A Work Project, presented as part of the requirements for the Award of a Master's degree in International Development and Public Policy from the Nova School of Business and Economics.

DETERMINANTS OF AI COLLABORATIONS, PATENTS, AND INVESTMENTS: A GRAVITY EQUATION APPROACH TOWARDS POLICY RECOMMENDATIONS FOR HUMAN DEVELOPMENT

RICHARD STEINERT

Work project carried out under the supervision of:

José Tavares

20/01/2025

Abstract

This work investigates the determinants of AI-related collaborations, patents, and investments using a gravity equation approach, addressing critical gaps in understanding the development phase of AI. Analyzing panel data spanning 2013-2022 across 176 countries, the findings reveal that economic size, geographic proximity, and development levels significantly influence AI activity. While developed countries dominate across all observation metrics in absolute numbers, developing regions show notable growth in research collaborations. Targeted policy interventions are proposed to bridge disparities, strengthen knowledge creation, and foster inclusive AI ecosystems. The results emphasize the strategic importance of the development phase in achieving equitable human development outcomes.

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Professor José Tavares, for his invaluable guidance and support throughout this thesis. His expertise, genuine interest in the subject, and continuous feedback were instrumental in shaping this work and ensuring its academic rigor. I am equally grateful to José Antonio Reyes González from the United Nations Human Development Report Office for his professional guidance and mentorship throughout this research collaboration. His expertise in policy research, international development, and network analysis has greatly enriched this thesis. I am also deeply thankful for the opportunity to contribute to the upcoming Human Development Report 2025, an invaluable learning experience that broadened my expertise and further enhanced the quality of this work.

Keywords

Artificial Intelligence, AI Determinants, Gravity Equation, Inclusive Development, Global Governance, Policy

This work used infrastructure and resources funded by Fundação para a Ciência e a Tecnologia (UID/ECO/00124/2013, UID/ECO/00124/2019 and Social Sciences DataLab, Project 22209), POR Lisboa (LISBOA-01-0145-FEDER-007722 and Social Sciences DataLab, Project 22209) and POR Norte (Social Sciences DataLab, Project 22209).

Table of Contents

Abstract 1
Acknowledgements 1
Keywords 1
Table of Contents 2
- **1. Introduction** 3
- **2. Theoretical Framework** 3
  - 2.1. AI's Two-Stage Lifecycle 3
    - 2.1.1. The Development Phase 4
    - 2.1.2. The Deployment Phase 5
  - 2.2. Mitigating Biases: Why Early Intervention Matters 6
    - 2.2.1. Origins of Bias 6
    - 2.2.2. Amplification of Biases during Deployment 7
    - 2.2.3. The Necessity of Early Intervention 7
  - 2.3. Global Divide in AI Development and Deployment 8
    - 2.3.1. Infrastructure and Investment Disparities: The Compute Divide 9
    - 2.3.2. Bridging the Divide: A Development-Phase Perspective 10
- **3. Data and Methodology** 11
  - 3.1. Data Collection 11
  - 3.2. Data Description 11
  - 3.3. The Gravity Equation Approach 11
  - 3.4. Descriptive Trends and Regional Clusters 13
- **4. Results and Analysis** 13
  - 4.1. Cluster-based Insights 13
  - 4.2. Regression Analysis 16
    - 4.2.1. Intellectual Outcomes 16
    - 4.2.2. Material Outcomes 18
    - 4.2.3. Quantifying the Impact of HDI on Dependent Variables 19
  - 4.3. Policy Implications 21
- **5. Limitations and Further Research** 23
- **6. Conclusion** 24

Appendix 26
List of Figures: 30
List of Tables: 34
List of References 38

1. Introduction

Artificial Intelligence (AI) is transforming both digital and physical environments, shaping user behavior, market trends, and societal outcomes (OECD 2024a; Vinuesa et al. 2020). However, these benefits are unevenly distributed due to disparities in resources for AI development and deployment, presenting significant challenges for global human development (Lehdonvirta, Wu, and Hawkins 2023; United Nations 2024). Given AI's rapid pace of change, there is an urgent need for action to ensure its benefits are equitably shared.

While much of the literature focuses on the deployment phase of AI, this study addresses a critical research gap by focusing on the development phase - a formative stage where interventions are particularly impactful and cost-efficient (Wachter et al. 2016; Shestakova 2021). Understanding the determinants of AI development, including knowledge creation, innovation (patents), and investment flows, is essential to inform effective policy responses that can mitigate systemic inequalities.

To achieve this, the study employs the gravity equation framework to analyze the economic, developmental, and geographic factors driving AI activity. Using comprehensive datasets spanning 2013–2022, the study quantifies the determinants of global AI collaborations, patents, and investments. Based on this analysis, it proposes evidence-based policy recommendations aimed at tackling disparities and fostering inclusive AI development, ensuring that AI's benefits contribute to broader human development.

2. Theoretical Framework

2.1. AI's Two-Stage Lifecycle

One can conclude from this that AI should be understood as a dynamic system that not only interacts with but also influences its environment. To further illustrate this interactive relationship, the AI lifecycle can be divided into two stages: development and deployment (OECD 2024a).

2.1.1. The Development Phase

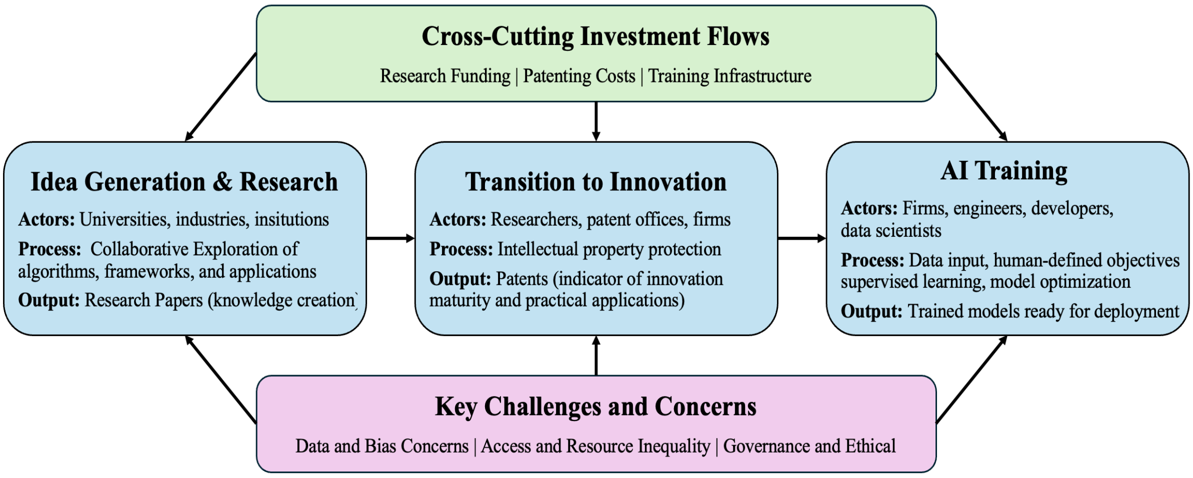

Figure 1: The AI Development Phase: Processes, Outputs, and Challenges

The development phase of AI, much like other scientific and technological research fields, begins with generating ideas and conducting research, where universities, industries, and specialized institutions collaborate to explore algorithms, frameworks, and applications. These efforts culminate in research papers, which not only disseminate knowledge but also shape discourse and influence future advancements (Russell and Norvig 2020; OECD 2017). Such knowledge creation is not merely an academic exercise but a fundamental driver of technological evolution, laying the groundwork for subsequent stages of AI development. As research transitions toward specific innovations, researchers seek to protect their intellectual property through patents. Hence, patents mark a critical transition from theoretical advancements to practical and protectable applications. One can conclude from this that patents not only serve as an indicator of innovation maturity but also bridge the gap between abstract concepts and real-world usability (WIPO 2022; Hall and Harhoff 2012).

The next stage known as "training", is where AI's foundational "intelligence" is built (Russell and Norvig 2020). This phase involves providing the system with inputs, such as data and human-specified objectives. This is in line with Mökander et al. (2023) study, which shows the importance of supervised learning techniques, where training data is labeled with correct answers, which allows AI systems to learn by example, refining their performance iteratively. This training process is increasingly resource-intensive, necessitating significant financial investments. For instance, optimizing and training advanced computational models demands powerful hardware, vast quantities of labeled data, and rigorous testing frameworks, all of which require substantial funding. As outlined earlier, human involvement in this phase is crucial for defining objectives, selecting datasets, and designing the architecture of the AI system (ITU 2024); however, this is also where one primary concerns lays. Namely, the system often absorbs the values, biases, and socio-political perspectives embedded within the data and goals set by its developers (Leavy, O'Sullivan, and Siapera 2020). Since these development efforts are often centered in high-income regions, AI systems may unintentionally perpetuate regional or cultural biases, which can be amplified as they are deployed globally. For instance, Vinuesa et al. (2019) highlighted that AI-based healthcare diagnostic systems which are trained primarily on medical data from developed countries may overlook or misdiagnose conditions like malaria or tuberculosis, which are more prevalent in developing regions. The role of local knowledge creation and investments in ensuring equitable representation and diversity in datasets is thus crucial for mitigating these risks. This lack of region-specific data not only reduces diagnostic accuracy in underrepresented areas but may also influence the conclusions drawn in analyses like this one, particularly for lower-HDI countries. The absence of localized data can skew results, as AI systems or analyses might over-rely on data from high-income regions.

2.1.2. The Deployment Phase

Once trained and tested, AI systems enter the deployment phase, where they begin to operate in real-world or simulated environments. It is in this phase that AI systems begin to influence their environments, a defining characteristic that sets AI apart from traditional automation (European Commission 2018). Furthermore, the deployment phase introduces the adaptability of AI systems. Unlike automation, which follows fixed rules, AI systems continue to evolve through a process known as post-deployment adaptation (Grobelnik, Perset, and Russell 2024) - another distinguishing characteristic. Because of this continuous adaptation, AI has far-reaching effects on both digital and physical environments, shaping user behavior, market trends, and societal outcomes (ITU 2024; Vinuesa et al. 2019). In digital environments, for example, personalized algorithms influence consumer preferences and behaviors, while in physical environments, autonomous AI systems - such as drones or self-driving cars - alter the logistics and transportation sectors (Vinuesa et al. 2019). While this adaptability drives innovation, it also introduces significant ethical and regulatory challenges. For instance, the OECD (2024a) and the European Commission (2018) highlight how AI systems can evolve in unexpected ways, raising concerns about accountability and oversight. As Floridi et al. (2018) point out, establishing robust oversight mechanisms and governance frameworks is crucial for ensuring that AI serves broader societal goals rather than inadvertently reinforcing existing inequalities.

2.2. Mitigating Biases: Why Early Intervention Matters

2.2.1. Origins of Bias

Building on the discussion in the previous section, a major concern with AI systems is the potential for biases, which can manifest in multiple forms:

Form of Bias	**Definition**
Dataset Bias	Underrepresentation of certain groups or regions
Algorithmic Bias	Prioritization inherently favors certain outcomes over others (influenced by the socio-political contexts in which these systems are developed).
Emergent Bias	Unexpected system-user interactions

Table 1: Sources of Biases (Ferrara 2024)

A notable example of dataset bias is provided by Obermeyer et al. (2019), who revealed how a widely used healthcare algorithm systematically underestimated the health needs of black patients compared to white patients with similar conditions. This disparity stemmed from the algorithm's reliance on healthcare costs - a socio-economically skewed proxy variable - as a measure of health. Such biases disproportionately harm underrepresented groups, including populations in low HDI countries. These regions are substantially underrepresented in training datasets, which amplifies the impact of bias by reducing diagnostic accuracy for region-specific health conditions, such as malaria or tuberculosis. In summary, as emphasized earlier, the development phase is foundational for shaping AI systems, making it a critical juncture where biases can originate and become embedded within the system (Ntoutsi et al. 2020).

