

When AI Meets Science

Research Diversity, Interdisciplinarity, Visibility, and Retractions across Disciplines in a Global Surge

Andrés F. Castro Torres

26.10.2025

Computational Social Sciences and Humanities Laboratory

The **Science and Technology Studies** Research Line

This research line examines how knowledge is produced, focusing on **global inequalities**.

We analyze scientific publications, researcher mobility, collaboration, and methods to assess their societal and technological impact.



MERCE CROSAS
COMPUTATIONAL SOCIAL SCIENCES AND HUMANITIES LABORATORY DIRECTOR
Computational Social Sciences and Humanities Laboratory



ANDRES FELIPE CASTRO TORRES
ESTABLISHED RESEARCHER
Computational Social Sciences and Humanities Laboratory - Computational Social Sciences

PALOMA NAVARRO
MASTER'S STUDENT - INTERN
Computational Social Sciences and Humanities Laboratory

The team behind this work

The authors



ANDRES FELIPE CASTRO TORRES
ESTABLISHED RESEARCHER

Computational Social Sciences and Humanities Laboratory - Computational Social Sciences



JOAN GINER MIGUELEZ
SENIOR RESEARCH ENGINEER

Computational Social Sciences and Humanities Laboratory - Data infrastructure Models and Methods



MERCE CROSAS
COMPUTATIONAL SOCIAL SCIENCES AND HUMANITIES LABORATORY DIRECTOR

Computational Social Sciences and Humanities Laboratory

Methods and Engineering for the Social Sciences and Humanities (MESSH)



CLARA GUASCH FORTUNY
SENIOR RESEARCH ENGINEER

Computational Social Sciences and Humanities Laboratory - Data infrastructure Models and Methods



PAOLO MARANGIO
RESEARCH ENGINEER

Computational Social Sciences and Humanities Laboratory - Data infrastructure Models and Methods



ALEXANDRA MALAGA MELLADO
SENIOR RESEARCH ENGINEER

Computational Social Sciences and Humanities Laboratory - Data infrastructure Models and Methods



ADRIAN CARRASCOSA LOPEZ
RESEARCH ENGINEER

Computational Social Sciences and Humanities Laboratory - Data infrastructure Models and Methods

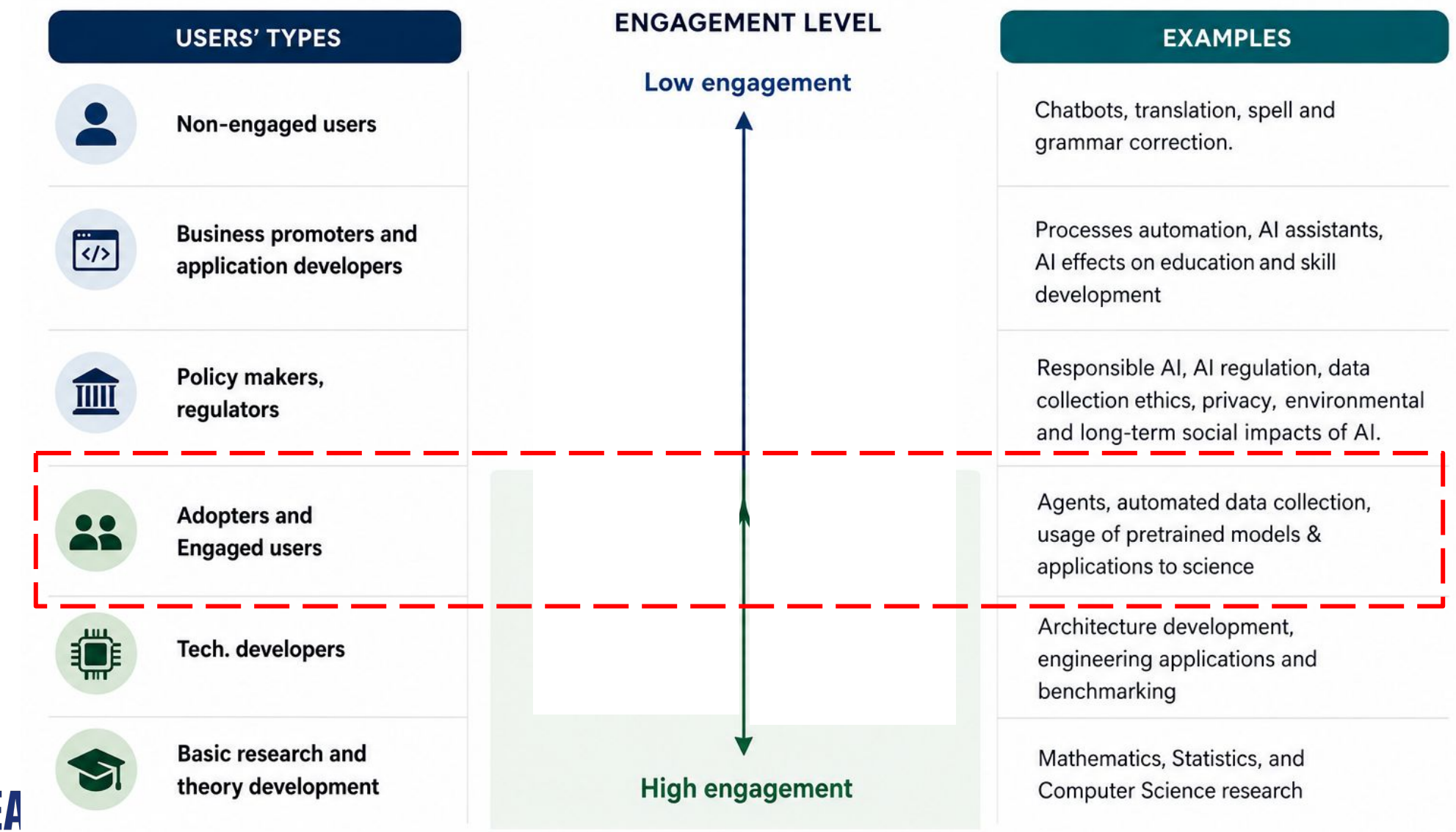


FELIPE LEONARDO GOMEZ CORTES
SENIOR RESEARCH ENGINEER

Computational Social Sciences and Humanities Laboratory - Data infrastructure Models and Methods

20 BSC YEARS

Multiple AI usages



Comprehensive bibliometric data



Title
Journal
Year
Authors
Affiliation
Abstract
Keywords
Topics
Methods
References
Citations

OpenAlex

OpenAlex

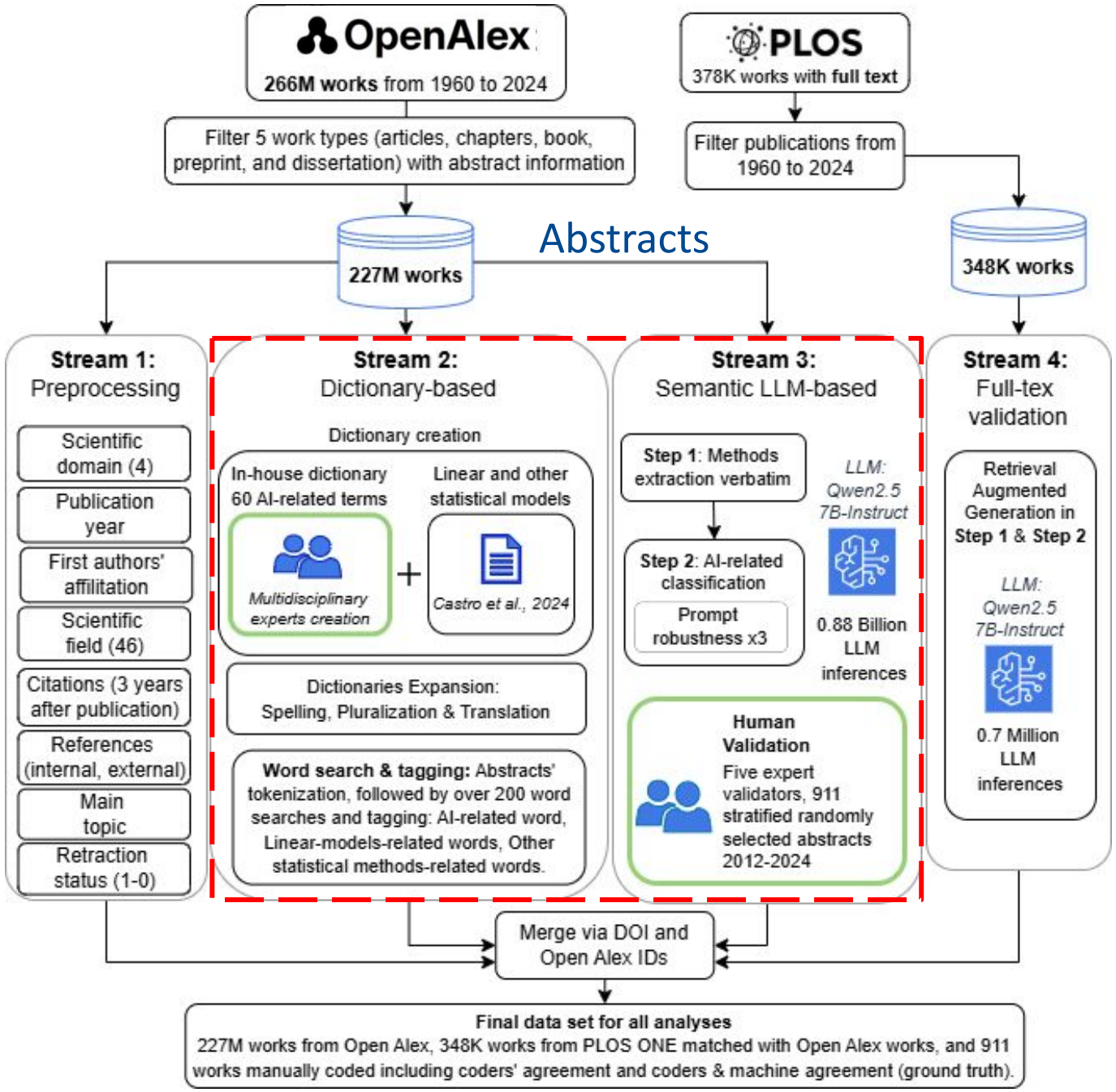
OpenAlex logo since 2025

Producer	OurResearch
History	2022; 4 years ago
Coverage	
Disciplines	Science, social science, arts, humanities (supports 256 disciplines)
Record depth	Citation indexing, author, topic title, subject keywords, abstract, periodical title, author's address, publication year
Format coverage	Articles, reviews, editorials, chronologies, abstracts, proceedings (journals and book-based), technical papers
Links	
Website	openalex.org

https://en.wikipedia.org/wiki/OpenAlex

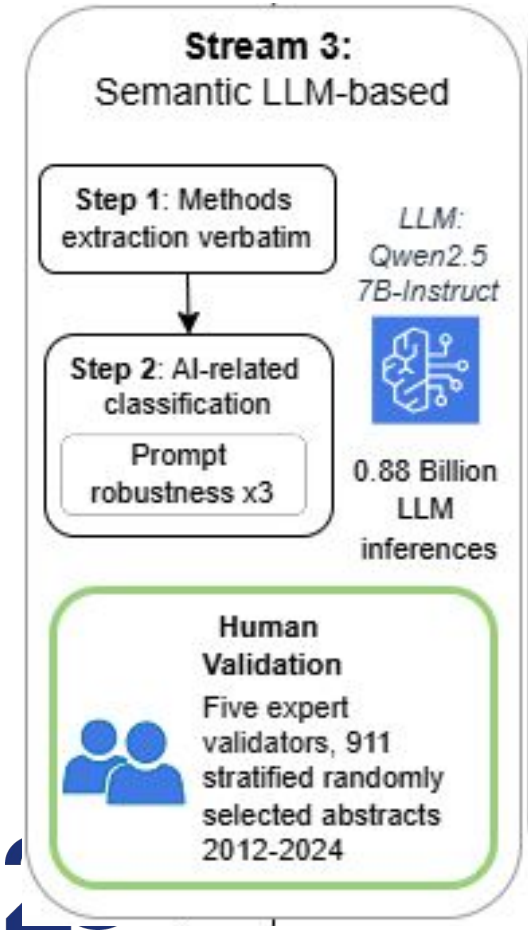
20 BSC YEARS

Computationally intensive data processing



BSC 20 YEARS

Example of methods extraction and classification.



