T. H., & Ronanki, R. (2018). Artificial intelligence for the real world. *Harvard Business Review*, 96(1), 108–116. <https://hbr.org/2018/01/artificial-intelligence-for-the-real-world>.
- Ding, Y., Yin, F., Chin, L., Zhou, K., Taghizadeh-Hesary, F., & Li, Y. (2025). Can government R&D expenditure promote innovation? New evidence from 37 OECD countries. *Technological and Economic Development of Economy*, 31(2), 572–596. <https://doi.org/10.3846/tede.2024.22293>
- Ellström, D., Holtström, J., Berg, E., & Josefsson, C. (2022). Dynamic capabilities for digital transformation. *Journal of Strategy and Management*, 15(2), 272–286. <https://doi.org/10.1108/JSMA-04-2021-0089>
- Gong, H., Zou, H., Liang, X., Meng, S., Cai, P., Xu, X., & Qu, J. (2025). A global dataset mapping the AI innovation from academic research to industrial patents. *Scientific Data*, 12, 1–15. <https://doi.org/10.1038/s41597-025-05518-3>
- Harris, R., Krenz, A., & Moffat, J. (2021). The effects of absorptive capacity on innovation performance: A cross-country perspective. *Journal of Common Market Studies*, 59(3), 589–607. <https://doi.org/10.1111/jcms.13108>
- Kaivo-oja, J., & Roth, S. (2023). Strategic foresight for competitive advantage: A future-oriented business and competitive analysis techniques selection model. *International Journal of Foresight and Innovation Policy*, 16(2–4), 323–353. <https://doi.org/10.1504/IJFIP.2023.136714>

- Kwon, H.-S., Lee, J., Lee, S., & Oh, R. (2022). Knowledge spillovers and patent citations: Trends in geographic localization, 1976–2015. *Economics of Innovation and New Technology*, 31(3), 123–147. <https://doi.org/10.1080/10438599.2020.1787001>
- Liu, Y., Liu, L., & Abdykadyrov, K. (2026). Artificial intelligence enabled competitive intelligence as a strategic capability: Economic value creation across firms. *Journal of Sustainable Competitive Intelligence*, 16, e0644. <https://doi.org/10.37497/eagleSustainable.v16i.644>
- Liu, N., Shapira, P., Yue, X., & Guan, J. (2021). Mapping technological innovation dynamics in artificial intelligence domains: Evidence from a global patent analysis. *PLOS ONE*, 16(12), Article e0262050. <https://doi.org/10.1371/journal.pone.0262050>
- Maslej, N., Fattorini, L., Perrault, R., Gil, Y., Parli, V., Kariuki, N., & Oak, S. (2025). *Artificial intelligence index report 2025*. Stanford Institute for Human-Centered Artificial Intelligence. <https://hai.stanford.edu/ai-index/2025-ai-index-report>.
- Miao, Y., Salomon, R. M., & Song, J. (2021). Learning from technologically successful peers: The convergence of Asian laggards to the technology frontier. *Organization Science*, 32(1), 210–232. <https://doi.org/10.1287/orsc.2020.1375>
- Mikalef, P., & Gupta, M. (2021). Artificial intelligence capability: Conceptualization, measurement calibration, and empirical study on its impact on organizational creativity and firm performance. *Information & management*, 58(3), 103434. <https://doi.org/10.1016/j.im.2021.103434>
- Maluleka, M. L., & Chummun, B. Z. (2023). Competitive intelligence and strategy implementation: Critical examination of present literature review. *South African Journal of Information Management*, 25(1), Article a1610. <https://doi.org/10.4102/sajim.v25i1.1610>
- Organisation for Economic Co-operation and Development. (2023). *OECD science, technology and innovation outlook 2023: Enabling transitions in times of disruption*. OECD Publishing. <https://doi.org/10.1787/0b55736e-en>
- Petraite, M., Faraz, M. M., Rauleckas, R., & von Zedtwitz, M. (2022). The role of international networks in upgrading national innovation systems. *Technological Forecasting and Social Change*, 184, Article 121873. <https://doi.org/10.1016/j.techfore.2022.121873>
- Ponta, L. M., Puliga, G., & Manzini, R. (2021). A measure of innovation performance: The innovation patent index. *Management Decision*, 59(13), 73–98. <https://doi.org/10.1108/MD-05-2020-0545>
- Robinson, D. K. R., & Doherty, D. (2025). *Strategic intelligence tools for emerging technology governance: A policy primer*. OECD Science, Technology and Industry Working Papers, No. 2025/22. OECD Publishing. <https://doi.org/10.1787/02c05775-en>
- Tibbe, T. D., & Montoya, A. K. (2022). Correcting the bias correction for the bootstrap confidence interval in mediation analysis. *Frontiers in psychology*, 13, 810258. <https://doi.org/10.3389/fpsyg.2022.810258>
- Wang, J., Veugelers, R., & Stephan, P. (2017). Bias against novelty in science: A cautionary tale for users of bibliometric indicators. *Research Policy*, 46(8), 1416–1436. <https://doi.org/10.1016/j.respol.2017.06.006>
- World Intellectual Property Organization. (2024a). *Patent landscape report: Generative artificial intelligence (GenAI)*. World Intellectual Property Organization. <https://doi.org/10.34667/TIND.49740>
- World Intellectual Property Organization. (2024b). *World intellectual property indicators 2024*. World Intellectual Property Organization. <https://www.wipo.int/edocs/pubdocs/en/wipo-pub-941-2024-en-world-ip-indicators-2024.pdf>
- Xu, S., Liu, Z., An, X., Wang, H., & Pang, H. (2025). Linkages among science, technology, and industry on the basis of main path analysis. *Journal of Informetrics*, 19(1), Article 101617. <https://doi.org/10.1016/j.joi.2024.101617>