2.2.2. Amplification of Biases during Deployment

When biases remain unchecked during the development phase, they are often magnified in the deployment phase. For instance, recommendation algorithms that prioritize popular content can further marginalize underrepresented voices, reinforcing existing disparities in visibility and access (Buolamwini and Gebru 2018). The risks of amplification are especially evident in AI-based language models. Adam et al. (2022) demonstrated that fine-tuning GPT-2 with just 2,000 biased sentences resulted in systematically biased outputs, propagating harmful stereotypes in user interactions. This example highlights how small biases in the development stage can cascade into large-scale societal impacts during deployment.

2.2.3. The Necessity of Early Intervention

Addressing biases during the development phase is not only more feasible but also significantly more cost-effective than intervening post-deployment. During the development stage, AI systems are in their formative phase, where datasets, algorithms, and model architectures are being selected and refined (ITU 2024). Hence, changes made at this stage, such as improving data diversity or integrating fairness constraints, are inherently easier to implement because they occur before the system's processes and behaviors become embedded and operationalized. Conversely, post-deployment modifications require altering established systems, retraining on larger datasets, and navigating the complexities of systems already integrated into real-world environments (Wachter et al. 2016). This is further reinforced by Shestakova (2021), which highlighted the operational hurdles in retroactively correcting biases in deployed facial recognition systems. Retraining the models with additional datasets post-deployment increased operational costs by 40% and required a complete overhaul of the model's existing infrastructure. However, integrating fairness constraints during the development stage only marginally increased upfront resource allocation without requiring costly post-release fixes. Similarly, Mitchell et al. (2019) demonstrated how early investments in representative datasets not only reduced downstream disparities by 25% but also avoided the logistical and ethical challenges of rectifying biases post-deployment. Adding to this evidence, O'Connor and Liu (2023) showcased the effectiveness of early interventions in natural language processing, where their debiasing algorithm reduced the presence of gender stereotypes in word embeddings from 19% to 6%. In summary, proactively ensuring inclusivity and fairness during development, prevents biases from embedding deeply into AI systems. Thus, intervening during the development phase is not only more practical but also cost-effective, reducing the need for costly corrections post-deployment and fostering more equitable outcomes from the start.

2.3. Global Divide in AI Development and Deployment

As illustrated in the previous section, ensuring diversity in datasets during the development phase is a critical step toward mitigating biases in AI systems. However, achieving this objective is fraught with challenges stemming from pronounced global disparities in resources and capabilities. High-income regions often possess the advanced infrastructure, investments, and datasets necessary to shape AI systems, while lower-income regions face significant barriers to participation. For instance, the United States hosts 19 times more top-tier data centers than India, the leading country among emerging markets (United Nations 2024). Similarly, developed countries outspend developing nations in AI research and development by wide margins, with global AI investments disproportionately concentrated - low-income regions receive only 1% of total digital investments, despite representing nearly 40% of the world's population (World Bank 2023). This stark imbalance not only highlights disparities in computational resources but also reflects the scale at which infrastructure and investment gaps translate into reduced capacity for AI innovation. These disparities reflect a broader global divide that influences not only AI development but also its deployment and adoption (Lehdonvirta, Wu, and Hawkins 2023).

2.3.1. Infrastructure and Investment Disparities: The Compute Divide

These disparities are further exemplified by the compute divide, which highlights the unequal distribution of critical computational resources necessary for AI development. Nations in the Compute North, such as the United States and China, dominate AI training and innovation largely due to their superior infrastructure, including advanced GPUs and cloud regions (Krishna 2024; Soare 2020). These countries collectively host nearly half of the global public cloud regions equipped with state-of-the-art hardware, such as Nvidia H100 GPUs, which are indispensable for training cutting-edge AI models (Lehdonvirta, Wu, and Hawkins 2023).

In contrast, Compute South nations, such as Brazil and South Africa, rely predominantly on legacy hardware like Nvidia V100 GPUs, limiting their capacity to engage in AI training and innovation. The disparity is even more pronounced in Compute Desert regions, where access to computational infrastructure is virtually nonexistent. These nations lack not only high-speed internet but also the institutional capacity to establish local AI ecosystems (Božić 2023; Lehdonvirta, Wu, and Hawkins 2023). Consequently, they remain reliant on external actors for AI solutions, perpetuating a cycle of technological dependence (Efe 2022). The implications of the global divide extend beyond infrastructure. Economic disparities compound the issue, with high-income nations capturing the majority of AI-driven economic benefits. For instance, AI is projected to contribute $13 trillion to global GDP by 2030, yet most of this growth is concentrated in advanced economies (McKinsey Global Institute 2018). The U.S. Air Force's predictive maintenance systems save $3 - $5 billion annually through AI-driven optimization, demonstrating the tangible economic advantages accessible to nations with robust AI ecosystems (Soare 2020). Such outcomes remain unattainable for regions lacking foundational infrastructure and investments (Božić 2023). Additionally, disparities in user adoption and digital literacy further marginalize underserved populations. In education, schools in resource-rich districts integrate AI-powered personalized learning platforms, whereas students in underfunded schools face barriers to accessing these technologies (Božić 2023; Krishna 2024). The following table summarizes the key activities and manifestations of the global divide across both the development phase and deployment phase:

Phase	Key Activities	Manifestation of the Global Divide
Development	Knowledge creation (research collaborations)	Dominance of high-income regions in research infrastructure and networks
	Innovation (patent filings and advancements)	Limited capacity for innovation due to funding and IP access barriers
	Infrastructure (such as computational resources)	Unequal access to GPUs, data centers, and training infrastructure
	Early-stage financing	Minimal investments in AI R&D in low-income regions
Deployment	AI tool adoption	Limited adoption due to infrastructure deficits and affordability issues
	Sectoral integration (e.g., healthcare, education)	Developed regions benefit from sector-wide AI solutions
	Regulatory and ethical implementation	Absence of robust governance frameworks in low-income regions
	Capacity building (skills and digital literacy)	Shortage of AI-trained professionals limits local deployment
	Data localization and relevance	Poorly contextualized AI systems reduce effectiveness

Table 2: Key Activities and Manifestations of the Global Divide

2.3.2. Bridging the Divide: A Development-Phase Perspective

While parts of the development phase, particularly computational capabilities for training and innovation, have been widely studied (Krishna 2024; Lehdonvirta, Wu, and Hawkins 2023), other defining aspects of this phase - such as knowledge creation, patent activity, and cross-border investments - have not received the same level of attention. However, as shown earlier, these factors are foundational to AI development, shaping the intellectual and material resources upon which AI systems are built. By addressing disparities at this early stage, stakeholders can influence the trajectory of AI ecosystems in a way that minimizes downstream inequalities and reduces the need for costly corrections during deployment. To effectively guide the trajectory of AI development, it is imperative to understand the driving determinants of AI during this formative phase. Accordingly, the subsequent quantitative analysis will examine the driving forces behind these critical factors during the development phase of AI.

3. Data and Methodology

3.1. Data Collection

The data for this study was sourced from the Country Activity Tracker (CAT), a tool developed by the Emerging Technology Observatory (ETO), a project of the Center for Security and Emerging Technology (CSET) at Georgetown University. The CAT tool encompasses metrics across three key dimensions: research, patents, and investments (ETO 2024b).[1]

3.2. Data Description

The research dataset comprises panel data on AI-related academic co-authorship across 176 countries over 2013–2022, covering 14,077 unique country pairs. These pairs are consistently distributed across the decade, ensuring robust analysis of international collaboration trends. The patent dataset tracks AI patents filed and granted across jurisdictions for 66 countries, with 7,830 observations over ten years. Its panel structure distinguishes between applications and grants, enabling insights into innovation stages and intellectual property evolution. The investment dataset spans 106 countries from 2013 to 2022, featuring 11,310 observations and 1,311 unique country pairs. By differentiating between incoming and outgoing investments, it provides a detailed view of financial flows within the global AI ecosystem.[2]

3.3. The Gravity Equation Approach

The gravity equation, originally inspired by Newton's law of gravity, has become a cornerstone in economic analysis for modeling bilateral interactions such as trade, migration, and investment flows (Anderson 2011). In its basic form, the equation posits that the volume of bilateral trade between two countries is directly proportional to the product of their economic

[1] Detailed information regarding the data collection is provided in Appendix A.
[2] Detailed Information regarding the data description and preprocessing are provided in Appendix B and C

masses (measured why GDP) and inversely proportional to the distance between them (Head and Mayer 2014). This conceptual framework reflects the intuitive notion that larger economies generate more trade, while greater distances pose higher transaction costs and barriers. Over time, the model has been refined into structural gravity equations, which account for multilateral resistance terms - factors that measure the relative accessibility of countries within a global economic network (Anderson and Van Wincoop 2003). These advancements have solidified the gravity equation's empirical reliability and extended its application beyond trade, encompassing areas like migration, foreign direct investment, and knowledge spillovers (Bergstrand and Egger 2013 ; Eaton and Kortum 2002).

In this study, the gravity equation framework is applied to analyze interactions in AI-related activities, including research collaborations, patent filings, and cross-border investments. These interactions, much like trade flows, involve measurable exchanges between countries that are influenced by economic size, geographic proximity, and development levels. Research collaborations often require shared resources and institutional partnerships, making economic capacity and spatial proximity key determinants of feasibility and frequency (Leamer and Levinsohn 1995; Keller 2002). Similarly, patent activity reflects the movement of ideas and innovation across borders, which is shaped by the economic and technological infrastructure of participating nations. Geographic proximity also plays a role in facilitating access to the advanced computational resources needed for training AI models, which is a crucial step in the innovation process leading to patent filings. It generates local externalities that enhance collaboration and resource sharing, further influencing cross-border patent activity and research outputs (Eaton and Kortum 2002). Investments, particularly private-market flows in emerging technologies like AI, may still be influenced by geographic distance despite the globalization of financial markets. Unlike public investments, private investments often require on-site visits, in-person scrutiny, and due diligence, which can be hindered by physical distance. Cultural and institutional differences between countries further heighten the role of proximity, as navigating these complexities often necessitates closer interactions to establish trust and mutual understanding. Furthermore, the use of geographic distance aligns with the model's theoretical foundations, reflecting the physical and institutional barriers that can influence the intensity of these interactions (Redding and Weinstein 2019). By adapting the gravity model to these dimensions of AI development, the study provides a new framework for analyzing the determinants of international knowledge creation, innovation, and financial flows.

3.4. Descriptive Trends and Regional Clusters

In addition to the regression analysis, this study employs a descriptive approach to provide a contextual understanding of trends in AI development by identifying disparities and growth patterns over the period from 2013 to 2022 across two primary dimensions: geographical regional clusters and HDI groupings. Regional clustering enables the study to explore geographic proximity and shared economic conditions, which often shape patterns of international collaboration and investment. HDI groupings provide an additional lens to assess disparities across different levels of socio-economic development, capturing how foundational capabilities influence AI-related activities. By focusing on trends over a full decade, this approach complements the regression analysis by uncovering broader patterns and disparities that may not be immediately apparent in statistical models.

4. Results and Analysis

4.1. Cluster-based Insights

This analysis reveals significant disparities in AI-related outcomes across regions and HDI groups. The following graph illustrates the growth across regions and HDI groups in terms of both absolute numbers and compound annual growth rates.