Work

Family change and variation through the lens of family configurations in low- and middle-income countries

HTML

Year: 2021

Type: article

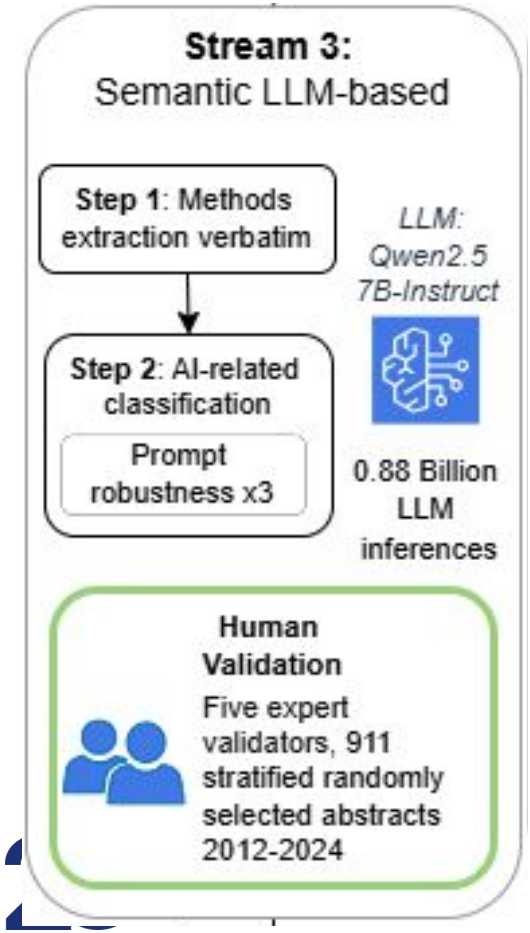
Abstract: Using 254 Demographic and Health Surveys from 75 low- and middle-income countries, this study shows how the joint examination of family characteristics across rural and urban areas provides new insights for understanding global family change. We operationalise this approach by building family configurations: a set of interrelated features that describe different patterns of family formation and structure. These features include partnership (marriage/unions) regimes and their stability, gender relations, household composition and reproduction. Factorial and clustering techniques allow us to summarise these family features into three factorial axes and six discrete family configurations. We provide an in-depth description of these configurations, their spatial distribution and their changes over time. Global family change is uneven because it emerges from complex interplays between the relative steadiness of longstanding arrangements for forming families and organising gender relations, and the rapidly changing dynamics observed in the realms of fertility, contraception, and timing of family formation.

(less)

Source: Population Space and Place

DOI: 10.1002/psp.2531

Example of methods extraction and classification.



{Factorial and clustering techniques}

Work

Family change and variation through the lens of family configurations in low- and middle-income countries

HTML

Year: 2021

Type: article

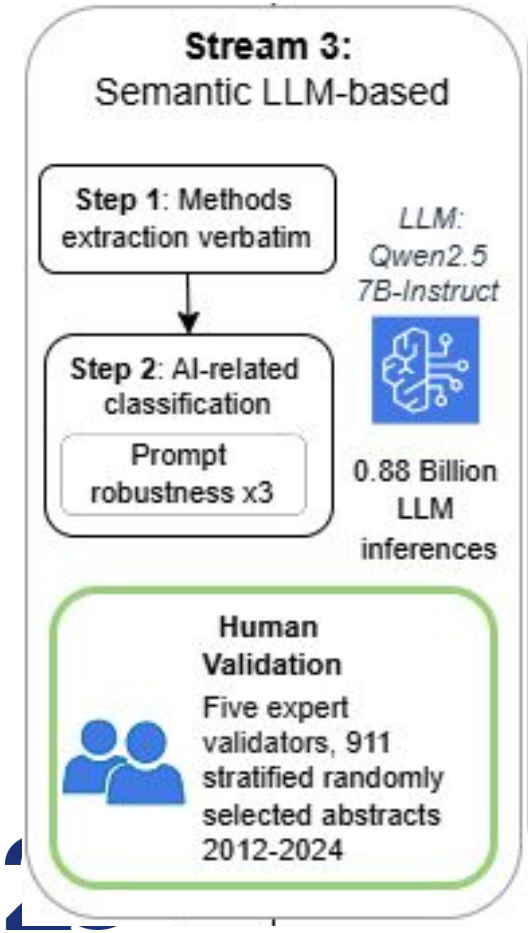
Abstract: Using 254 Demographic and Health Surveys from 75 low- and middle-income countries, this study shows how the joint examination of family characteristics across rural and urban areas provides new insights for understanding global family change. We operationalise this approach by building family configurations: a set of interrelated features that describe different patterns of family formation and structure. These features include partnership (marriage/unions) regimes and their stability, gender relations, household composition and reproduction. Factorial and clustering techniques allow us to summarise these family features into three factorial axes and six discrete family configurations. We provide an in-depth description of these configurations, their spatial distribution and their changes over time. Global family change is uneven because it emerges from complex interplays between the relative steadiness of longstanding arrangements for forming families and organising gender relations, and the rapidly changing dynamics observed in the realms of fertility, contraception, and timing of family formation.

(less)

Source: Population Space and Place

DOI: 10.1002/psp.2531

Example of methods extraction and classification.



{Factorial and clustering techniques}

Are any of these AI-based?

NO

Work

Family change and variation through the lens of family configurations in low- and middle-income countries

HTML

Year: 2021

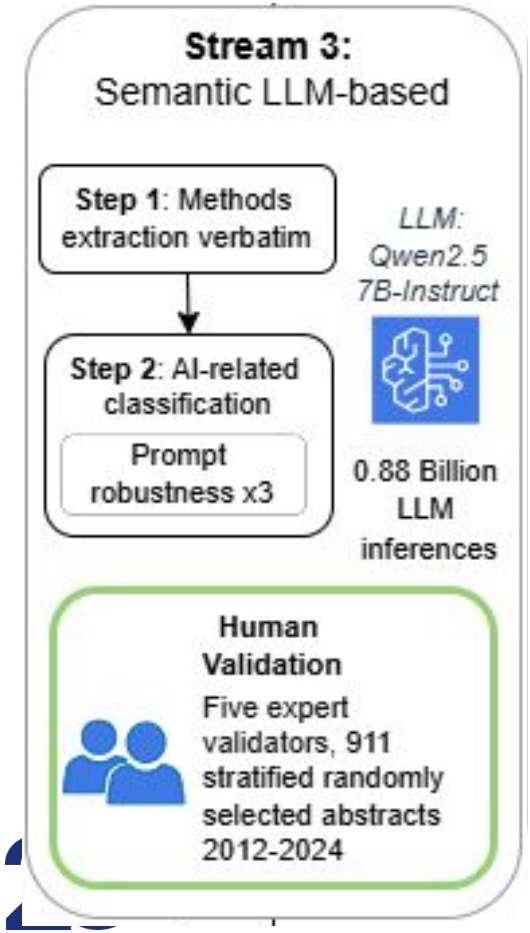
Type: article

Abstract: Using 254 Demographic and Health Surveys from 75 low- and middle-income countries, this study shows how the joint examination of family characteristics across rural and urban areas provides new insights for understanding global family change. We operationalise this approach by building family configurations: a set of interrelated features that describe different patterns of family formation and structure. These features include partnership (marriage/unions) regimes and their stability, gender relations, household composition and reproduction. Factorial and clustering techniques allow us to summarise these family features into three factorial axes and six discrete family configurations. We provide an in-depth description of these configurations, their spatial distribution and their changes over time. Global family change is uneven because it emerges from complex interplays between the relative steadiness of longstanding arrangements for forming families and organising gender relations, and the rapidly changing dynamics observed in the realms of fertility, contraception, and timing of family formation.
(less)

Source: Population Space and Place

DOI: 10.1002/psp.2531

Example of methods extraction and classification.



{Multiple Correspondence and Cluster Analysis….community detection algorithms}

Work

A global perspective on social stratification in science

HTML PDF

Year: 2024

Type: article

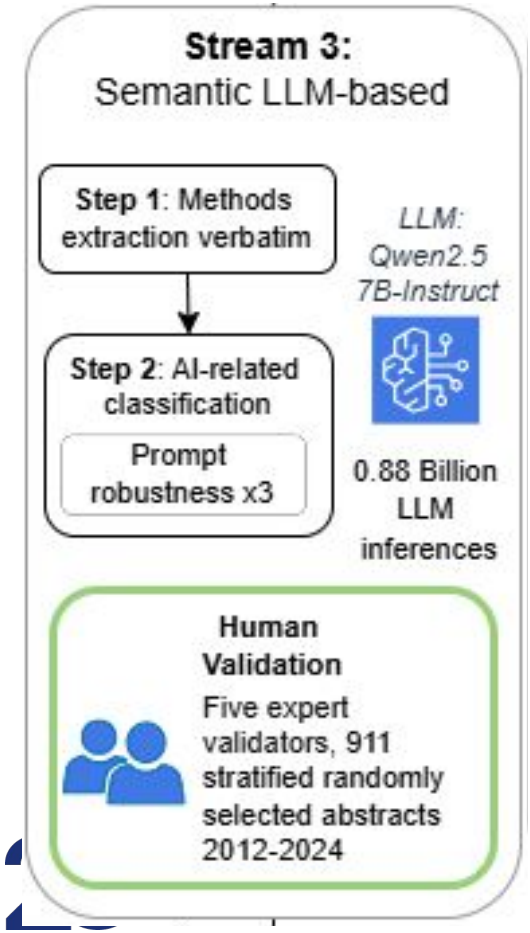
Abstract: To study stratification among scientists, we reconstruct the career-long trajectories of 8.2 million scientists worldwide using 12 bibliometric measures of productivity, geographical mobility, collaboration, and research impact. While most previous studies examined these variables in isolation, we study their relationships using Multiple Correspondence and Cluster Analysis. We group authors according to their bibliometric performance and academic age across six macro fields of science, and analyze co-authorship networks and detect collaboration communities of different sizes. We found a stratified structure in terms of academic age and bibliometric classes, with a small top class and large middle and bottom classes in all collaboration communities. Results are robust to community detection algorithms used and do not depend on authors' gender. These results imply that increased productivity, impact, and collaboration are driven by a relatively small group that accounts for a large share of academic outputs, i.e., the top class. Mobility indicators are the only exception with bottom classes contributing similar or larger shares. We also show that those at the top succeed by collaborating with various authors from other classes and age groups. Nevertheless, they are benefiting disproportionately from these collaborations which may have implications for persisting stratification in academia.

(less)

Source: Humanities and Social Sciences Communications

DOI: 10.1057/s41599-024-03402-w

Example of methods extraction and classification.



{Multiple Correspondence and Cluster Analysis….community detection algorithms}

Are any of these AI-based?