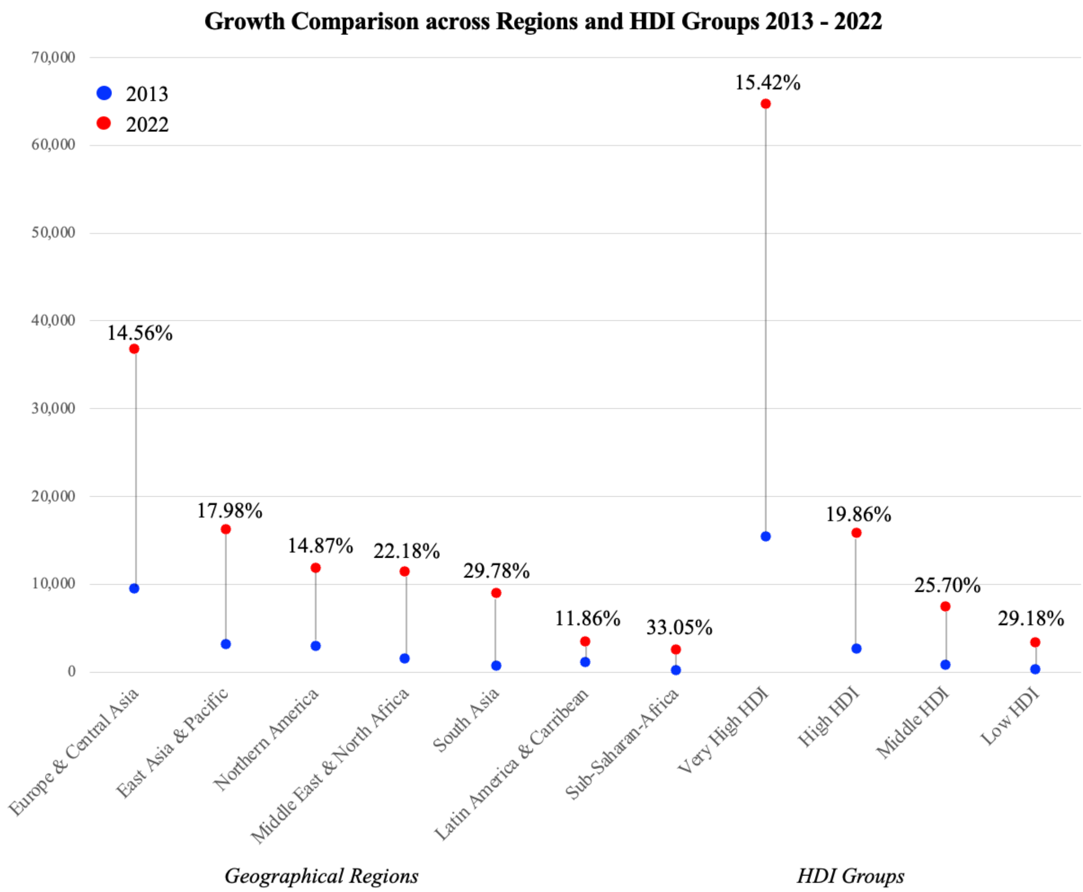

Figure 2: Growth Comparison across Regions and HDI Groups 2013-2022

While very high HDI countries continue to dominate in both co-authorship and patent activities, developing regions show notable growth, especially in co-authorship, where low and middle HDI countries experienced impressive compound annual growth rates (CAGR), with Sub-Saharan Africa leading at 33.05%. Similarly, middle HDI countries saw robust growth, with a CAGR of 25.70%. Regionally, Sub-Saharan Africa experienced the highest growth rate with a CAGR of 33.05% closely followed by South Asia with a CAGR of 29.78%, reflecting a marked increase in research activity. Despite starting from significantly lower baselines, these developing regions are catching up in terms of knowledge creation. This progress stands in stark contrast to the more modest growth in Europe & Central Asia (CAGR of 14.56%) and Northern America (14.87%), which, while still leaders in absolute numbers, appear to have reached a plateau with regard to publishing new articles.

Patent data, on the other hand, highlights a starkly different dynamic. The overwhelming majority of patent applications and grants are concentrated in very high HDI countries and regions such as East Asia & Pacific and Northern America, confirming the well-documented trend of patent activity being heavily concentrated in innovation hubs and major markets with advanced technological infrastructures and economic resources (Crescenzi et al. 2019; Paunov et al. 2019; Lengyel, Sebestyén, and Leydesdorff 2013). For instance, in 2022, very high HDI countries accounted for 3,694 granted patents, while high HDI countries followed with 36,682. Middle HDI countries and low HDI countries, however, reported no patent activity in the same year, underscoring the stark disparity. Regional analysis further supports this trend. In 2022, East Asia & Pacific and Northern America led patent grants with 39,108 and 741, respectively, compared to the near absence of patent activity in regions like Sub-Saharan Africa (140 granted patents in 2022) and South Asia (0 patents granted in 2022). While the co-authorship data suggests an encouraging trend of increasing collaborative research activity in developing regions, potentially narrowing the gap in knowledge creation, the findings from the patent data stand in sharp contrast to this. The lack of corresponding patent activity highlights significant barriers in translating knowledge into commercially viable technologies.

The investment data across different HDI groups highlights significant global disparities.[3] Low HDI countries are largely marginalized, receiving only $19 million from 2013 to 2022, mainly from very high HDI countries (47%). Middle HDI countries attracted a more substantial $28.86 billion, with very high HDI countries contributing 88% of this amount. Very high HDI countries overwhelmingly dominated both as recipients and sources of investments, attracting $340 billion, with 97% originating from within their own group. This intra-group investment emphasizes the self-reinforcing economic power and concentration of wealth among the highest HDI nations.

[3] Table 6 in the Appendix provides a detailed overview of investment patterns across HDI groups

Overall, very high HDI countries are the principal investors globally, while low HDI countries remain on the fringes of investment flows. Although there's a gradual increase in investments from middle HDI countries into their own group and into low HDI countries, the disparities remain vast. Addressing these inequalities necessitates targeted policies to promote inclusive growth and enhance access to financial resources for developing nations.

4.2. Regression Analysis

This section presents the regression results for the six dependent variables, grouped into intellectual and material outcomes. Each table reports the results for one of the dependent variables, with three specifications: the base specification, which includes GDP and distance; the second specification, which adds the Human Development Index (HDI) as an additional explanatory variable; and the third specification, which adds GDP per capita to the base specification.

4.2.1. Intellectual Outcomes

Consistent with the gravity equation, economic size, as measured by GDP, exhibits a positive and significant relationship with all outcomes. To complement this, geographic proximity, consistently shows a negative relationship with all outcomes. The negative coefficients for distance, ranging from -0.22 for co-authorship to -0.08 for investments, align with the gravity model's assertion that greater distance hinders cross-border AI interactions. Notably, co-authorship, a critical indicator of international collaboration in AI research, is strongly influenced by the development level of the primary country. A one-unit increase in the primary country's HDI leads to a 2.16% increase in co-authorship activity, compared to just 0.65% for a 1% rise in GDP. This significant difference highlights the crucial role of development, particularly in human capital and institutional capacity, in fostering collaborative research.

	CoAuthorship		
	(1)	(2)	(3)
	logArticles	logArticles	logArticles
log GDPPrimary	0.65 ***	0.58 ***	0.58 ***
	(.01)	(.01)	(.01)
logGDPSecondary	0.28 ***	0.23 ***	0.23 ***
	(.01)	(.01)	(.01)
logDistance	-0.22 ***	-0.13 ***	-0.15 ***
	(.01)	(.01)	(.01)
HDIPrimary		2.16 ***	
		(.09)	
HDISecondary		1.10 ***	
		(.12)	
logGDPpcPrimary			0.21 ***
			(.01)
logGDPpcSecondary			0.11 ***
			(.01)
_cons	-20.46 ***	-20.47 ***	-20.80 ***
	(.29)	(.28)	(.28)
Observations	8,667	8,667	8,667
R-squared	0.5389	0.5769	0.5738

t statistics in parenthes

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3: Regression Results Co-authorship

	Panel A: Patents applied			Panel B: Patents granted		
	(1)	(2)	(3)	(4)	(5)	(6)
	logPatents	logPatents	logPatents	logPatents	logPatents	logPatents
log GDPPrimary	1.00 ***	0.99 ***	0.98 ***	0.75 ***	0.75 ***	0.74 ***
	(.02)	(.02)	(.20)	(.02)	(.02)	(.022)
logGDPSecondary	0.30 ***	0.29 ***	0.29 ***	0.35 ***	0.36 ***	0.34 ***
	(.03)	(.03)	(.03)	(.03)	(.03)	(.03)
logDistance	-0.10 ***	-0.10 ***	-0.10 ***	-0.10 ***	-0.10 ***	-0.09 ***
	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
HDIPrimary		1.68 ***			1.47 ***	
		(.35)			(.35)	
HDISecondary		0.82 *			1.90 ***	
		(.40)			(.44)	
logGDPpcPrimary			0.14 ***			0.12 ***
			(.03)			(.03)
logGDPpcSecondary			0.10 *.			0.20 ***
			(.04)			(.04)
_cons	-32.86 ***	-34.68 ***	-34.50 ***	-28.17 ***	-31.24 ***	-30.70 ***
	(.89)	(1.00)	(.98)	(1.00)	(1.14)	(1.10)
Observations	2,020	2,020	2,020	1,675	1,675	1,675
R-squared	0.5676	0.5734	0.5733	0.4977	0.5083	0.5084

t statistics in parenthes

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4: Regression Results Patents

Interestingly, for patents applied and patents granted, the influence of HDI on innovation is more aligned with that of GDP. Specifically, for patents applied, the primary country's HDI coefficient is 1.68%, while GDP shows a similar influence of 1.00%. For patents granted, HDI remains slightly stronger at 1.47%, compared to 0.99% for GDP. These findings suggest that while development plays a critical role in early-stage innovation, the economic size of a country becomes just as important when converting research into patentable technologies.

4.2.2. Material Outcomes

	Panel A: Investments overall			Panel B: Investments incoming			Panel C: Investments outgoing		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	logInvestment	logInvestment	logInvestment	logInvestment	logInvestment	logInvestment	logInvestment	logInvestment	logInvestment
log GDPPrimary	0.53 ***	0.52 ***	0.51 ***	0.51 ***	0.50 ***	0.49 ***	0.53 ***	0.50 ***	0.49 ***
	(.02)	(.02)	(.02)	(.03)	(.03)	(.03)	(.02)	(.02)	(.02)
logGDPSecondary	0.32 ***	0.31 ***	0.31 ***	0.37 ***	0.36 ***	0.36 ***	0.31 ***	0.30 ***	0.30 ***
	(.02)	(.02)	(.02)	(.03)	(.03)	(.03)	(.03)	(.03)	(.03)
logDistance	-0.08 ***	-0.08 ***	-0.08 ***	-0.08 ***	-0.08 ***	-0.07 ***	-0.08 ***	-0.08 ***	-0.08 ***
	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
HDIPrimary		3.10 ***			1.68 **			3.95 ***	
		(.34)			(.57)			(.41)	
HDISecondary		.53			2.13 ***			-0.42	
		(.40)			(.60)			(.53)	
logGDPpcPrimary			0.35 ***			0.19 ***			0.43 ***
			(.03)			(.05)			(.04)
logGDPpcSecondary			0.10 **			0.26 ***			0.01
			(.04)			(.05)			(.05)
_cons	-20.75 ***	-23.32 ***	-24.26 ***	-21.40 ***	-24.28 ***	-25.29 ***	-20.18 ***	-22.46 ***	-23.56 ***
	(.76)	(.87)	(.86)	(1.12)	(1.31)	(1.30)	(1.04)	(1.18)	(1.16)
Observations	3,132	3,132	3,132	1,296	1,296	1,296	1,836	1,836	1,836
R-squared	0.3148	0.3356	0.3486	0.3352	0.3497	0.3599	0.2982	0.3308	0.3486

t statistics in parenthes
* p < 0.05, ** p < 0.01, *** p < 0.001

Table 5: Regression Results Investment

The role of HDI becomes even more pronounced when examining investments. For overall investments, a one-unit increase in the primary country's HDI leads to a 3.10% increase in investment activity, a notable effect compared to a 0.53% increase linked to a 1% rise in GDP. This highlights the dominant role of development in attracting investments. When examining incoming investments, the primary country's HDI coefficient of 1.68% is substantial, whereas GDP per capita plays a secondary role with a coefficient of 0.19%. Conversely, for outgoing investments, the primary country's HDI exerts the strongest influence with a coefficient of 3.95%, while the secondary country's HDI also has a meaningful effect at 2.13%.

Interestingly, GDP coefficients in the investment regressions are smaller compared to those in co-authorship and patents. While GDP consistently shows a positive effect on intellectual outcomes, its influence on material outcomessuch as investments is more modest. This suggests that while economic size is pivotal for generating research and innovation, it plays a less dominant role in attracting and channeling financial flows compared to HDI. GDP per capita, although significant, consistently shows smaller coefficients than GDP and HDI, particularly in material outcomes. For instance, GDP per capita has a coefficient of 0.21% for co-authorship, 0.14% for patents applied, and 0.19% for incoming investments. This indicates that while individual wealth is important, it does not exert as strong an influence as overall economic size or development. A stark contrast emerges when comparing intellectual and material outcomes in terms of the explanatory power of the models. Intellectual outcomes, such as co-authorship and patents applied, exhibit relatively higher R-squared values (0.54 and 0.57, respectively), suggesting that the gravity model more effectively explains knowledge creation and innovation patterns. These outcomes are more directly shaped by economic size and proximity. In contrast, material outcomes, including investments, show significant lower R-squared values (0.30 for investments overall and 0.33 for outgoing investments), suggesting that additional factors, beyond those captured in the gravity equation, have a stronger influence on financial flows.

4.2.3. Quantifying the Impact of HDI on Dependent Variables

This section complements the regression analysis by quantifying the standardized effects of a one-standard-deviation (1 SD) increase of HDI of the primary and secondary countries on each dependent variable, expressed in their respective standard deviations. Unlike other independent variables, HDI operates on a scale from 0 to 1. By standardizing its effects, this analysis makes

HDI's influence directly comparable across dependent variables, clarifying its relative impact. This enhances the interpretability of the findings and provides a critical contribution to the literature by demonstrating the concrete link between development levels and AI outcomes.

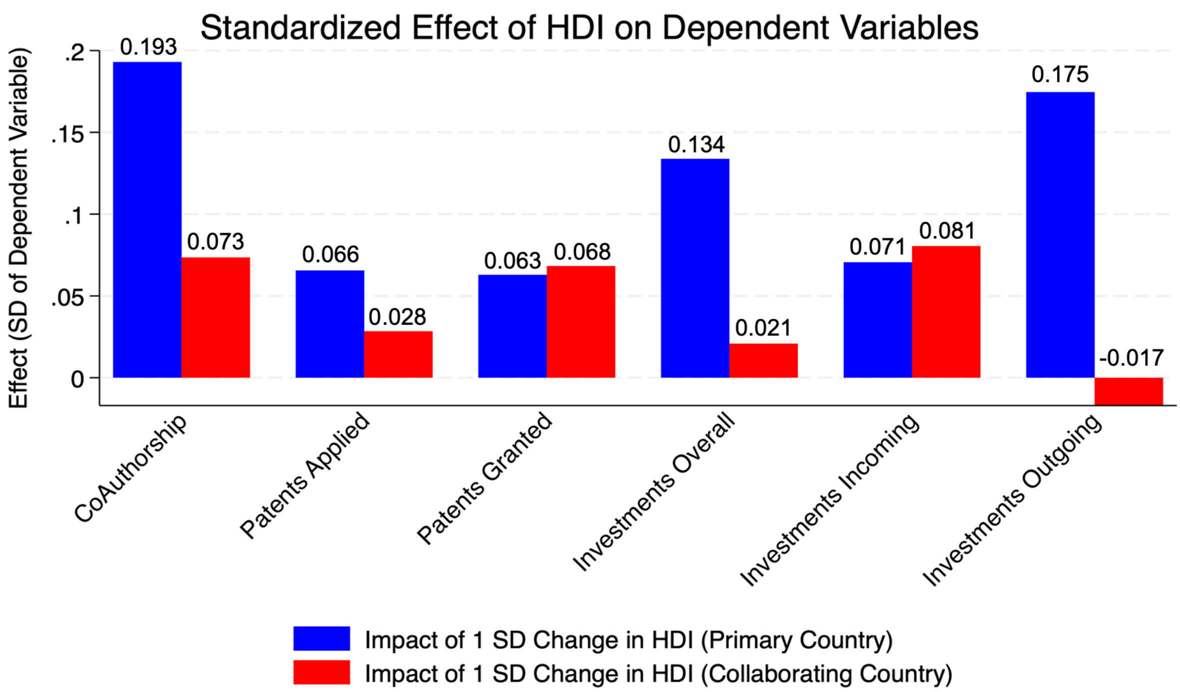

Figure 3: Standardized Effect of HDI on Dependent Variables

Among all outcomes, co-authorship displays the largest sensitivity to HDI changes. A 1 SD increase in the HDI of the primary country boosts co-authorship by 0.19 SDs, nearly tripling the effect of the secondary country's HDI at 0.07 SDs. This stark disparity emphasizes the dominant role of the primary country's development level in facilitating research collaboration, as supported by the regression coefficients (primary country: 2.16; secondary country: 1.10). Similarly, outgoing investments show the second most pronounced effect, with a 1 SD increase in the primary country's HDI leading to a 0.18 SD increase compared to a negligible -0.02 SD for the secondary country. This substantial asymmetry underlines that outward investment flows are almost entirely shaped by the development and economic capacity of the investing country, which is consistent with the regression findings for material outcomes. Following this, overall investments exhibit a significant but smaller effect (0.13 SD). This contrasts with both patent outcomes, which show a substantially smaller increases compared to co-authorship and investment outcomes. This illustrates that while development is still important, the economic size of the country may play a more dominant role in patent activity. Additionally, the impact of HDI on patents is more balanced between the primary and secondary countries, though still somewhat higher for the primary country. The regression analysis corroborates these findings with a similar pattern of stronger coefficients for the primary country.

4.3. Policy Implications

Building on the findings from the quantitative analysis, several key challenges and opportunities have been identified, highlighting critical areas for policy intervention. The following recommendations outline targeted policy implications to address these issues effectively.

Fostering International Research Partnerships

The standardized analysis reveals the critical role that the HDI of the primary country plays in driving co-authorship. This significant influence illustrates the importance of human capital, infrastructure, and institutional capacity in fostering international research collaborations. To capitalize on this, policymakers should prioritize initiatives that encourage international research partnerships, particularly led by high HDI countries. These partnerships could be facilitated through bilateral and multilateral funding streams, as well as the establishment of research collaboration platforms that lower barriers to entry for researchers in low and middle HDI countries. Strengthening the digital and physical infrastructure in these regions, such as the provision of advanced computational resources and access to training programs, would also be essential for enabling meaningful participation in global AI research efforts.

Scaling Research Growth from High-Growth Regions

The descriptive analysis highlights the impressive growth in research output in Sub-Saharan Africa and South Asia, which have both experienced significant increases in co-authorship.

These regions show that there is substantial potential for catching up in terms of knowledge creation. Policymakers should build on this momentum by identifying and supporting the factors driving this growth. Moreover, successful practices from these regions, should be studied and potentially extrapolated to other developing regions like Latin America and the Caribbean. By replicating these strategies in other regions with similar developmental challenges, countries can foster more inclusive knowledge creation and mitigate disparities in global AI collaboration.

Leveraging Geographic Proximity for Local AI Hubs

The study also found a consistent negative relationship between geographic distance and AI-related outcomes. This suggests that geographic proximity is a crucial factor in fostering collaboration, knowledge creation, and the flow of investments. Policymakers should prioritize the development of local AI ecosystems in regions where AI research and innovation have yet to reach critical mass. Strengthening local innovation hubs, increasing access to high-speed internet, and investing in computational resources can help create the infrastructure necessary to support domestic AI development. Such initiatives will reduce the dependence on external actors and empower regions to build their own capacity for research, innovation, and commercialization.

Overcoming Barriers to Patents and Innovation

While research output in developing regions is on the rise, the lack of corresponding patent activity in these regions emerged as significant barriers to turning knowledge into commercially viable technologies. Low and middle HDI countries, particularly those with high growth in research outputs, face substantial challenges in patenting their innovations. These barriers include limited access to intellectual property (IP) support, the high cost of patent filings, and a lack of capacity for managing and commercializing innovations. To address this gap, policymakers should prioritize improving access to IP resources, such as by creating subsidized legal and technical advisory services tailored to the needs of researchers in developing regions. The creation of regional patent hubs, which is in line with the previous recommendation, can also facilitate the filing process and reduce costs, empowering innovators in low-resource settings to protect and scale their technologies.

Channeling Investments for Inclusive AI Development

The regression analysis further confirms that investments, particularly outgoing investments, are strongly driven by the primary country's HDI. This suggests that investments from low and middle HDI countries into each other (South-South flows) are less significant and less impactful. Furthermore, the effect of the secondary country's HDI is comparatively smaller, highlighting its limited role in driving investment flows compared to the primary country. Given these findings, we recommend focusing on facilitating investments from high and very high HDI countries into developing regions, rather than prioritizing South-South investment flows. This could involve creating incentives for high-income nations to direct more funding toward low and middle HDI countries, ensuring that financial resources are channeled to where they can have the most substantial impact. Blended finance mechanisms, combining public and private investments, could be used to de-risk early-stage AI projects in developing regions.

5. Limitations and Further Research

While this study offers a comprehensive analysis of the determinants of AI development during the critical development phase, certain limitations must be acknowledged. The reliance on co-authorship metrics to measure knowledge creation provides valuable insights but is constrained to traditional institutions such as universities and research centers. Knowledge generation within private firms, an increasingly important driver of AI innovation, is not directly captured. Nevertheless, this limitation is mitigated by the fact that a majority of formal knowledge creation still occurs within these traditional institutions. Another limitation arises from the scope of the patent data, which is based on the location of filings rather than the nationality of inventors or the actual site of innovation. While this captures important market and jurisdictional trends, it may overlook innovation activity occurring outside major hubs. Similarly, the investment data focuses exclusively on private-market activity, excluding public-market investments. This could result in an incomplete picture of global AI financing trends. Finally, the use of the HDI as a proxy for development provides a comprehensive measure, but it may overlook other critical factors relevant to AI development, such as digital infrastructure and computational capacity.

Building on this the findings and presented limitations, future research could enhance understanding of global AI activity by addressing the aforementioned limitations and broadening the scope of analysis. Expanding data sources to include public investments and alternative knowledge creation metrics, such as open-source contributions and preprints, could offer a more nuanced view of AI ecosystems. Investigating interdisciplinary and cross-sector collaborations, such as academia-industry partnerships, could further illuminate underexplored dimensions of AI knowledge creation. While HDI offers a strong foundation for comparative analysis, future research could complement it with more specific technological and infrastructure-related metrics. A particularly impactful avenue for further research lies in assessing the downstream effects of investments and patents. Evaluating their role in fostering equitable AI adoption in low- and middle-income regions could inform policies aimed at reducing global disparities. Similarly, extending the temporal scope of analysis to capture post-2022 trends and focusing on emerging hubs in developing regions would help trace evolving patterns in AI development.

6. Conclusion

This study highlights the critical role of the development phase in shaping global AI ecosystems, addressing a key research gap by identifying it as a strategic and underexplored intervention point. While much of the literature focuses on computational capabilities, this work emphasizes the unique opportunities for mitigating systemic disparities during the early stages of AI development, where interventions are both more feasible and cost-effective. Focusing on the determinants of AI activity, the study employs the gravity equation framework to provide a comprehensive empirical analysis of intellectual and material outcomes. It quantifies the influence of economic size, development levels, and geographic proximity on global AI collaborations, innovation, and investment flows. Consistent with the gravity equation approach, the analysis finds that economic size positively influences AI outcomes, while distance consistently shows a negative relationship. The results also reveal that development levels play a decisive role across all outcomes, especially for co-authorship and investments. HDI's influence is particularly strong in driving investment flows, whereas GDP's effect on material outcomes is less pronounced. Despite growth in knowledge creation within low and middle HDI countries, it was found that these regions face barriers in translating research into patents and attracting investments.