YES

Work

A global perspective on social stratification in science

HTML PDF

Year: 2024

Type: article

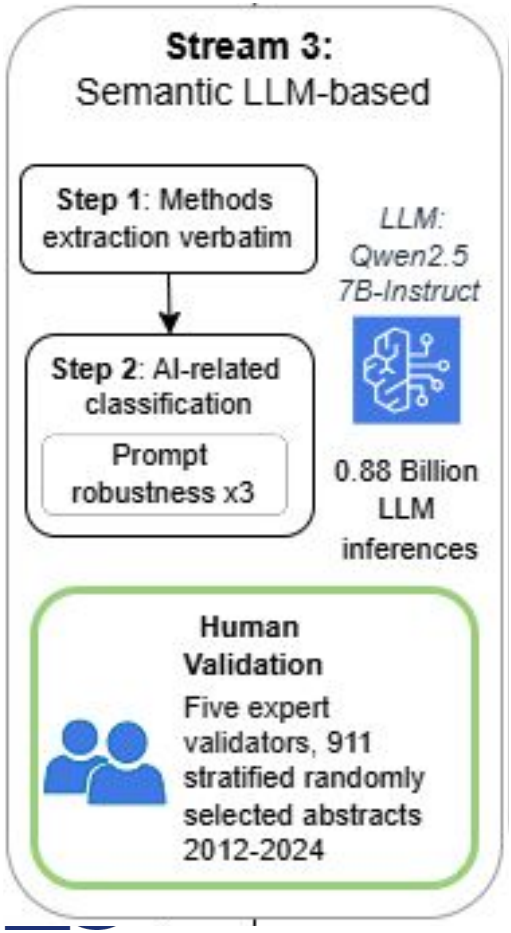
Abstract: To study stratification among scientists, we reconstruct the career-long trajectories of 8.2 million scientists worldwide using 12 bibliometric measures of productivity, geographical mobility, collaboration, and research impact. While most previous studies examined these variables in isolation, we study their relationships using Multiple Correspondence and Cluster Analysis. We group authors according to their bibliometric performance and academic age across six macro fields of science, and analyze co-authorship networks and detect collaboration communities of different sizes. We found a stratified structure in terms of academic age and bibliometric classes, with a small top class and large middle and bottom classes in all collaboration communities. Results are robust to community detection algorithms used and do not depend on authors' gender. These results imply that increased productivity, impact, and collaboration are driven by a relatively small group that accounts for a large share of academic outputs, i.e., the top class. Mobility indicators are the only exception with bottom classes contributing similar or larger shares. We also show that those at the top succeed by collaborating with various authors from other classes and age groups. Nevertheless, they are benefiting disproportionately from these collaborations which may have implications for persisting stratification in academia.

(less)

Source: Humanities and Social Sciences Communications

DOI: 10.1057/s41599-024-03402-w

Example of methods extraction and classification.



A Comprehensive Survey of IoT and Machine Learning Innovations in Agriculture

HTML PDF

Year: 2024

Type: article

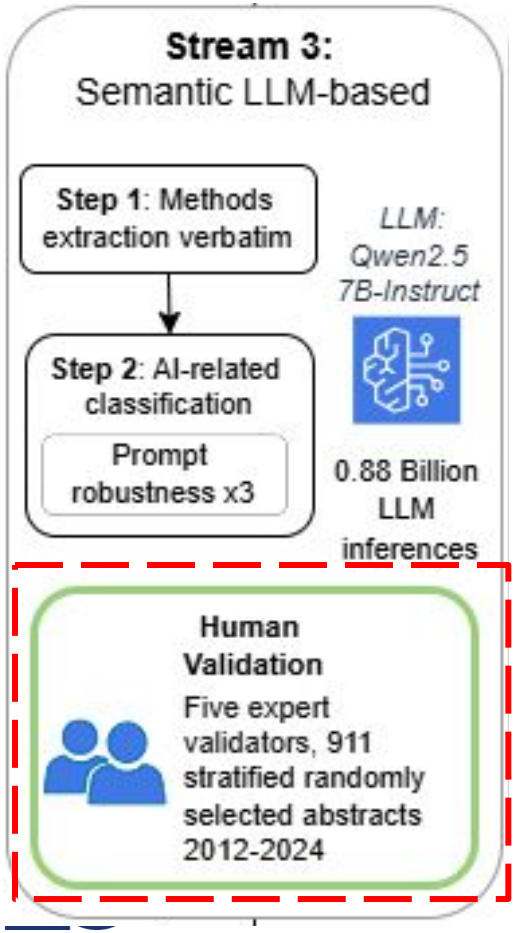
Abstract: The efficient management of soil moisture and water levels is crucial for optimizing crop yield and sustainable agricultural practices. This paper presents a comprehensive review of over 20 research studies focusing on the prediction of soil moisture levels and water requirements for crops using Internet of Things (IoT) devices and Machine Learning (ML) algorithms. By systematically analysing and synthesizing key findings from these studies, the paper highlights the methodologies, technologies, and algorithms that have been most effective in this domain. Artificial Intelligence (AI) and Machine Learnng enhances precision agriculture by enabling the analysis of large datasets to identify patterns and insights that improve decision-making. AI facilitates real-time monitoring and predictive analytics, optimizing resource usage and crop management. Through automation and intelligent systems, AI contributes to increased efficiency, reduced costs, and sustainable agricultural practices. Machine Learning (ML) algorithms, including regression, classification, and neural networks, are essential for modelling and predicting soil moisture levels and water requirements. These algorithms learn from historical and real-time data to provide accurate predictions and recommendations for irrigation scheduling. By continuously improving with new data, ML algorithms enhance the reliability and effectiveness of agricultural management systems. IoT devices, such as soil moisture sensors and weather stations, enable the continuous collection of real-time data from agricultural fields.

(less)

Source: International Research Journal on Advanced Science Hub

DOI: 10.47392/irjash.2024.032

Example of methods extraction and classification.



{comprehensive review}

Are any of these AI-based?

NO

A Comprehensive Survey of IoT and Machine Learning Innovations in Agriculture

HTML PDF

Year: 2024

Type: article

Abstract: The efficient management of soil moisture and water levels is crucial for optimizing crop yield and sustainable agricultural practices. This paper presents a comprehensive review of over 20 research studies focusing on the prediction of soil moisture levels and water requirements for crops using Internet of Things (IoT) devices and Machine Learning (ML) algorithms. By systematically analysing and synthesizing key findings from these studies, the paper highlights the methodologies, technologies, and algorithms that have been most effective in this domain. Artificial Intelligence (AI) and Machine Learnng enhances precision agriculture by enabling the analysis of large datasets to identify patterns and insights that improve decision-making. AI facilitates real-time monitoring and predictive analytics, optimizing resource usage and crop management. Through automation and intelligent systems, AI contributes to increased efficiency, reduced costs, and sustainable agricultural practices. Machine Learning (ML) algorithms, including regression, classification, and neural networks, are essential for modelling and predicting soil moisture levels and water requirements. These algorithms learn from historical and real-time data to provide accurate predictions and recommendations for irrigation scheduling. By continuously improving with new data, ML algorithms enhance the reliability and effectiveness of agricultural management systems. IoT devices, such as soil moisture sensors and weather stations, enable the continuous collection of real-time data from agricultural fields.
(less)

Source: International Research Journal on Advanced Science Hub

DOI: 10.47392/irjash.2024.032

Rapid adoption after the mid-1990s, followed by exponential growth from 2015 onward.

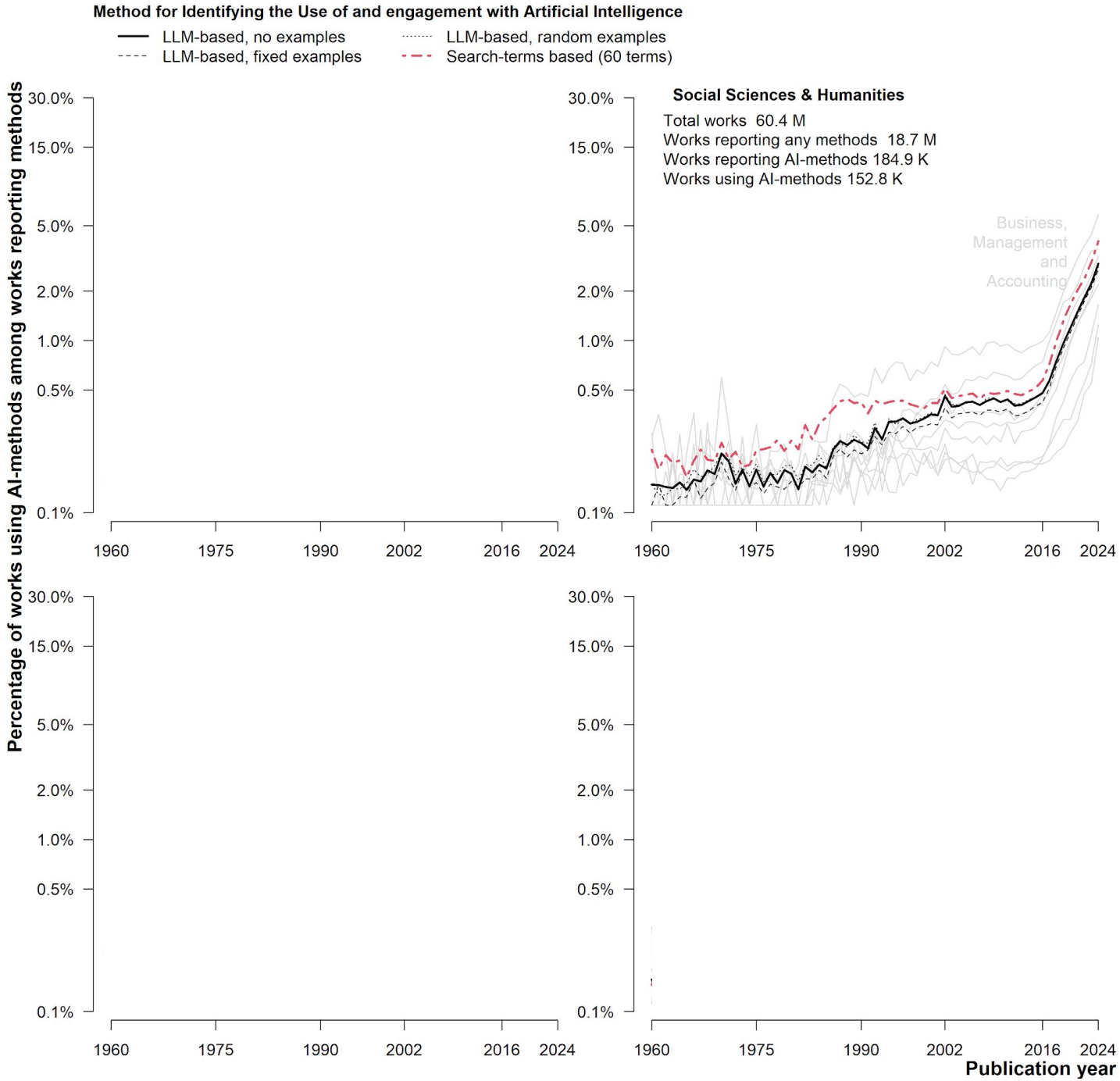


Figure 2. Log-scaled yearly percentage of academic works using AI methods (black and grey lines) and engaged with AI (red-dashed line) as a fraction of works using any method from 1960 to 2024 by scientific domains (bold lines) and 46 nested scientific fields and subfields (gray lines).

Source: Authors' calculation based on the Open Alex collection (n = 227,854,643 works, Data downloaded on Feb 27th, 2025).

20 YEARS BSC

Rapid adoption after the mid-1990s, followed by exponential growth from 2015 onward.