To address these disparities, the study proposes targeted policy interventions that focus on strengthening knowledge creation, enhancing local AI ecosystems, overcoming barriers to commercialization, and strategically prioritizing investment flows from high-HDI countries. These recommendations aim to bridge the gap between research and commercialization, empowering underrepresented regions to meaningfully participate in global AI development. By quantifying the economic and developmental determinants of AI activity and emphasizing the strategic importance of the development phase, this study advances our understanding of global AI dynamics while offering actionable insights for fostering equitable AI ecosystems. These findings aim to guide policymakers and stakeholders in ensuring the benefits of AI development are more inclusively shared across regions and development levels.

Appendix

Appendix A: Data Collection

Research metrics in CAT are derived from ETO's *Merged Academic Corpus* (MAC), a dataset comprising over 260 million scholarly articles. An automated classifier identifies AI-related publications and articles are attributed to countries based on the affiliations of their authors, with multi-country collaborations credited to all participating nations, thereby serving as a proxy for knowledge creation. Patent data in CAT is compiled from sources such as *PATSTAT*, *The Lens*, and *1790 Analytics*. AI-related patents are identified through a combination of keywords and classification codes, with each patent attributed to its "priority country" - the jurisdiction of its initial filing. Drawing on data from patent offices worldwide, including national bodies like the *U.S. Patent and Trademark Office* and international entities such as the *European Patent Office*, the dataset captures the geographic distribution of innovation, differentiated into patent applications and patents granted. Investment data within CAT focuses exclusively on equity-based transactions involving privately held AI-related companies, sourced from Crunchbase. This dataset encompasses venture capital investments, private equity deals, and mergers and acquisitions, while excluding investments in publicly traded companies, such as major "Big Tech" firms, and non-equity financing such as grants or loans (ETO 2024b). To complement these AI-specific metrics, additional macroeconomic and geographic data were sourced independently. GDP and GDP per capita data were retrieved from the World Bank, expressed in constant 2015 USD to ensure consistency and comparability across countries (World Bank 2024). Human Development Index (HDI) values were obtained from the United Nations Human Development Report Office (UNDP 2024). For the gravity model analysis, geographic data on countries' latitudinal and longitudinal coordinates was sourced from the "GeoDist" database, maintained by the French Center for Prospective Studies and International Information (CEPII 2024). For the descriptive cluster analysis, information on regional classifications for each country was sourced from the United Nations Statistics Division (n.d.). Taken together, these contextual variables allow for the inclusion of economic conditions, development levels, and spatial proximity in analyzing global AI activity.

Appendix B: Data Description

Firstly, the research dataset is structured as panel data, encompassing cross-country co-authorship information for AI-related academic publications over a ten-year period from 2013 to 2022. This full-cycle coverage strengthens the reliability of the analysis by capturing long-term trends and shifts in international collaboration. The dataset spans 176 countries, resulting in 14,077 unique pairs of countries, defined by the nationality of the main author and the co-author. These unique country pairs are distributed consistently across the years, with 1,412 pairs recorded from 2013 to 2018, 1,403 pairs from 2019 to 2021, and 1,396 pairs in 2022. The inclusion of such a wide range of countries, coupled with a near-uniform distribution of country pairs across the period, emphasizes the dataset's robustness and its suitability for analyzing dynamic patterns in AI knowledge creation and collaboration. Secondly, the patent dataset tracks AI-related patents filed across multiple jurisdictions. Specifically, it counts patents initially filed in one patent office and subsequently filed or granted in another jurisdiction. This dataset includes metrics on patent activity from 66 countries, differentiated into patent applications and granted patents. Across the ten years, the dataset comprises 7,830 observations of country pairs, with 783 pairs recorded consistently each year. Its panel structure and even distribution allows for the examination of innovation dynamics and the evolution of intellectual property in AI over years. Additionally, the differentiation between patent applications and granted patents provides insights into both the initial stages of innovation and the subsequent realization of technological advancements. Thirdly, covering a ten-year period from 2013 to 2022, the investment dataset provides a longitudinal perspective on financial flows within the global AI ecosystem. It includes 11,310 observations across 106 countries, tracking 1,311 unique country pairs consistently over the decade. By distinguishing between incoming and outgoing investments, it captures the direction of financial flows between countries and provides a detailed view of cross-border investment dynamics in AI.[4]

Appendix C: Data Preprocessing

To ensure the integrity and reliability of the analysis, several preprocessing steps were undertaken across all datasets, aligning them with the study's methodological requirements. First, macroeconomic indicators, including GDP, GDP per capita, and HDI, along with geographical information such as latitude and longitude coordinates and regional classification, were assigned to each country for the corresponding year. This preprocessing step resulted in datasets comprising 176 countries for research, 66 for patents, and 106 for investments. Second, duplicates in the research dataset were addressed. Since the dataset captured co-authorship metrics, entries such as "USA-China" and "China-USA" were both recorded. While these reflect reciprocal collaborations, retaining both entries would disproportionately weight data toward major AI contributors like the United States and China. For this reason, duplicates were removed to ensure a balanced representation of collaborations across all countries. Third, the geographic distance between country pairs' capitals was calculated to support the gravity model analysis. This involved first assigning latitudinal and longitudinal coordinates to each country and then transforming these coordinates into distances using the haversine formula, which is widely utilized for determining great-circle distances on a spherical surface (Mahmoud and Akkari 2016). The formula is expressed as follows:

$$d = 2r \cdot \arcsin\left(\sqrt{\sin^2\left(\frac{\phi_2 - \phi_1}{2}\right) + \cos(\phi_1) \cdot \cos(\phi_2) \cdot \sin^2\left(\frac{\lambda_2 - \lambda_1}{2}\right)}\right)$$

[4] Detailed information regarding the data preprocessing is provided in Appendix B.

Here, d represents the distance between two points, r is the Earth's radius (approximately 6,371 km), ϕ_1 and ϕ_2 are the latitudes, and λ_1 and λ_2 are the longitudes of the two countries' capitals. Finally, due to the skewed distribution of research, patent, and investment data toward major AI contributors (e.g., the United States, China, and Western European Countries including France, Germany, United Kingdom), logarithmic transformations were applied to normalize the data. This adjustment ensured that the analysis accounted for relative differences across countries while minimizing distortions caused by extreme values. These preprocessing steps resulted in robust and standardized datasets suitable for both regression analysis and descriptive trend analysis.

List of Figures:

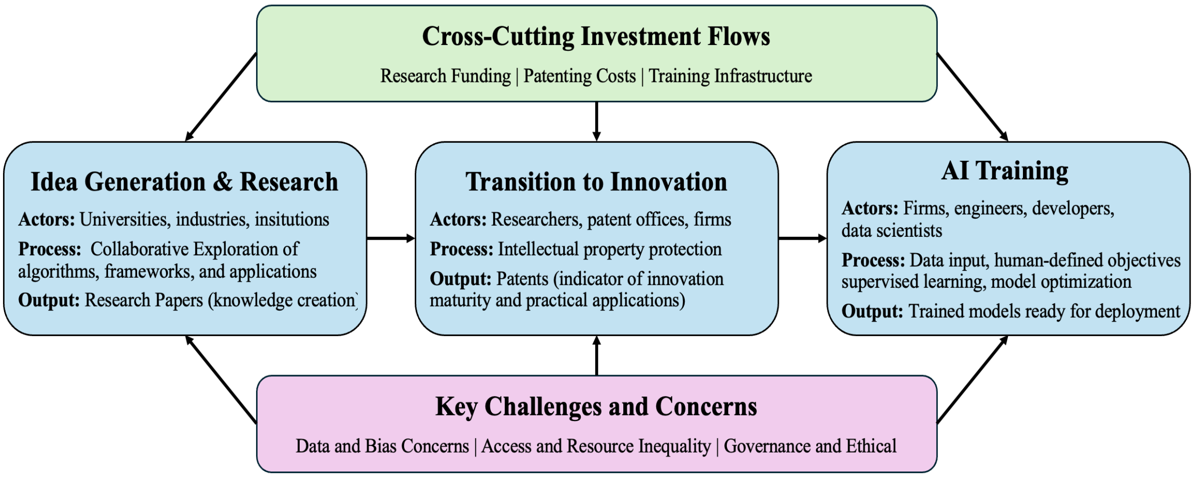

Figure 1: The AI Development Phase: Processes, Outputs, and Challenges

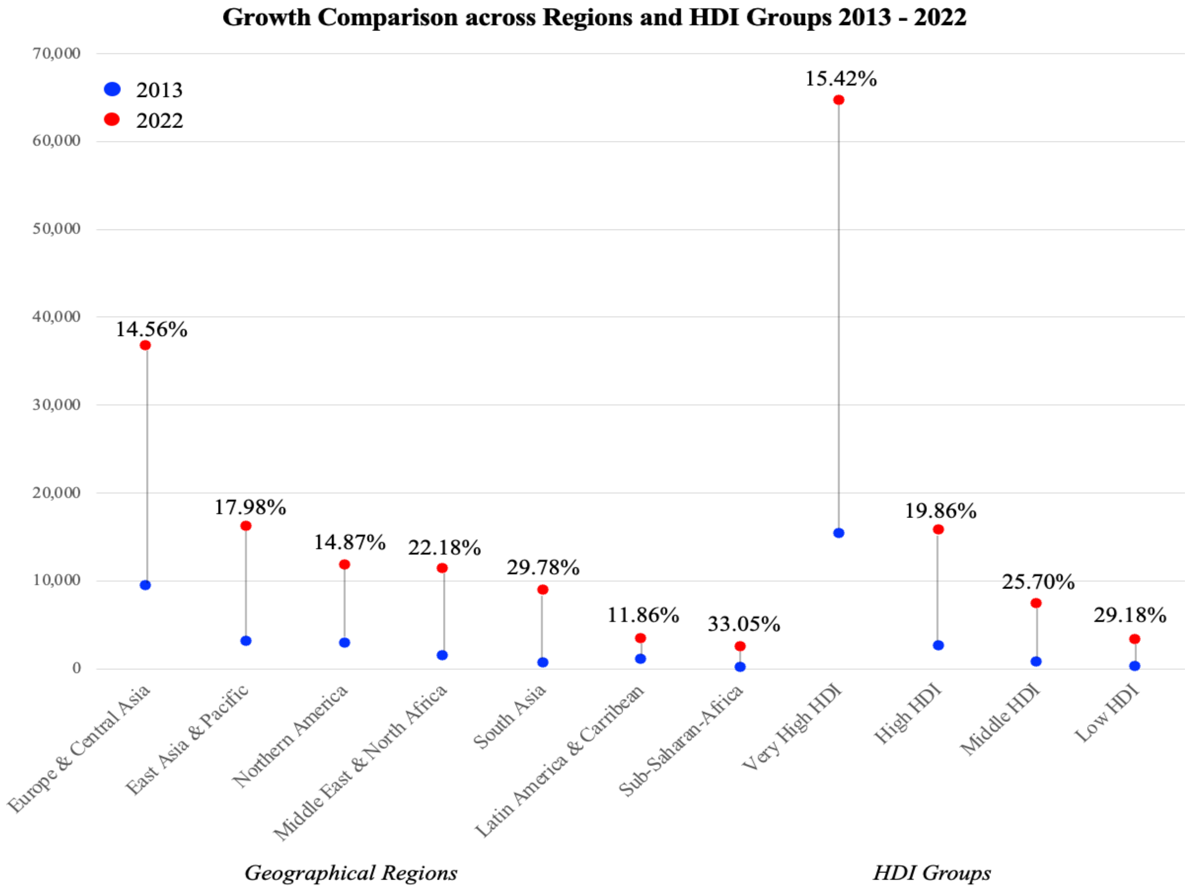

Figure 2: Growth Comparison across Regions and HDI Groups 2013-2022

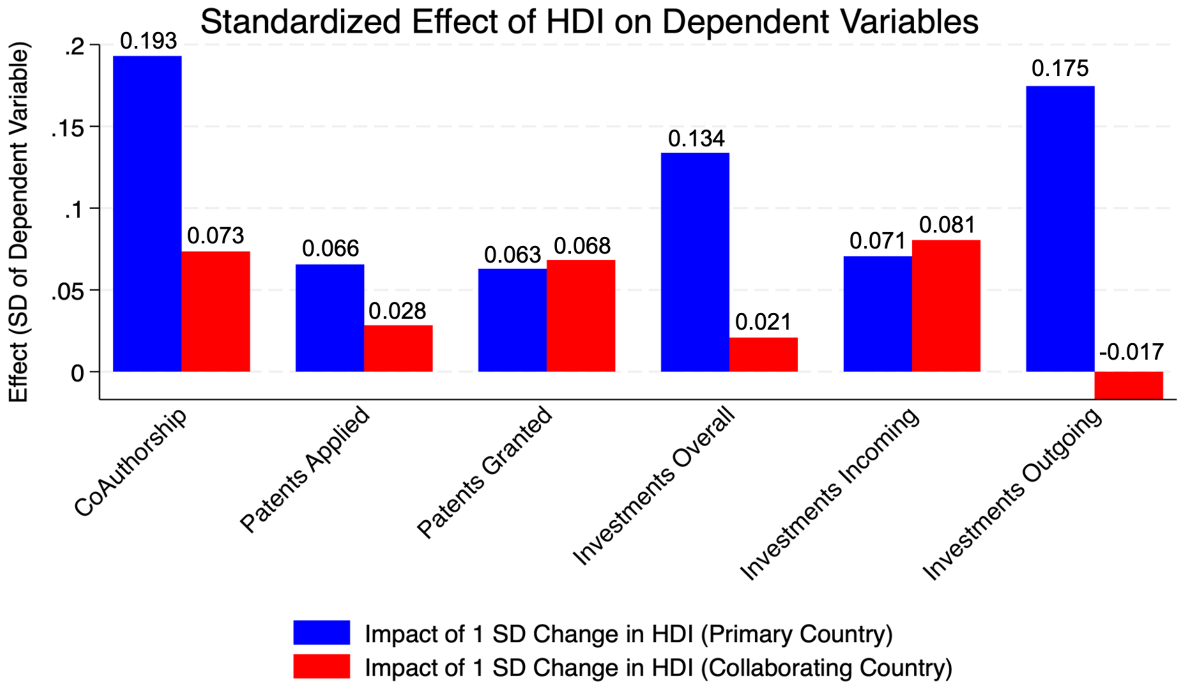

Figure 3: Standardized Effect of HDI on Dependent Variables

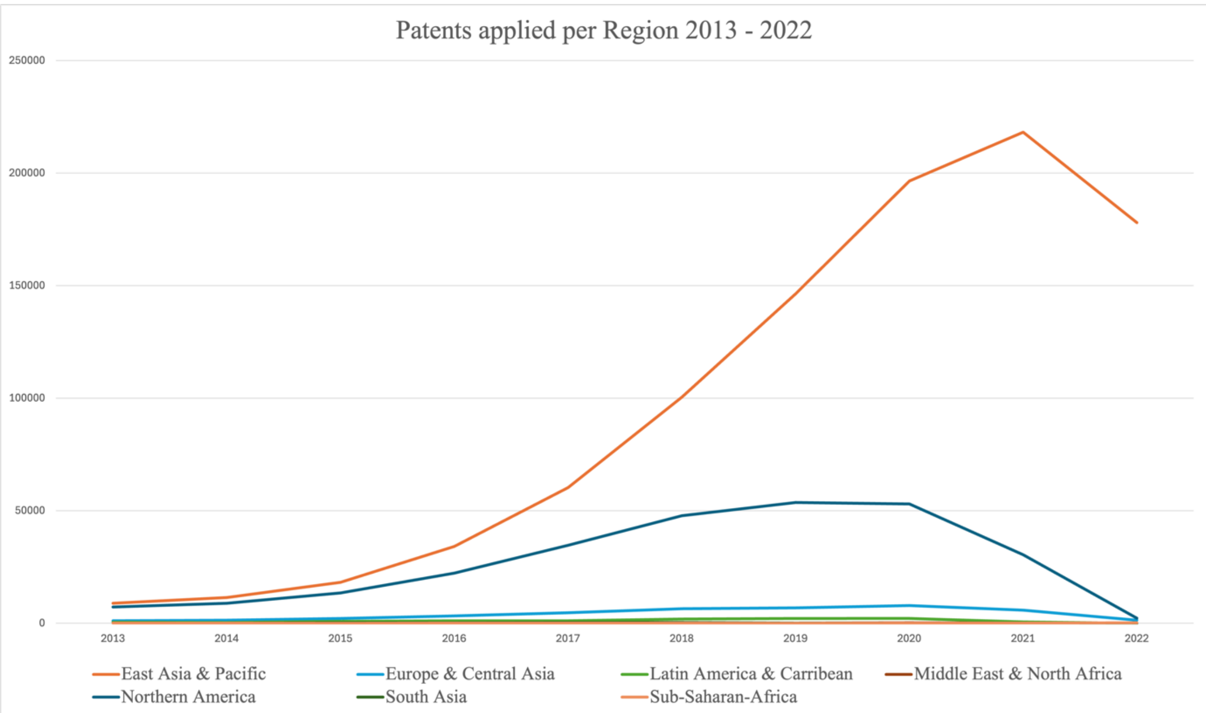

Figure 4: Patents applied per Region 2013 – 2022

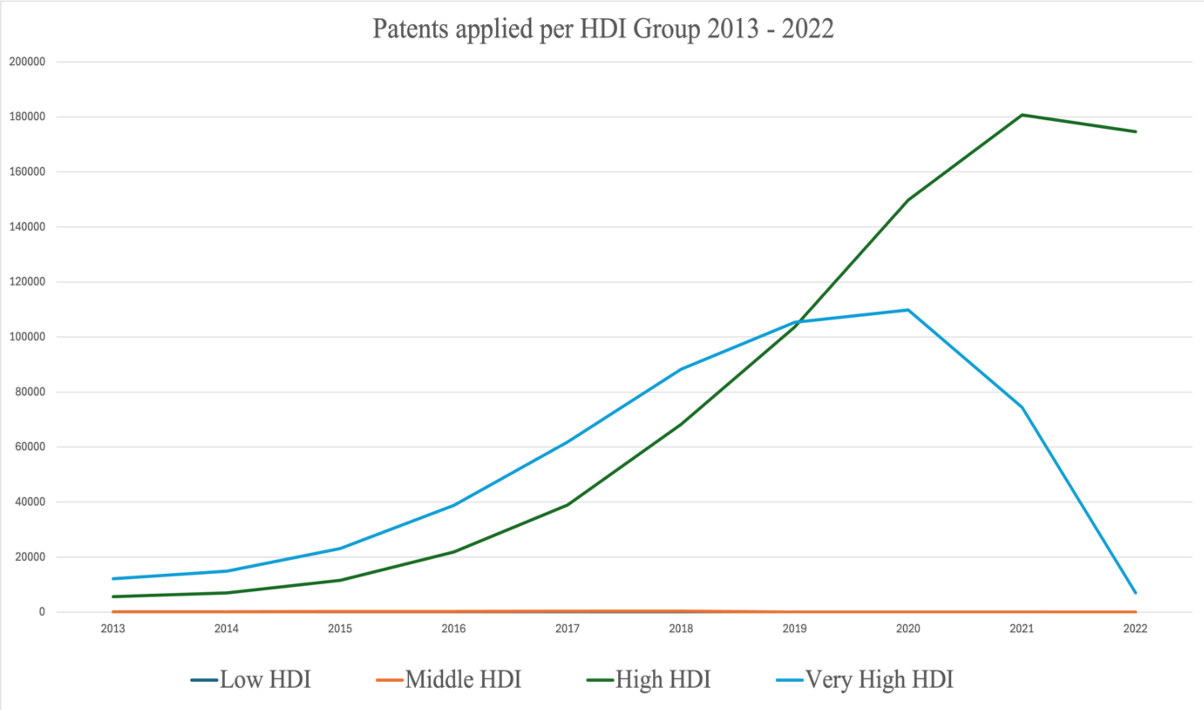

Figure 5: Patents applied per HDI Group 2013 – 2022

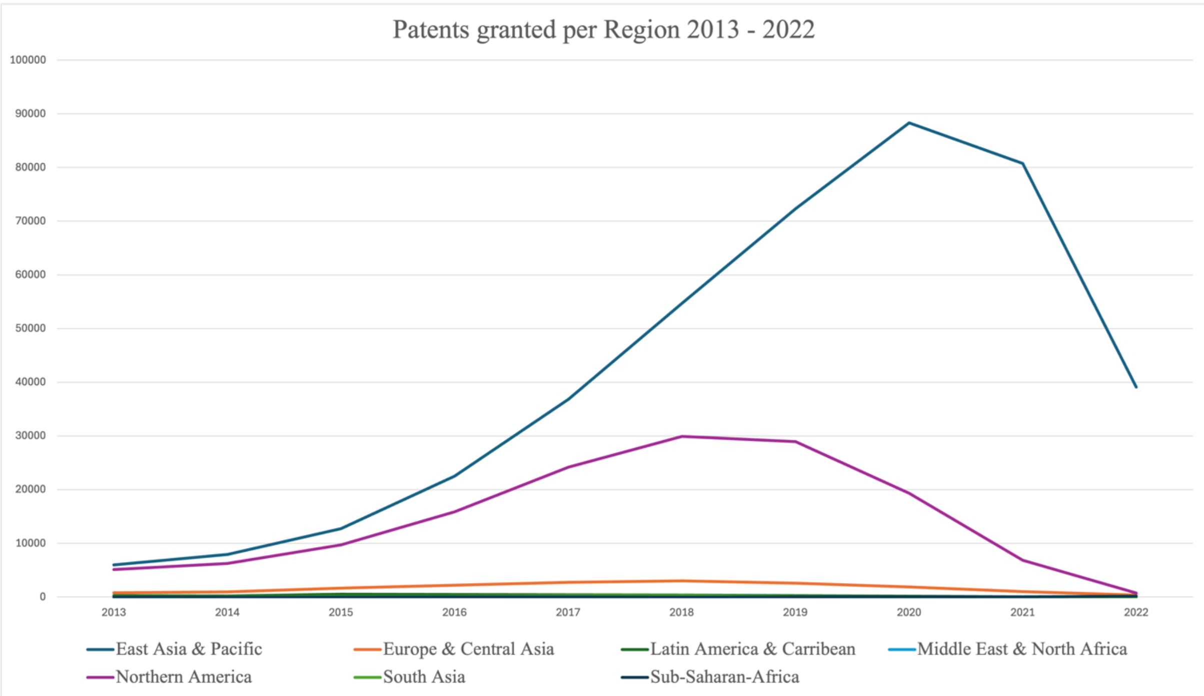

Figure 6: Patents granted per Region 2013 – 2022

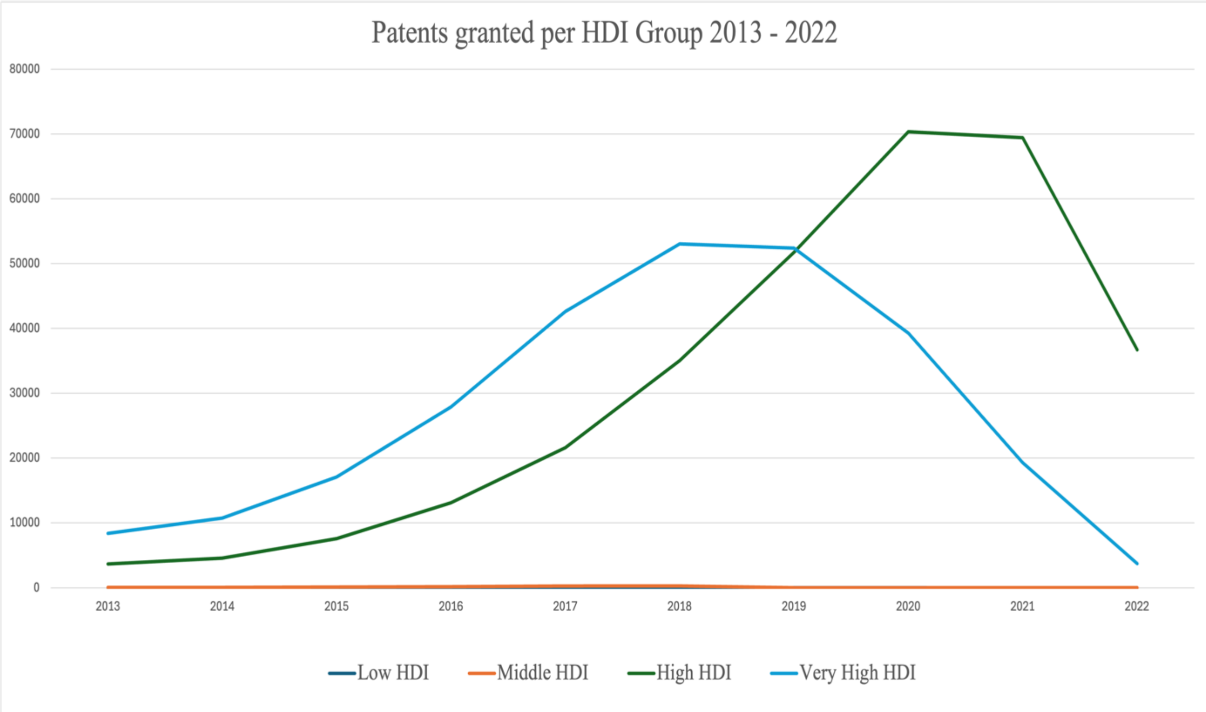

Figure 7: Patents granted per HDI Group 2013 – 2022

List of Tables:

Form of Bias	Definition
Dataset Bias	Underrepresentation of certain groups or regions
Algorithmic Bias	Prioritization inherently favors certain outcomes over others (influenced by the socio-political contexts in which these systems are developed).
Emergent Bias	Unexpected system-user interactions

Table 1: Sources of Biases (Ferrara 2024)

Phase	Key Activities	Manifestation of the Global Divide
Development	Knowledge creation (research collaborations)	Dominance of high-income regions in research infrastructure and networks
	Innovation (patent filings and advancements)	Limited capacity for innovation due to funding and IP access barriers
	Infrastructure (such as computational resources)	Unequal access to GPUs, data centers, and training infrastructure
	Early-stage financing	Minimal investments in AI R&D in low-income regions
Deployment	AI tool adoption	Limited adoption due to infrastructure deficits and affordability issues
	Sectoral integration (e.g., healthcare, education)	Developed regions benefit from sector-wide AI solutions
	Regulatory and ethical implementation	Absence of robust governance frameworks in low-income regions
	Capacity building (skills and digital literacy)	Shortage of AI-trained professionals limits local deployment
	Data localization and relevance	Poorly contextualized AI systems reduce effectiveness

Table 2: Key Activities and Manifestations of the Global Divide

	CoAuthorship		
	(1)	(2)	(3)
	logArticles	logArticles	logArticles
log GDPPrimary	0.65 ***	0.58 ***	0.58 ***
	(.01)	(.01)	(.01)
logGDPSecondary	0.28 ***	0.23 ***	0.23 ***
	(.01)	(.01)	(.01)
logDistance	-0.22 ***	-0.13 ***	-0.15 ***
	(.01)	(.01)	(.01)
HDIPrimary		2.16 ***	
		(.09)	
HDISecondary		1.10 ***	
		(.12)	
logGDPpcPrimary			0.21 ***
			(.01)
logGDPpcSecondary			0.11 ***
			(.01)
_cons	-20.46 ***	-20.47 ***	-20.80 ***
	(.29)	(.28)	(.28)
Observations	8,667	8,667	8,667
R-squared	0.5389	0.5769	0.5738

t statistics in parenthes

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3: Regression Results Co-authorship

	Panel A: Patents applied			Panel B: Patents granted		
	(1)	(2)	(3)	(4)	(5)	(6)
	logPatents	logPatents	logPatents	logPatents	logPatents	logPatents
log GDPPrimary	1.00 ***	0.99 ***	0.98 ***	0.75 ***	0.75 ***	0.74 ***
	(.02)	(.02)	(.20)	(.02)	(.02)	(.022)
logGDPSecondary	0.30 ***	0.29 ***	0.29 ***	0.35 ***	0.36 ***	0.34 ***
	(.03)	(.03)	(.03)	(.03)	(.03)	(.03)
logDistance	-0.10 ***	-0.10 ***	-0.10 ***	-0.10 ***	-0.10 ***	-0.09 ***
	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
HDIPrimary		1.68 ***			1.47 ***	
		(.35)			(.35)	
HDISecondary		0.82 *			1.90 ***	
		(.40)			(.44)	
logGDPpcPrimary			0.14 ***			0.12 ***
			(.03)			(.03)
logGDPpcSecondary			0.10 *.			0.20 ***
			(.04)			(.04)
_cons	-32.86 ***	-34.68 ***	-34.50 ***	-28.17 ***	-31.24 ***	-30.70 ***
	(.89)	(1.00)	(.98)	(1.00)	(1.14)	(1.10)
Observations	2,020	2,020	2,020	1,675	1,675	1,675
R-squared	0.5676	0.5734	0.5733	0.4977	0.5083	0.5084

t statistics in parenthes
* p < 0.05, ** p < 0.01, *** p < 0.001

Table 4: Regression Results Patents

	Panel A: Investments overall			Panel B: Investments incoming			Panel C: Investments outgoing		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	logInvestment	logInvestment	logInvestment	logInvestment	logInvestment	logInvestment	logInvestment	logInvestment	logInvestment
log GDPPrimary	0.53 ***	0.52 ***	0.51 ***	0.51 ***	0.50 ***	0.49 ***	0.53 ***	0.50 ***	0.49 ***
	(.02)	(.02)	(.02)	(.03)	(.03)	(.03)	(.02)	(.02)	(.02)
logGDPSecondary	0.32 ***	0.31 ***	0.31 ***	0.37 ***	0.36 ***	0.36 ***	0.31 ***	0.30 ***	0.30 ***
	(.02)	(.02)	(.02)	(.03)	(.03)	(.03)	(.03)	(.03)	(.03)