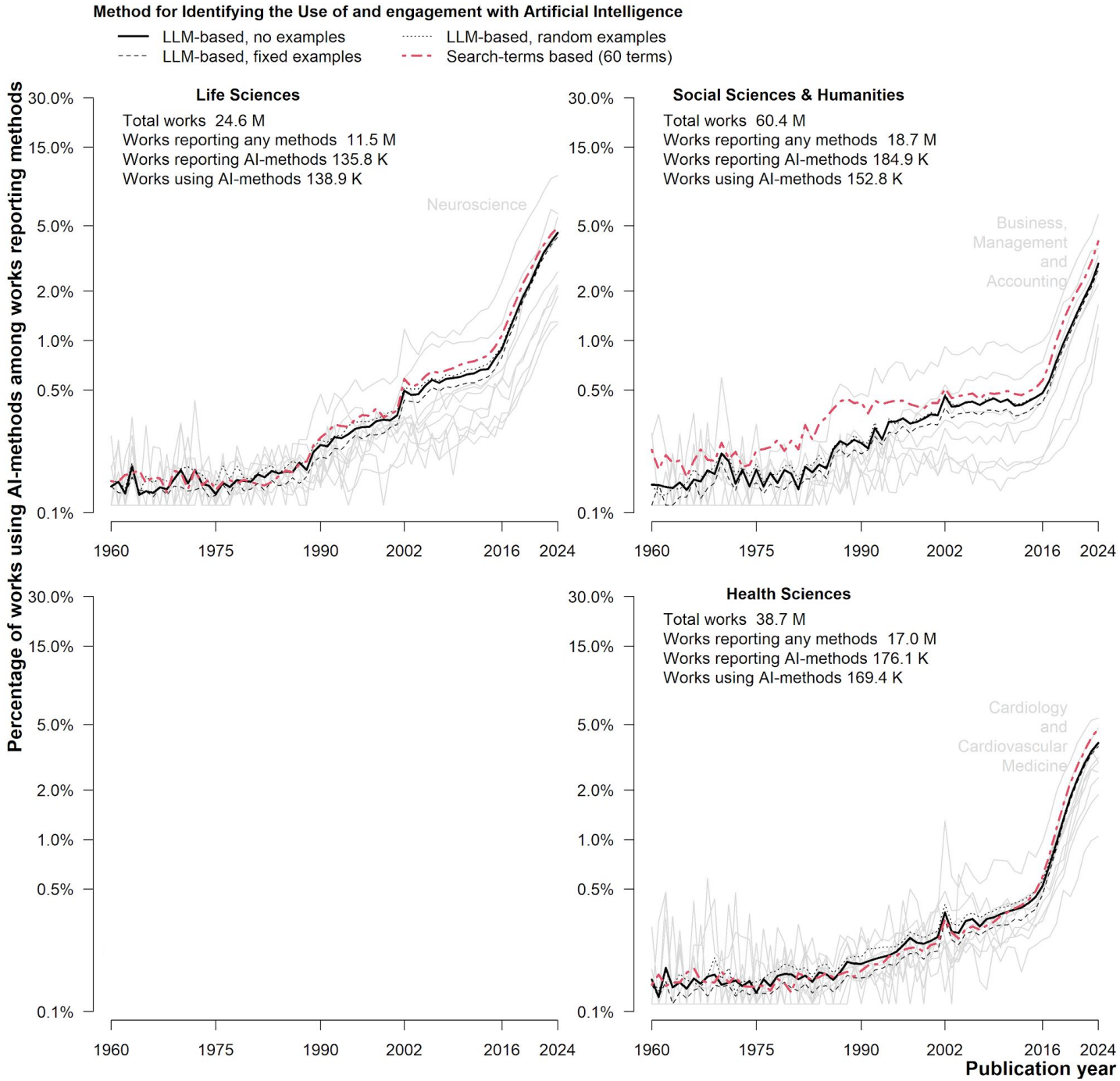


Figure 2. Log-scaled yearly percentage of academic works using AI methods (black and grey lines) and engaged with AI (red-dashed line) as a fraction of works using any method from 1960 to 2024 by scientific domains (bold lines) and 46 nested scientific fields and subfields (gray lines).

Source: Authors' calculation based on the Open Alex collection (n = 227,854,643 works, Data downloaded on Feb 27th, 2025).

20 BSC YEARS

Rapid adoption after the mid-1990s, followed by exponential growth from 2015 onward.

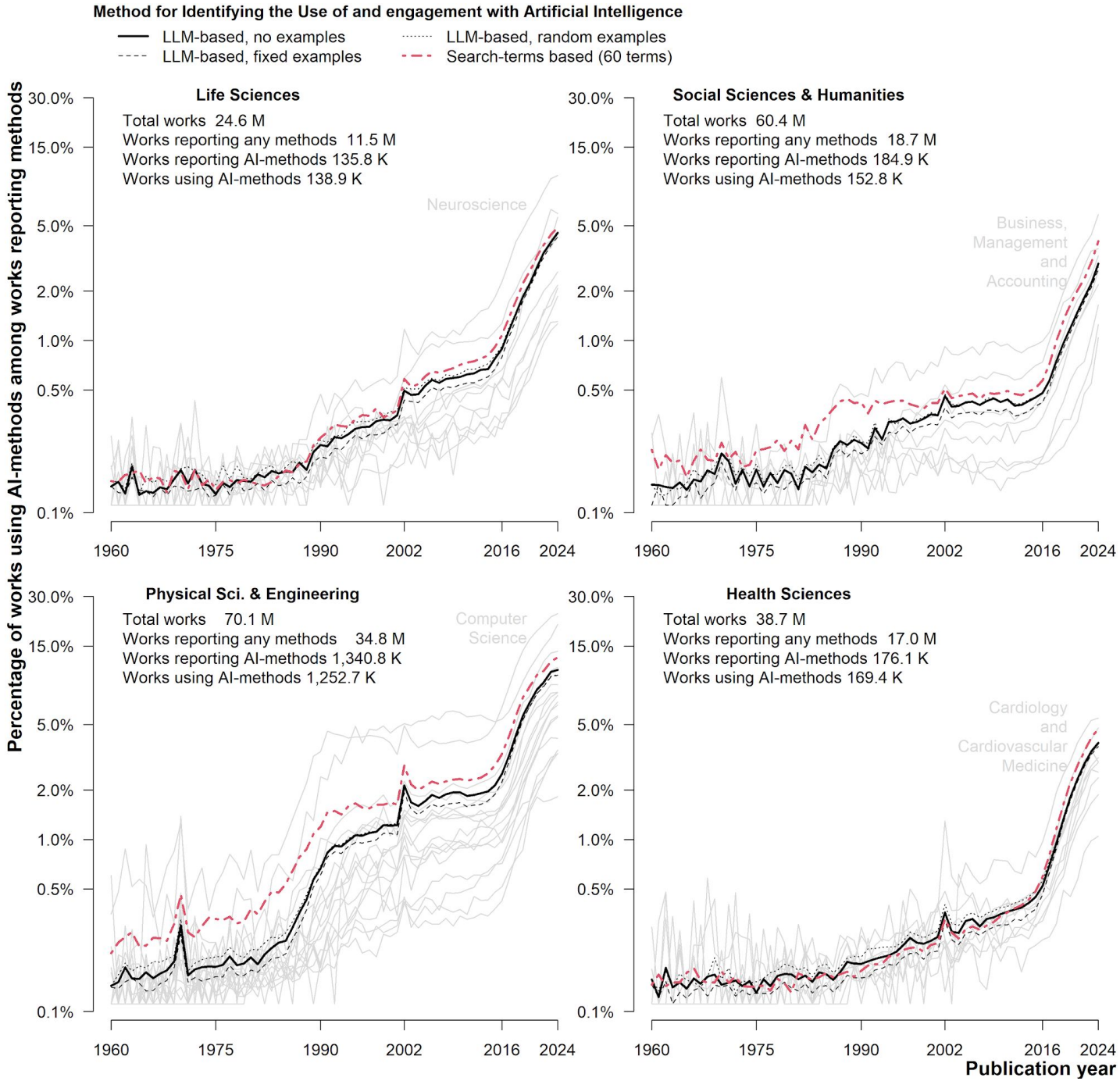
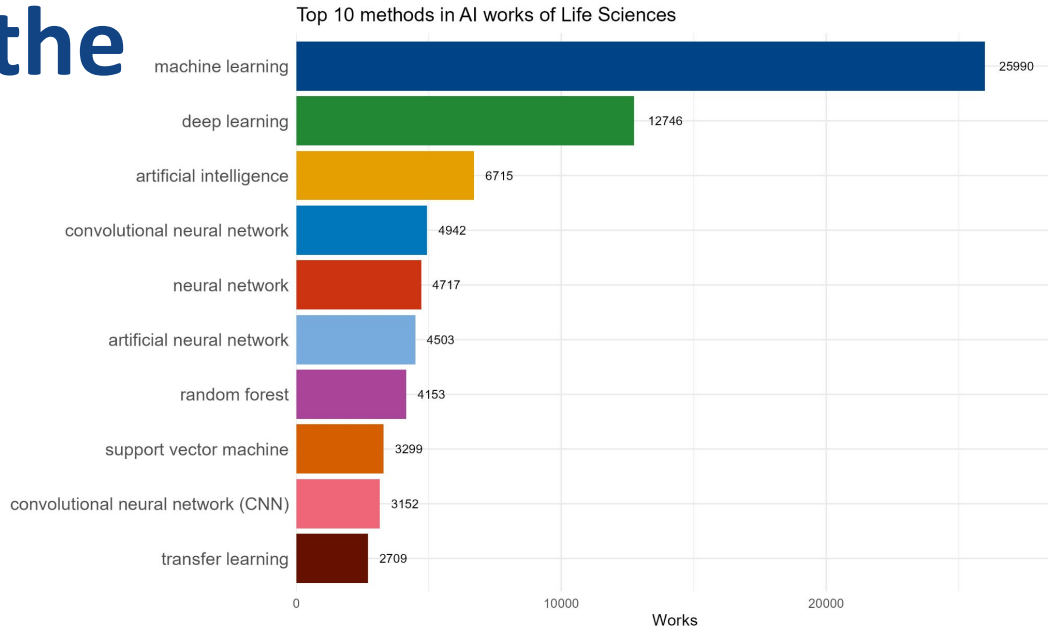


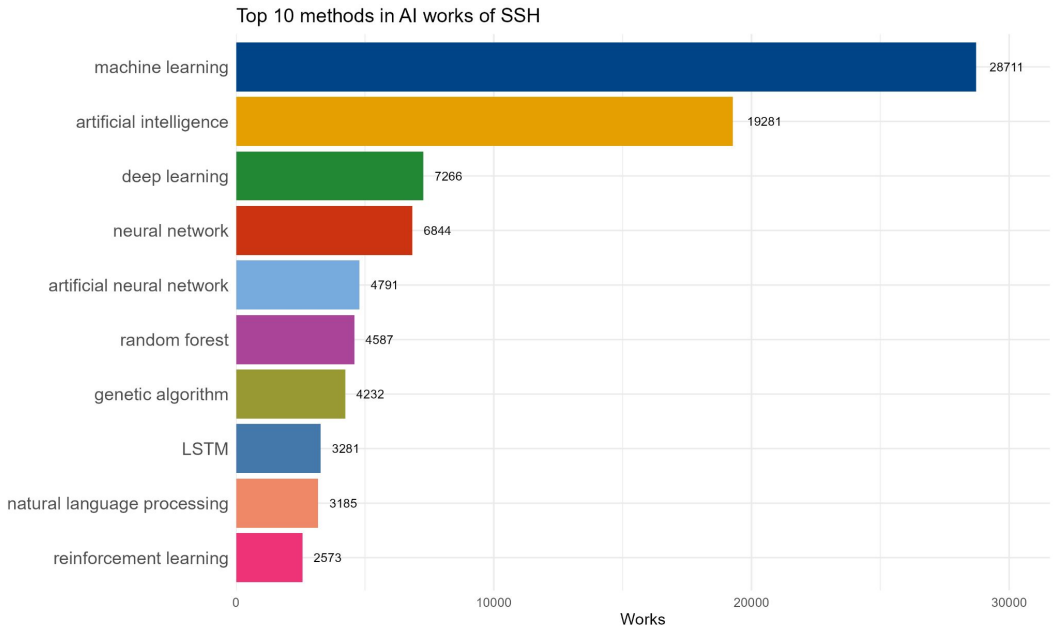
Figure 2. Log-scaled yearly percentage of academic works using AI methods (black and grey lines) and engaged with AI (red-dashed line) as a fraction of works using any method from 1960 to 2024 by scientific domains (bold lines) and 46 nested scientific fields and subfields (gray lines).

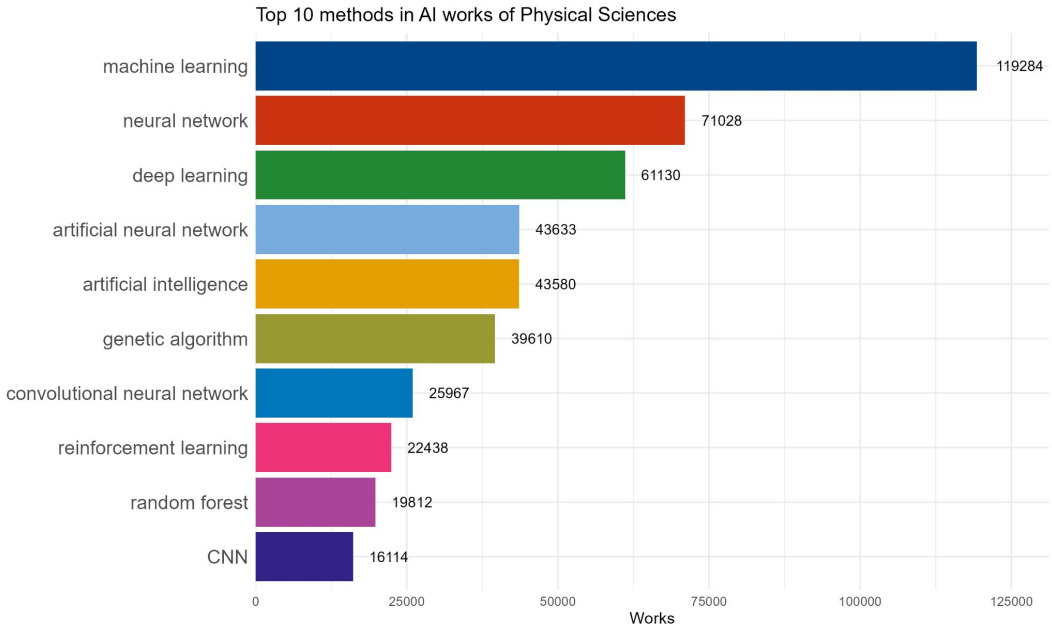
Source: Authors' calculation based on the Open Alex collection (n = 227,854,643 works, Data downloaded on Feb 27th, 2025).

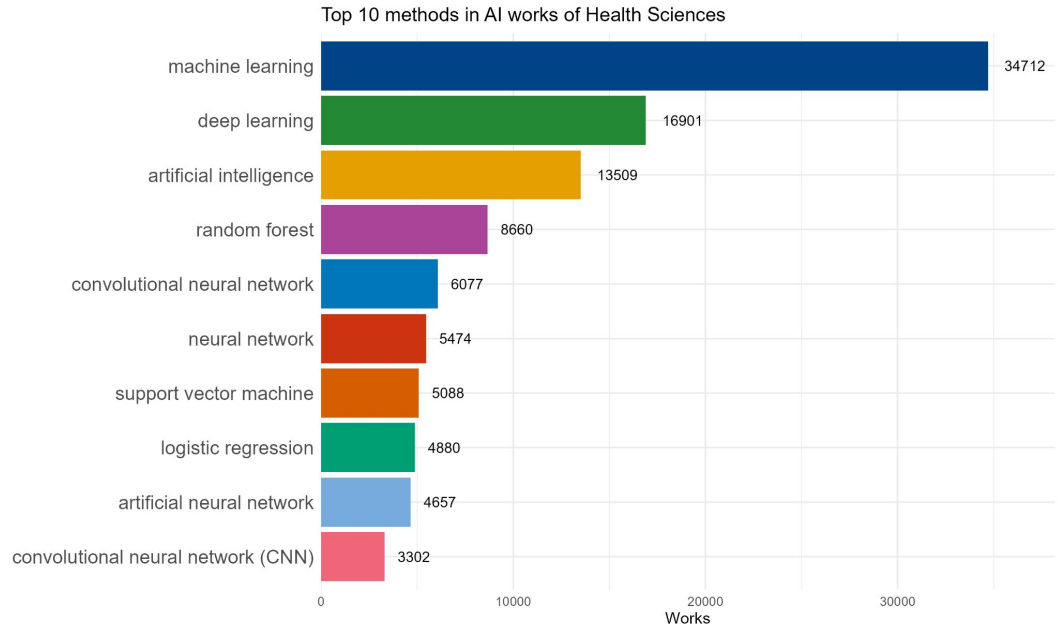
20 BSC YEARS

What are the top 10 AI methods?









20 YEARS BSC

Data analysis and visualizations developed by Paloma Navarro under PI's supervision

A patterns that holds across subfields in all scientific domains, with only a few exceptions…



Figure 3. Yearly percentage of scientific works using AI methods as a fraction of works using any method for four scientific domains (colored lines), 46 scientific fields and subfields (black lines), and for works with missing scientific domain information from 1985 to 2024.

The total number of academic works per field and the growth of AI use refer to the period 2005 to 2024 (n = 143,232,964 works).

20 BSC YEARS

A patterns that holds across subfields in all scientific domains, with only a few exceptions…

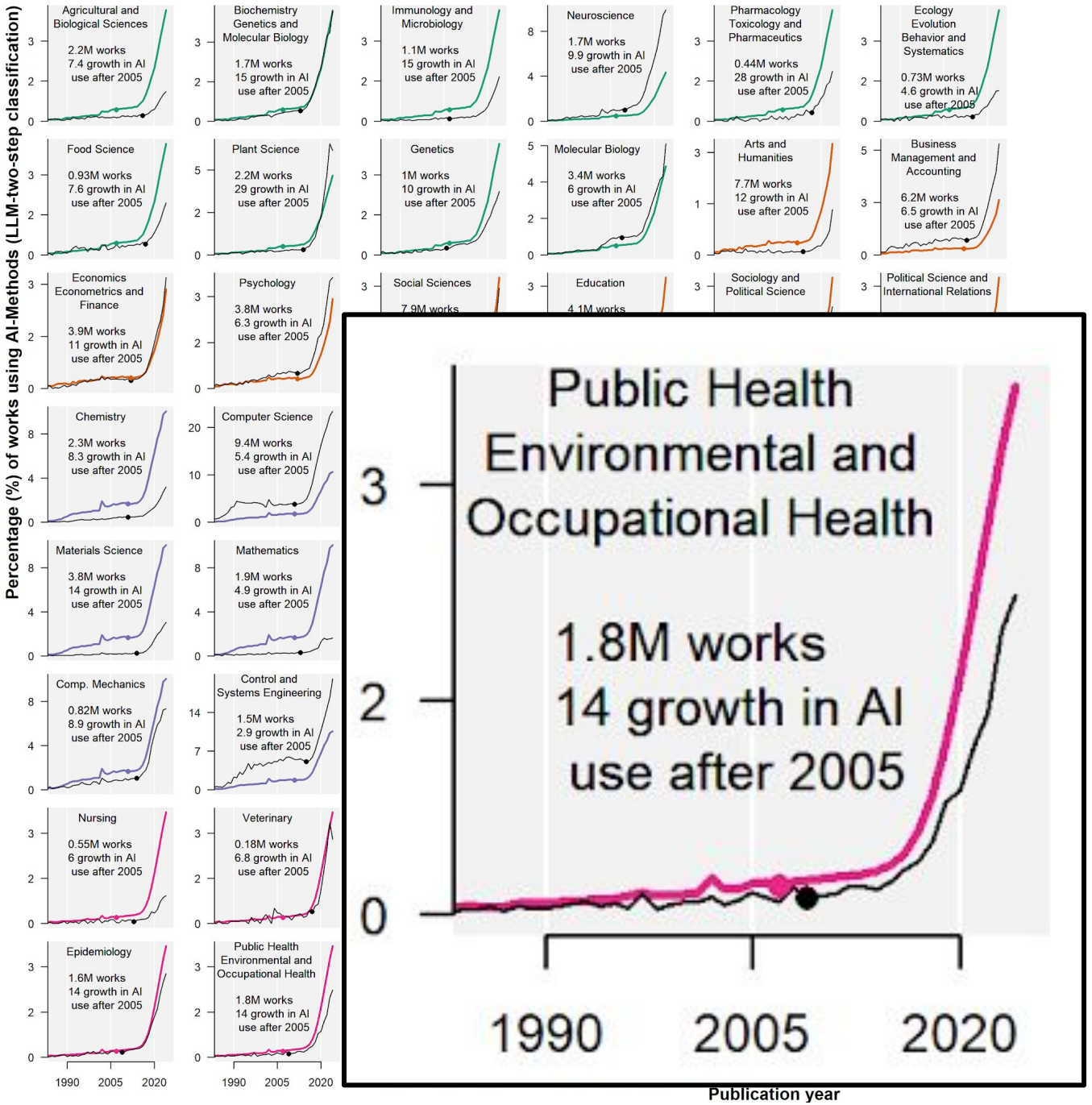


Figure 3. Yearly percentage of scientific works using AI methods as a fraction of works using any method for four scientific domains (colored lines), 46 scientific fields and subfields (black lines), and for works with missing scientific domain information from 1985 to 2024.

The total number of academic works per field and the growth of AI use refer to the period 2005 to 2024 (n = 143,232,964 works).

20 BSC YEARS

A patterns that holds across subfields in all scientific domain with only a few exceptions…

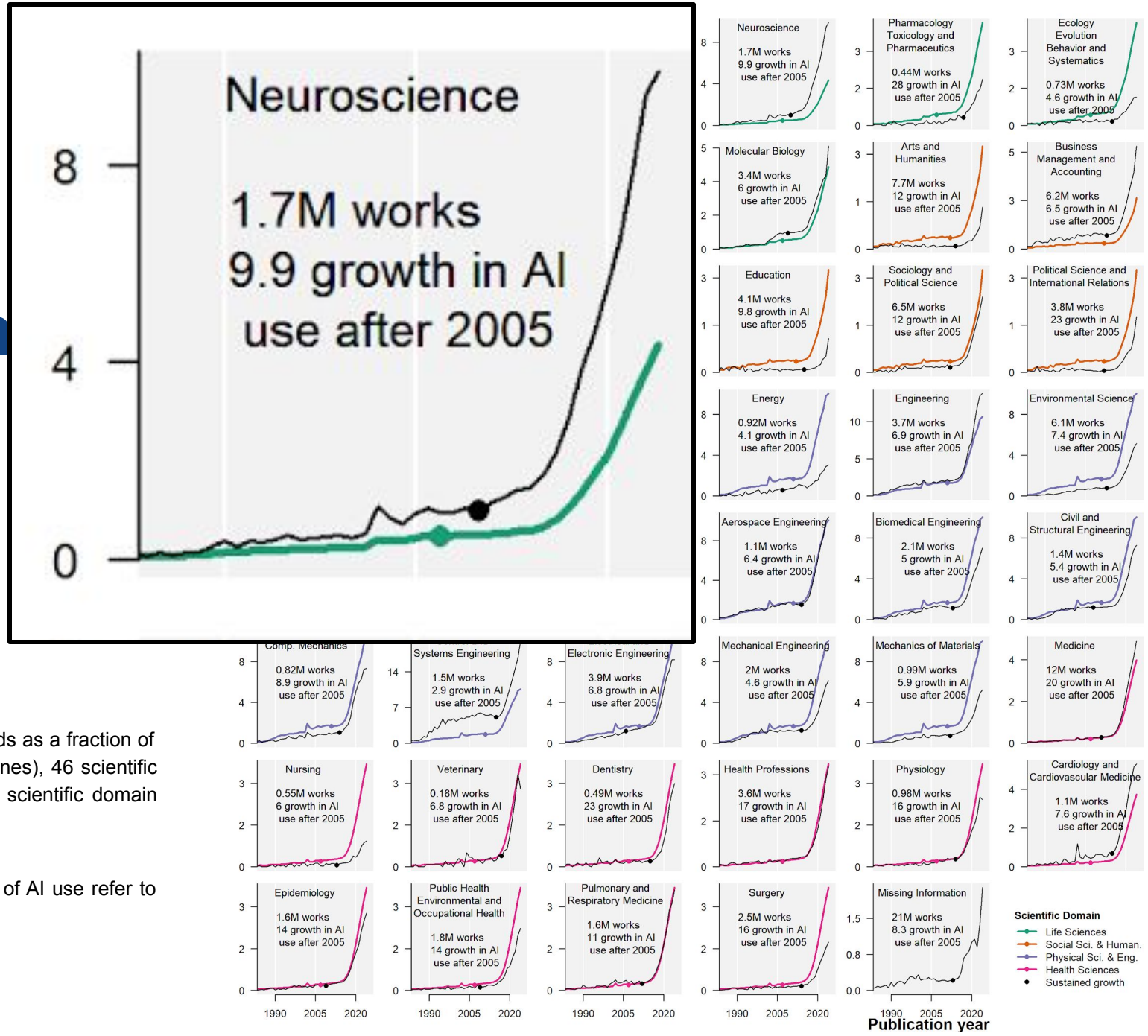


Figure 3. Yearly percentage of scientific works using AI methods as a fraction of works using any method for four scientific domains (colored lines), 46 scientific fields and subfields (black lines), and for works with missing scientific domain information from 1985 to 2024.