logDistance	-0.08 ***	-0.08 ***	-0.08 ***	-0.08 ***	-0.08 ***	-0.07 ***	-0.08 ***	-0.08 ***	-0.08 ***
	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)	(.01)
HDIPrimary		3.10 ***			1.68 **			3.95 ***	
		(.34)			(.57)			(.41)	
HDISecondary		.53			2.13 ***			-0.42	
		(.40)			(.60)			(.53)	
logGDPpcPrimary			0.35 ***			0.19 ***			0.43 ***
			(.03)			(.05)			(.04)
logGDPpcSecondary			0.10 **			0.26 ***			0.01
			(.04)			(.05)			(.05)
_cons	-20.75 ***	-23.32 ***	-24.26 ***	-21.40 ***	-24.28 ***	-25.29 ***	-20.18 ***	-22.46 ***	-23.56 ***
	(.76)	(.87)	(.86)	(1.12)	(1.31)	(1.30)	(1.04)	(1.18)	(1.16)
Observations	3,132	3,132	3,132	1,296	1,296	1,296	1,836	1,836	1,836
R-squared	0.3148	0.3356	0.3486	0.3352	0.3497	0.3599	0.2982	0.3308	0.3486

t statistics in parenthes
* p < 0.05, ** p < 0.01, *** p < 0.001

Table 5: Regression Results Investments

From	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	Total
Incoming Investments (in million USD) into Low HDI Countries											
Low HDI Countries	0	0	0	0	0	0	0	1	3	3	7
Middle HDI Countries	0	0	0	0	0	0	0	1	0	2	3
High HDI Countries	0	0	0	0	0	0	0	0	0	0	0
Very High HDI Countries	0	0	0	0	0	0	8	0	0	1	9
Total	0	0	0	0	0	0	8	2	3	6	19
Incoming Investments (in million USD) into Middle HDI Countries											
Middle HDI Countries	4	30	419	57	63	312	295	131	793	1,255	3,359
High HDI Countries	0	0	0	0	0	0	0	0	0	5	5
Very High HDI Countries	49	1,991	577	31	31	1,243	109	20,378	363	722	25,494
Total	53	2,021	996	88	94	1,555	404	20,509	1,156	1,982	28,858
Incoming Investments (in million USD) into High HDI Countries											
Low HDI Countries	0	0	0	0	0	0	0	0	0	5	5
Middle HDI Countries	0	0	0	0	1	0	0	1	20	9	31
High HDI Countries	79	453	1,146	5,333	2,879	7,701	3,877	5,211	11,144	6,835	44,658
Very High HDI Countries	8	167	140	117	813	631	757	2,338	2,231	1,476	8,678
Total	87	620	1,286	5,450	3,693	8,332	4,634	7,550	13,395	8,325	53,372
Incoming Investments (in million USD) into Very High HDI Countries											
Low HDI Countries	0	0	0	0	0	0	0	0	0	0	0
Middle HDI Countries	38	67	295	80	46	228	274	195	403	528	2,154
High HDI Countries	30	135	205	623	562	1,132	753	1,522	2,794	1,557	9,313
Very High HDI Countries	5,269	10,417	13,176	13,209	15,123	31,065	36,782	41,709	95,110	66,685	328,545
Total	5,337	10,619	13,676	13,912	15,731	32,425	37,809	43,426	98,307	68,770	340,012

Table 6: Incoming investments clustered by HDI Groups 2013 – 2022

List of References

Acemoglu, Daron, and Pascual Restrepo. 2019. “Automation and New Tasks: How Technology Displaces and Reinstates Labor.” *Journal of Economic Perspectives* 33, no. 2: 3–30. https://doi.org/10.1257/jep.33.2.3.

Adam, Hammaad, Aparna Balagopalan, Emily Alsentzer, Fotini Christia, and Marzyeh Ghassemi. 2022.“Mitigating the Impact of Biased Artificial Intelligence in Emergency Decision-Making.” *Communications Medicine* 2, no. 149. https://doi.org/10.1038/s43856-022-00214-4.

Anderson, James E. 2011. "The Gravity Model." *Annual Review of Economics* 3: 133–160. https://doi.org/10.1146/annurev-economics-111809-125114.

Anderson, James, E., and Eric van Wincoop. 2003. "Gravity with Gravitas: A Solution to the Border Puzzle ." *American Economic Review*, 93 (1): 170–192**.** https://doi.org/10.1257/000282803321455214.

Bergstrand, Jeffrey, and Peter Egger. 2013. “Gravity Equations and Economic Frictions in the World Economy.” In *Palgrave Handbook of International Trade*, 532–570. London: Palgrave Macmillan. https://doi.org/10.1007/978-0-230-30531-1_17.

Božić, Velibor. 2023. “Risks of digital divide in using artificial intelligence (AI).” https://doi.org/10.13140/RG.2.2.18156.13443.

Brynjolfsson, Erik, and Andrew McAfee. 2014. *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies.* New York: W.W. Norton & Company.

Buolamwini, Joy, and Timnit Gebru. 2018. “Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification.” *Proceedings of Machine Learning Research* 81: 1–15. Conference on Fairness, Accountability, and Transparency. https://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf.

Cazzaniga, Mauro, Florence Jaumotte, Longji Li, Giovanni Melina, Augustus J. Panton, Carlo Pizzinelli, Emma Rockall, and Marina M. Tavares. 2024. *Gen-AI: Artificial Intelligence and the Future of Work.* International Monetary Fund. https://www.imf.org/en/Publications/Staff-Discussion-Notes/Issues/2024/01/14/Gen-AI-Artificial-Intelligence-and-the-Future-of-Work-542379.