The total number of academic works per field and the growth of AI use refer to the period 2005 to 2024 (n = 143,232,964 works).

20 BSC YEARS

A patterns that holds across subfields in all scientific domains, with only a few exceptions…

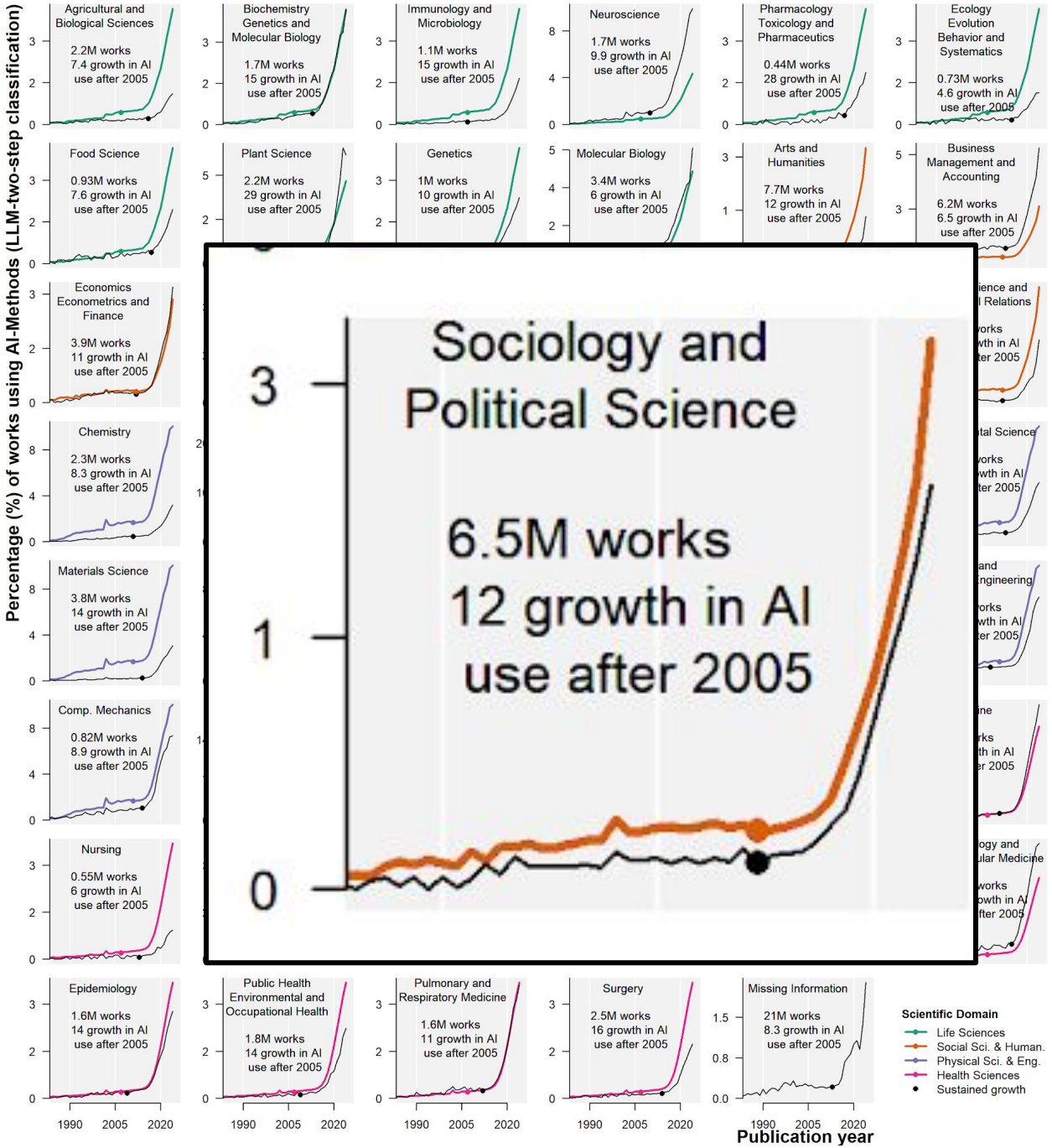


Figure 3. Yearly percentage of scientific works using AI methods as a fraction of works using any method for four scientific domains (colored lines), 46 scientific fields and subfields (black lines), and for works with missing scientific domain information from 1985 to 2024.

The total number of academic works per field and the growth of AI use refer to the period 2005 to 2024 (n = 143,232,964 works).

20 BSC YEARS

A patterns that holds across subfields in all scientific domains, with only a few exceptions…

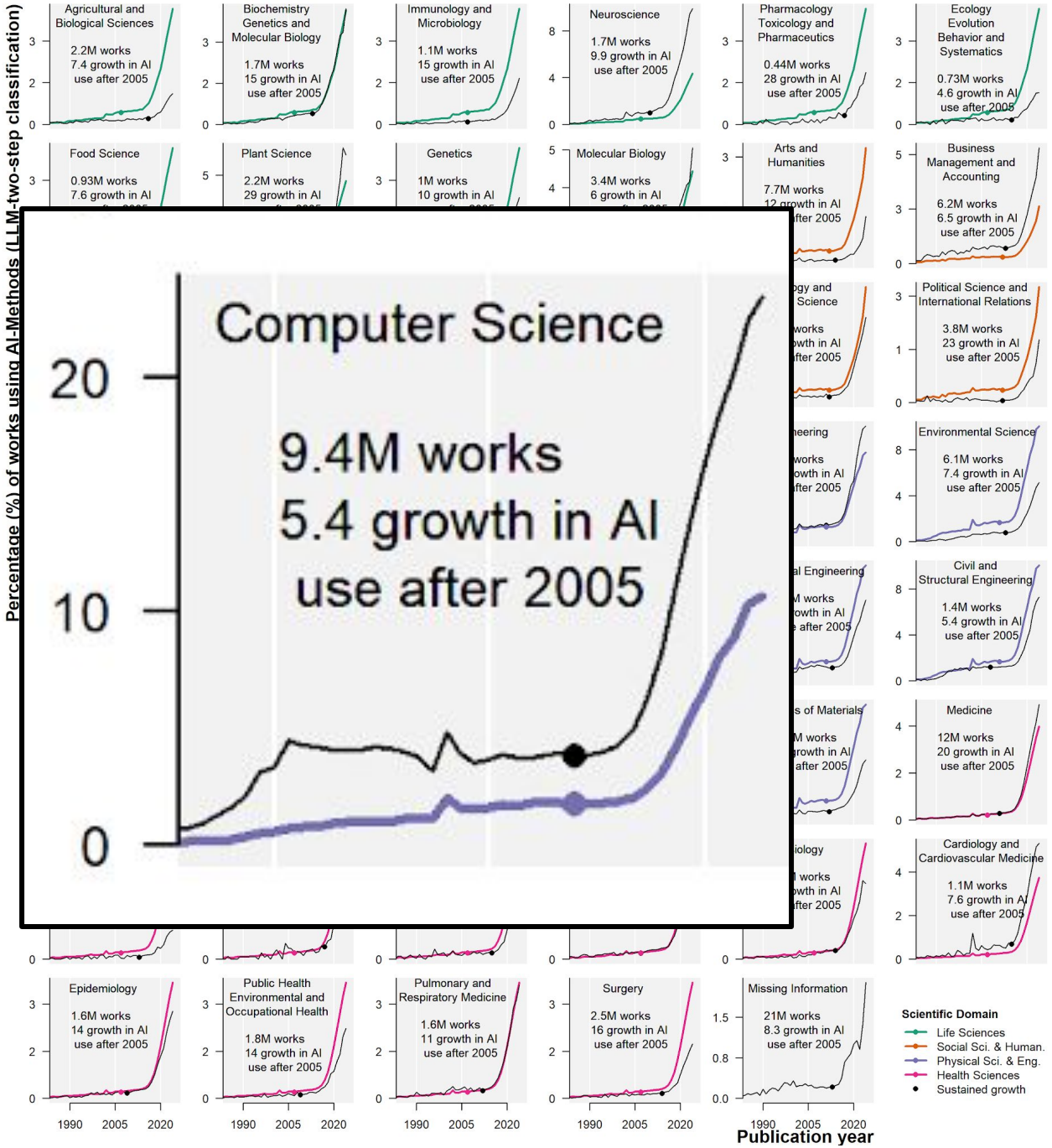


Figure 3. Yearly percentage of scientific works using AI methods as a fraction of works using any method for four scientific domains (colored lines), 46 scientific fields and subfields (black lines), and for works with missing scientific domain information from 1985 to 2024.

The total number of academic works per field and the growth of AI use refer to the period 2005 to 2024 (n = 143,232,964 works).

20 BSC YEARS

A patterns that holds across subfields in all scientific domains, with only a few exceptions…

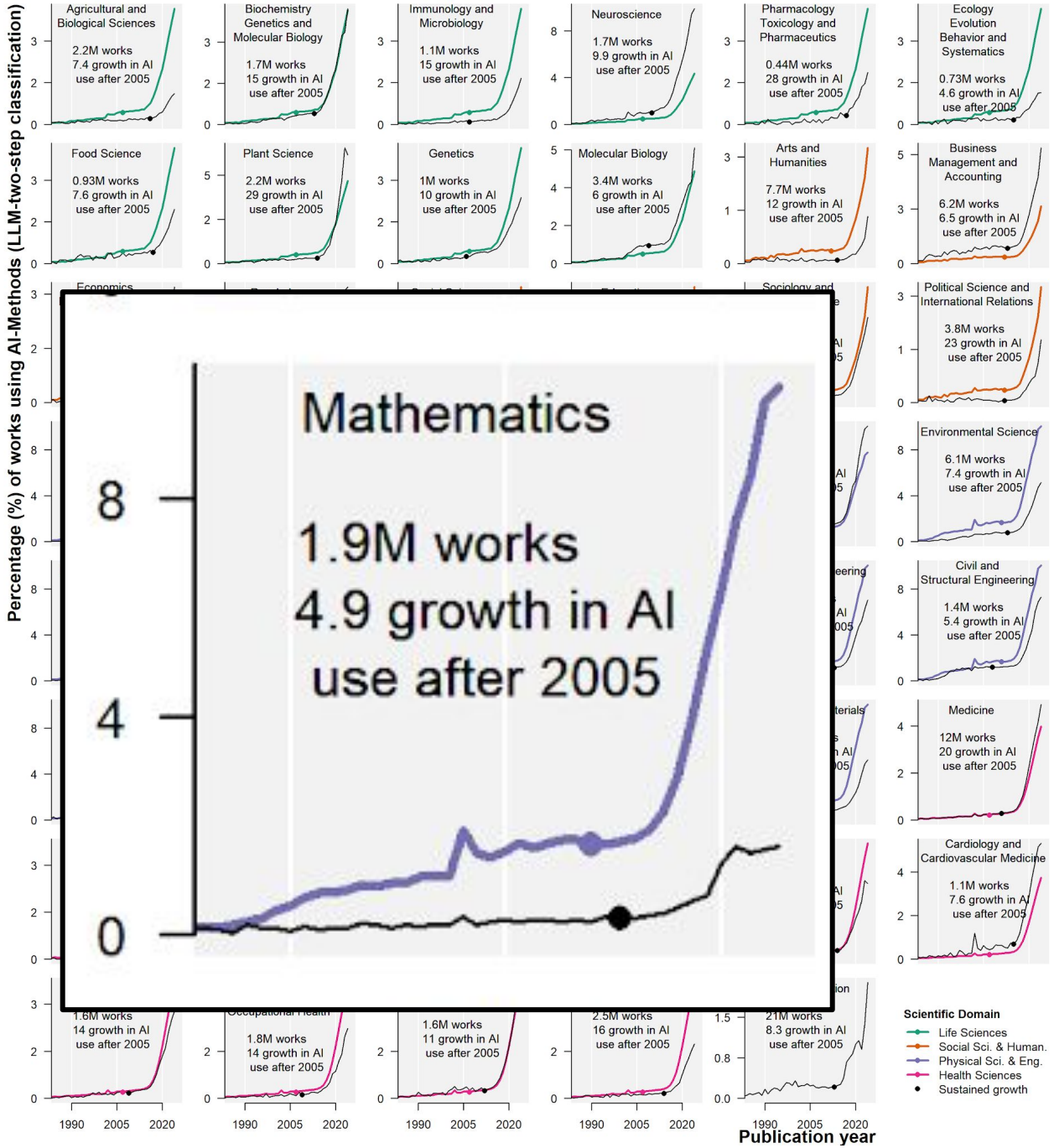


Figure 3. Yearly percentage of scientific works using AI methods as a fraction of works using any method for four scientific domains (colored lines), 46 scientific fields and subfields (black lines), and for works with missing scientific domain information from 1985 to 2024.

The total number of academic works per field and the growth of AI use refer to the period 2005 to 2024 (n = 143,232,964 works).

Life Sciences

Social and Hum.

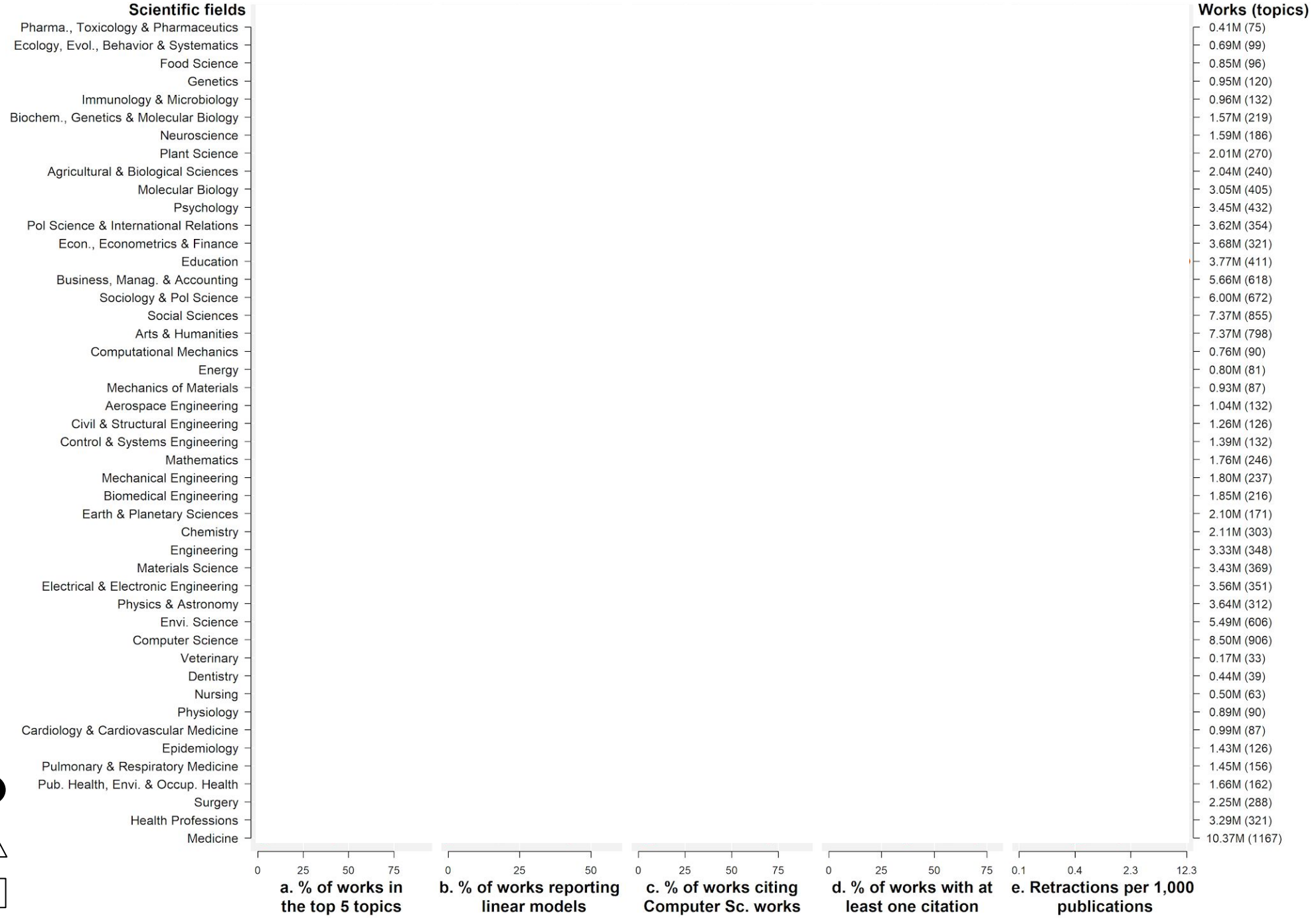
Physical and Eng.

Medical Sciences

AI methods ●

Non-AI methods △

No methods □



Life Sciences

Social and Hum.

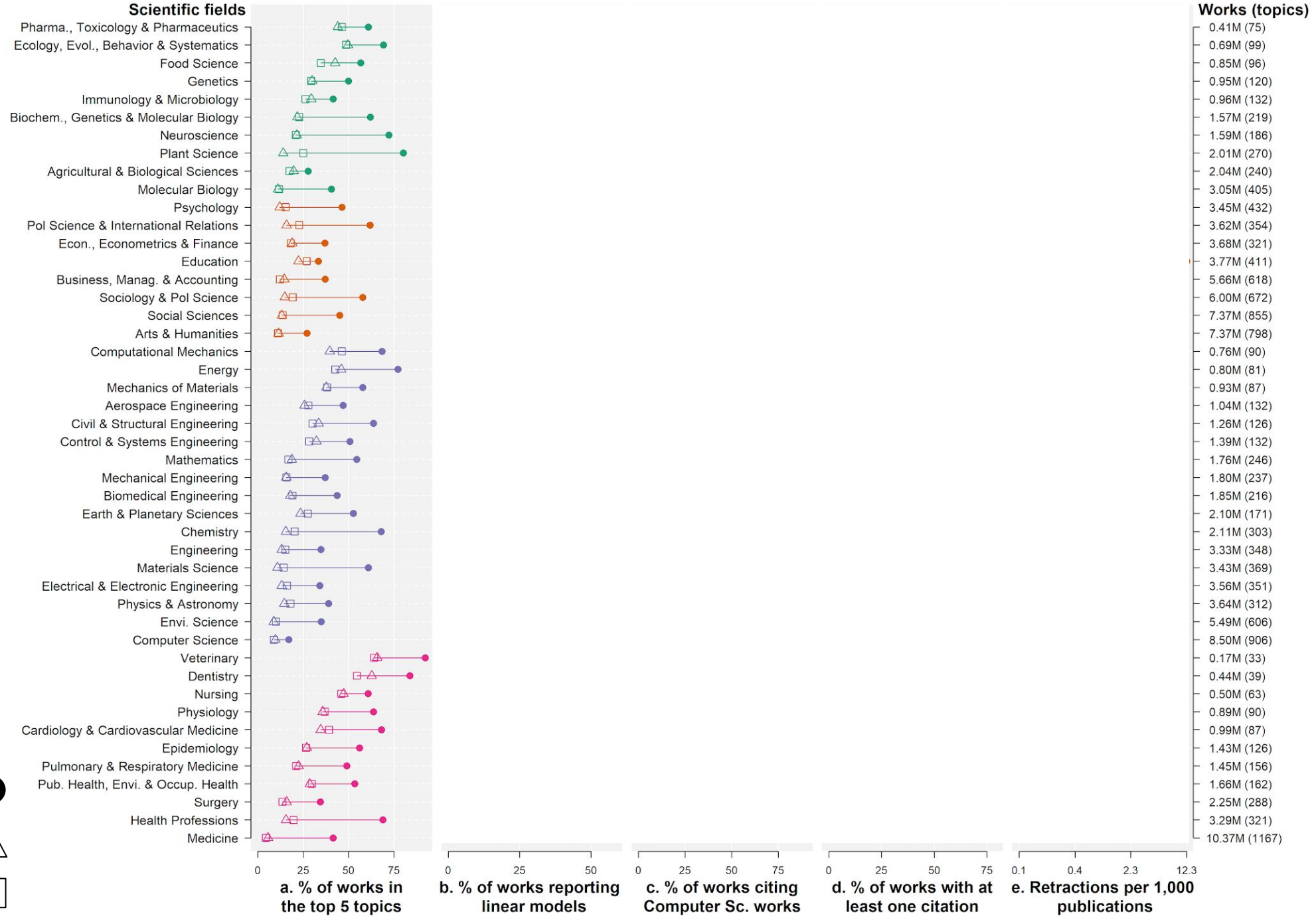
Physical and Eng.

Medical Sciences

AI methods ●

Non-AI methods △

No methods □



Life Sciences

Social and Hum.

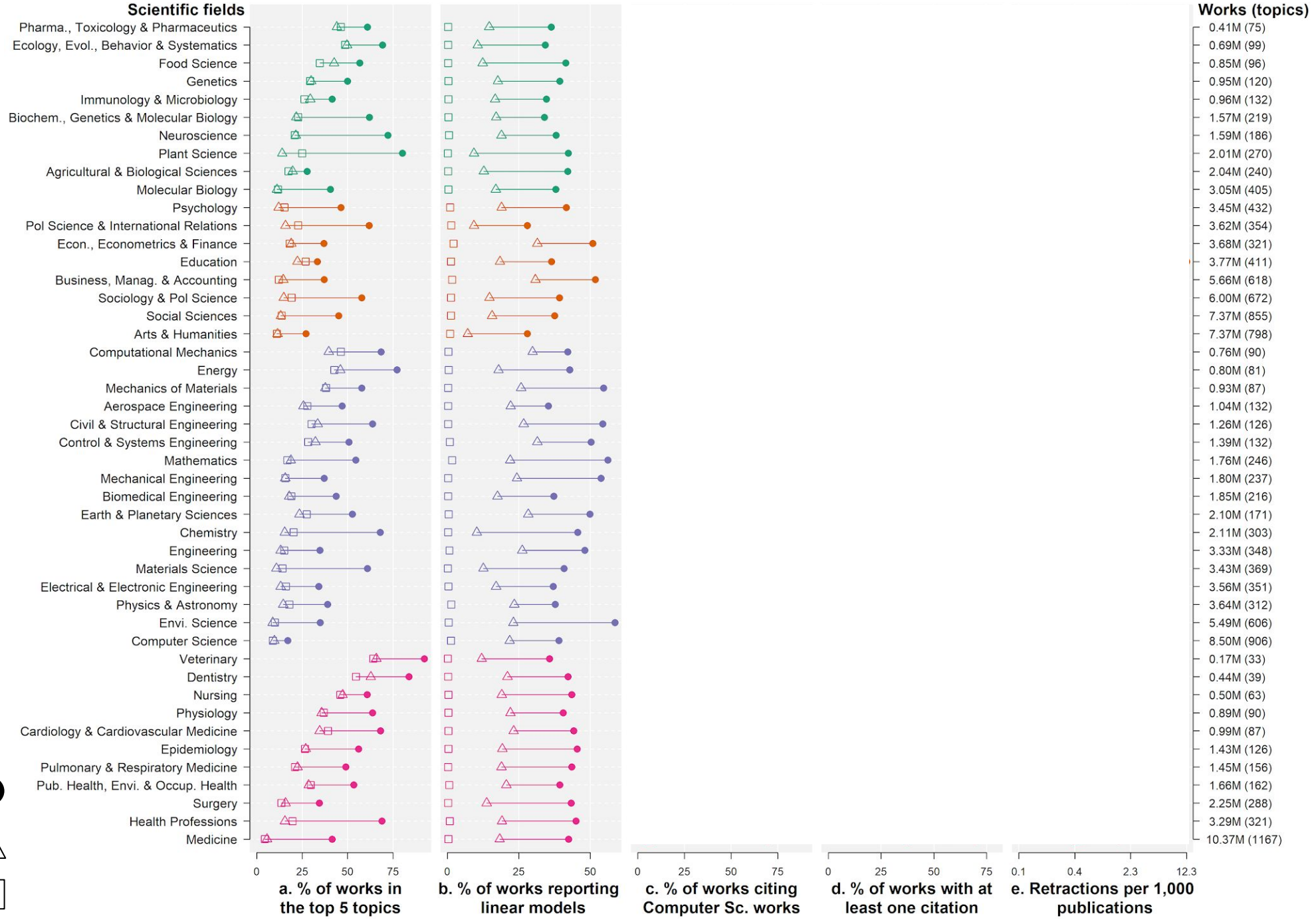
Physical and Eng.