CEPII. 2024. GeoDist Database. Accessed October 15, 2024. https://www.cepii.fr/CEPII/en/bdd_modele/bdd_modele_item.asp?id=6.

Chiodo, Simone. 2022. “Human autonomy, technological automation (and reverse).” *AI & Soc* 37: 39–48. https://doi.org/10.1007/s00146-021-01149-5.

Crescenzi, Riccardo, Simona Iammarino, Carolin Ioramashvili, Andrés Rodríguez-Pose, and Michael Storper.2019. “The Geography of Innovation: Local Hotspots and Global

Innovation Networks." *Economic Research Working Paper Series*, no. 57. Accessed December 07, 2024. https://doi.org/10.2139/ssrn.4424262.

Colombo, Emiliano, Fabio Mercorio, and Mario Mezzanzanica. 2019. "AI meets labor market: Exploring the link between automation and skills." *Information Economics and Policy* 47: 27–37. https://doi.org/10.1016/j.infoecopol.2019.05.003.

Eaton, Jonathan, and Samuel Kortum. 2002. "Technology, Geography, and Trade." *Econometrica* 70 (5): 1741–79. https://doi.org/10.2307/3082009.

Efe, Ahmet. 2022. "The Impact of Artificial Intelligence on Social Problems and Solutions: An Analysis on the Context of Digital Divide and Exploitation." *Yeni Medya Dergisi.* Accessed November 28, 2024. https://www.researchgate.net/publication/366530822_The_Impact_of_Artificial_Intelligence_on_Social_Problems_and_Solutions_An_Analysis_on_The_Context_of_Digital_Divide_and_Exploitation.

Emerging Technology Observatory (ETO). 2024a. *Country Activity Tracker: Artificial Intelligence* .https://cat.eto.tech/?expanded=Summary-metrics.

———. 2024b. *Documentation: Country Activity Tracker: Artificial Intelligence.* https://eto.tech/tool-docs/cat/.

European Commission. 2018. "A Definition of AI: Main Capabilities and Scientific Disciplines." *High-Level Expert Group on Artificial Intelligence.* Accessed October 5, 2024. https://ec.europa.eu/futurium/en/system/files/ged/ai_hleg_definition_of_ai_18_december_1.pdf.

European Union Institute for Security Studies, and Simona Soare. 2020."Digital Divide? Transatlantic Defence Cooperation on Artificial Intelligence." *Publications Office of the European Union.* Accessed December 1, 2024. https://data.europa.eu/doi/10.2815/268841.

Ferrara, Emilio. 2024. "Fairness and Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, and Mitigation Strategies." *Sci* 6, no. 1:3. https://doi.org/10.3390/sci6010003.

Floridi, Luciano, Josh Cowls, Matthew King, Mariarosaria Taddeo, Ramon Chatila, Pierre Chazerand, Virginia Dignum, et al. 2018. "AI4People - An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations." *Minds & Machines* 28: 689–707. Accessed October 13, 2024. https://doi.org/10.1007/s11023-018-9482-5.

Grobelnik, Marko, Karine Perset, and Stuart Russell. 2024. "What Is AI? Can You Make a Clear Distinction Between AI and Non-AI Systems?" *OECD.AI Policy Observatory*. Accessed October 03, 2024. https://oecd.ai/en/wonk/definition.

Hall, Bronwyn H., and Dietmar Harhoff. 2012. "Recent Research on the Economics of Patents." *NBER Working Paper* no. 17773. Cambridge, MA: National Bureau of Economic Research. Accessed October 20, 2024.

https://www.nber.org/papers/w17773.

Head, Keith, and Thierry Mayer. 2014. “Gravity Equations: Workhorse, Toolkit, and Cookbook.” *Handbook of International Economics* 4: 131–195. https://doi.org/10.1016/B978-0-444-54314-1.00003-3.

International Telecommunication Union. 2024. “ITU-T Focus Group Technical Report: Technical Report on Transformative AI Models for Natural Disaster Management.” *ITU Publications.* Accessed October 13, 2024. https://www.itu.int/en/ITU-T/focusgroups/ai4ndm/Documents/Deliverables/FG-AI4NDM%20WG-MODELING%20-%20AI%20for%20Modeling.pdf.

Keller, Wolfgang. 2002. "Geographic Localization of International Technology Diffusion." *American Economic Review*92 (1): 120–42. https://doi.org/10.1257/000282802760015630.

Krakowski, Sebastian, Julia Luger, and Sebastian Raisch. 2022. “Artificial Intelligence and the Changing Sources of Competitive Advantage.” *Strategic Management Journal* 44, no. 6: 1425–52. https://doi.org/10.1002/smj.3387.

Krishna, Ketana. 2024. “Artificial Intelligence: Bridging or Exacerbating the Digital Divide?” *SSRN*. https://doi.org/10.2139/ssrn.4866372.

Leamer, Edward E., and James Levinsohn. 1995. “International Trade Theory: The Evidence.” In *Handbook of International Economics* 3: 1339–94. Amsterdam: Elsevier. https://doi.org/10.1016/S1573-4404(05)80006-1.

Leavy, Susan, Barry O'Sullivan, and Eugenia Siapera. 2020. “Data, Power and Bias in Artificial Intelligence.” In *AI for Social Good: Harvard CRCS Workshop, 20–21 July 2020.* Accessed November 21, 2024. https://arxiv.org/abs/2008.07341.

Lehdonvirta, Vili, Bili Wú, and Zoe Hawkins. 2024. “Compute North vs. Compute South: The Uneven Possibilities of Compute-Based AI Governance Around the Globe.” *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* 7, no. 1: 828–38. https://doi.org/10.1609/aies.v7i1.31683.

Lengyel, Balázs, Tamás Sebestyén, and Loet Leydesdorff. 2013. “Challenges for Regional Innovation Policies in CEE Countries: Spatial Concentration and Foreign Control of US Patenting.” *Science and Public Policy* (forthcoming).

Mahmoud, Hagar, and Nadine Akkari. 2016. "Shortest Path Calculation: A Comparative Study for Location-Based Recommender System." *World Symposium on Computer Applications & Research*. King Abdulaziz University. https://doi.org/10.1109/WSCAR.2016.16.

McKinsey Global Institute. 2018. *Notes from the AI Frontier: Modeling the Impact of AI on the World Economy.* Accessed October 03, 2024. https://www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-ai-frontier-modeling-the-impact-of-ai-on-the-world-economy.

———. 2020. *The State of AI in 2020.* Accessed October 13, 2024. https://www.mckinsey.com/capabilities/quantumblack/our-insights/global-survey-the-state-of-ai-in-2020.

Mitchell, Margaret, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timnit Gebru. 2019. "Model Cards for Model Reporting." In *FAT '19: Conference on Fairness, Accountability, and Transparency*, January 29–31, 2019, Atlanta, GA, USA. https://doi.org/10.1145/3287560.3287596.

Mökander, Jakob, Margi Sheth, David S. Watson, and Luciano Floridi. 2023. "The Switch, the Ladder, and the Matrix: Models for Classifying AI Systems." *Minds & Machines* 33: 221–48. Accessed October 13, 2024. https://doi.org/10.1007/s11023-022-09620-y.

National Security Commission on Artificial Intelligence. 2021. *Final Report.* Accessed October 5, 2024. https://assets.foleon.com/eu-central-1/de-uploads/7e3kk3/48187/nscai_full_report_digital.04d6b124173c.pdf.

Ntoutsi, Eirini, Pavlos Fafalios, Ujwal Gadiraju, Vasileios Iosifidis, Wolfgang Nejdl, Maria-Esther Vidal, Salvatore Ruggieri, et al. 2020. "Bias in Data-Driven Artificial Intelligence Systems - An Introductory Survey." *WIREs Data Mining and Knowledge Discovery.* https://doi.org/10.1002/widm.1356.

Obermeyer, Ziad, Brian Powers, Christine Vogeli, and Sendhil Mullainathan. 2019. "Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations." *Science* 366, no. 6464: 447–53. https://doi.org/10.1126/science.aax2342.

O'Connor, Sinead, and Helen Liu. 2023. "Gender Bias Perpetuation and Mitigation in AI Technologies: Challenges and Opportunities." *AI & Society* 39, no. 4: 1–13. https://doi.org/10.1007/s00146-023-01675-4.

OECD. 2017. *The Next Production Revolution: Implications for Governments and Business.* Paris: OECD Publishing. Accessed October 15, 2024. https://doi.org/10.1787/9789264271036-en.

———. 2024a. Explanatory Memorandum on the Updated OECD Definition of an AI System. OECD *Artificial Intelligence Papers*, no. 8. OECD Publishing. Accessed October 5, 2024. https://doi.org/10.1787/623da898-en.

———. 2024b. *What Is AI? Can You Make a Clear Distinction Between AI and Non-AI Systems? OECD AI Policy Observatory.* https://oecd.ai/en/wonk/definition.

Paunov, Caroline, et al. 2019. "On the Concentration of Innovation in Top Cities in the Digital Age." *OECD Science, Technology and Industry Policy Papers*, no. 85. Paris: OECD Publishing. https://doi.org/10.1787/f184732a-en.

Popoola, Bankole. 2024. *AI and Automation.* https://doi.org/10.13140/RG.2.2.15359.21929.

Redding, Stephen J., and David E. Weinstein. 2019. "Aggregation and the Gravity Equation." *AEA Papers and Proceedings* 109: 450–55. https://doi.org/10.1257/pandp.20191006.

Russell, Stuart J., and Peter Norvig. 2020. *Artificial Intelligence: A Modern Approach.* 4th ed. Pearson.

Sarmah, Simanta Shekhar. 2019. "Artificial Intelligence in Automation." *Research Review Journals* 4, no. 6. https://www.researchgate.net/publication/336085049_Artificial_Intelligence_in_Automation.

Shestakova, Veronika. 2021. "Best Practices to Mitigate Bias and Discrimination in Artificial Intelligence." *Performance Improvement 60,* no. 1: 25–32. https://doi.org/10.1002/pfi.21987.

Spencer, David. 2024. "AI, Automation and the Lightening of Work." *AI & Society.* https://doi.org/10.1007/s00146-024-01959-3.

United Nations. 2024. *Mind the AI Divide: Shaping a Global Perspective on the Future of Work*. https://www.un.org/techenvoy/sites/www.un.org.techenvoy/files/MindtheAIDivide.pdf

United Nations Development Programme (UNDP). 2024. Human Development Index Data Center. Accessed October 03, 2024. https://hdr.undp.org/data-center/human-development-index#/indicies/HDI.

United Nations Statistics Division. n.d. *Standard Country or Area Codes for Statistical Use (M49): Methodology.* Accessed November 28, 2024. https://unstats.un.org/unsd/methodology/m49/.

U.S. Department of State. n.d. "Artificial Intelligence (AI)." Accessed October 5, 2024. https://www.state.gov/artificial-intelligence/.

Vinuesa, Ricardo, Hossein Azizpour, Iolanda Leite, Margot Balaam, Virginia Dignum, Sebastian Domisch, Anna Felländer, Stefan D. Langhans, Max Tegmark, and Francesco Fuso Nerini. 2020. "The Role of Artificial Intelligence in Achieving the Sustainable Development Goals." *Nature Communications* 11, no. 233. https://doi.org/10.1038/s41467-019-14108-y.

Wachter, Sandra, Brent Mittelstadt, and Luciano Floridi. 2016. "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation." *SSRN Electronic Journal.* https://doi.org/10.1093/idpl/ipx005.

World Bank. 2024. World Bank Open Data. Accessed October 15, 2024. https://data.worldbank.org.

World Intellectual Property Organization (WIPO). 2022. *World Intellectual Property Report 2022: The Direction of Innovation.* Geneva: World Intellectual Property Organization. Accessed October 20, 2024. https://doi.org/10.34667/tind.45356.