Medical Sciences

AI methods ●

Non-AI methods △

No methods □



Life Sciences

Social and Hum.

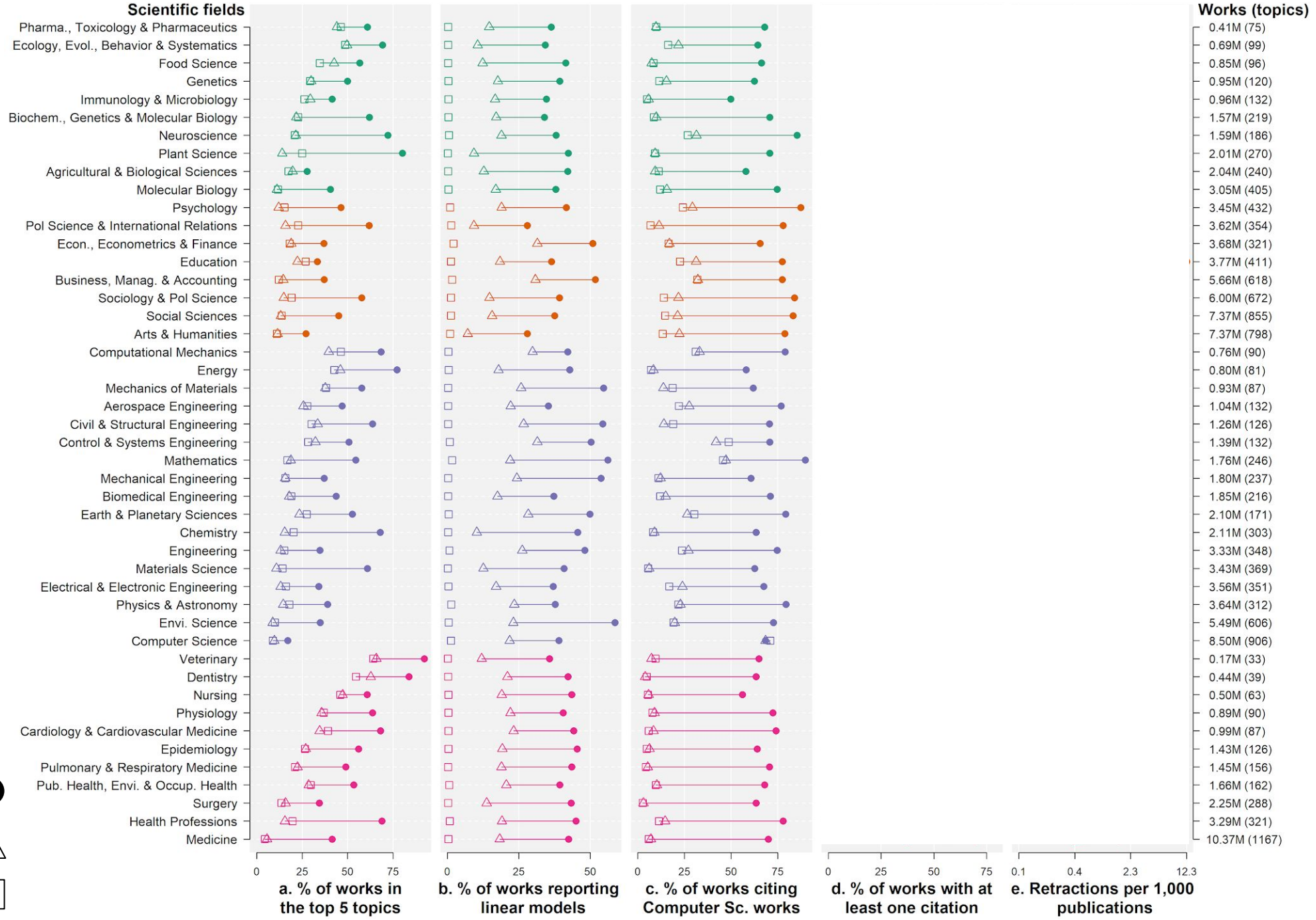
Physical and Eng.

Medical Sciences

AI methods ●

Non-AI methods △

No methods □



Life Sciences

Social and Hum.

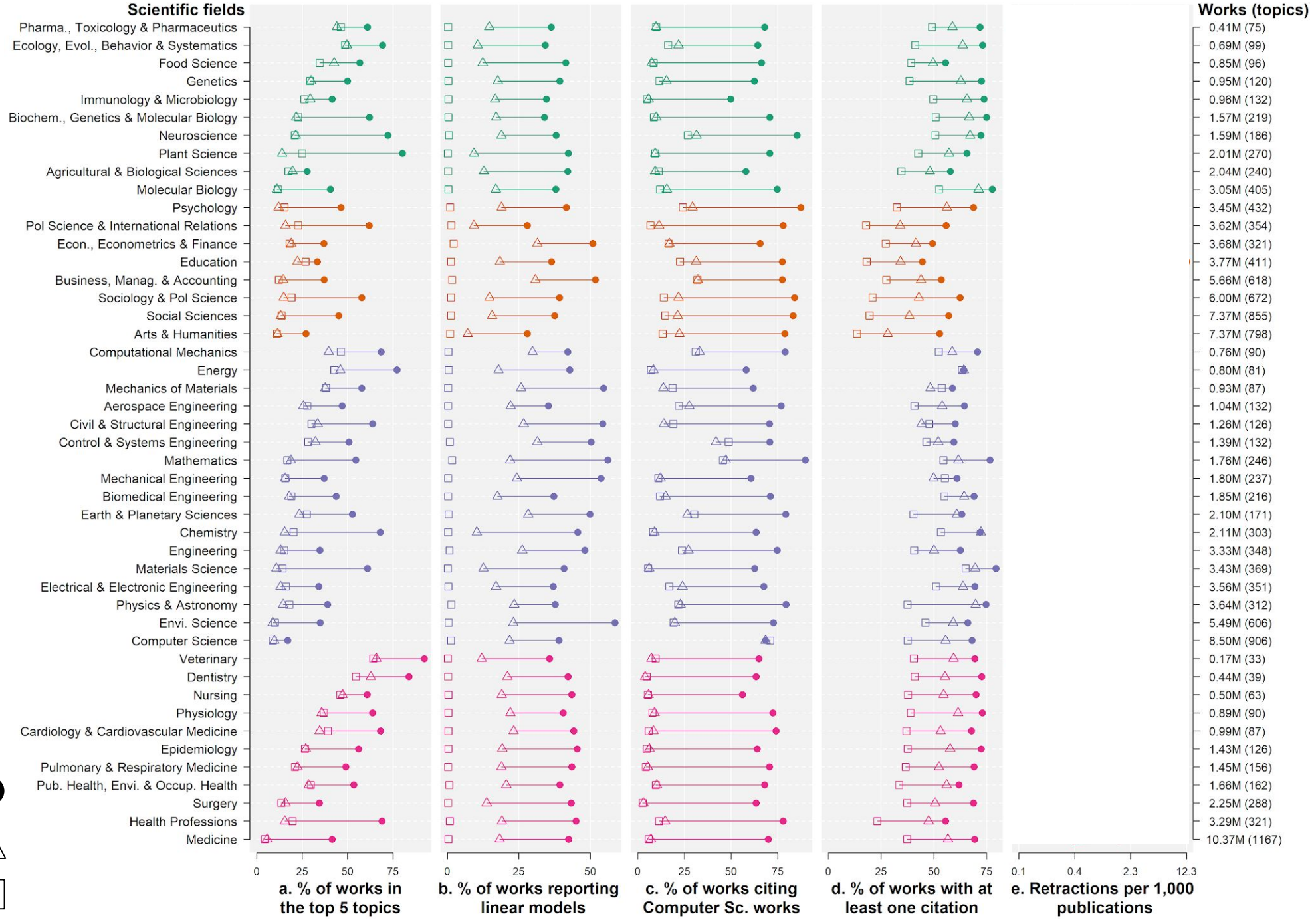
Physical and Eng.

Medical Sciences

AI methods ●

Non-AI methods △

No methods □



Life Sciences

Social and Hum.

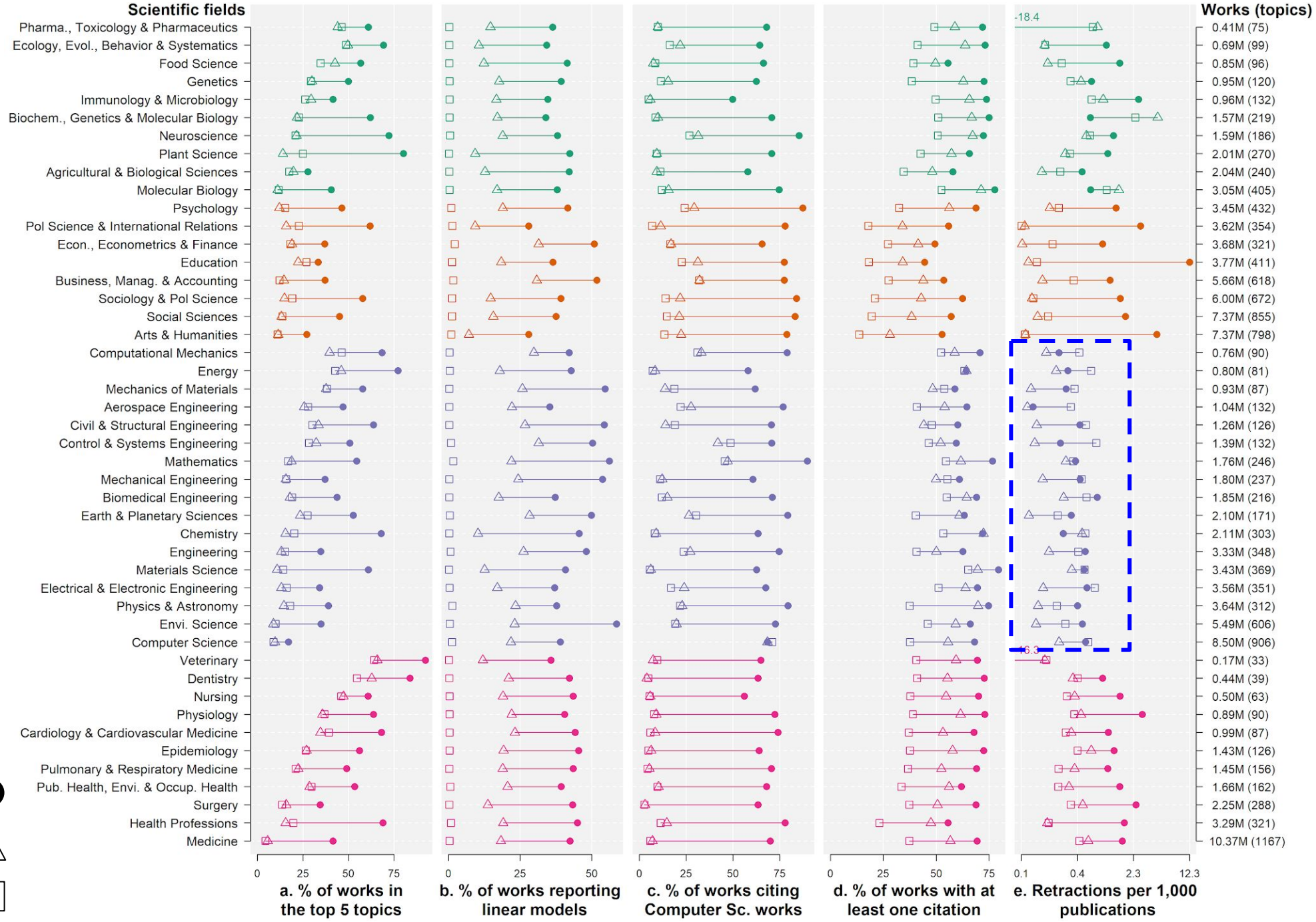
Physical and Eng.

Medical Sciences

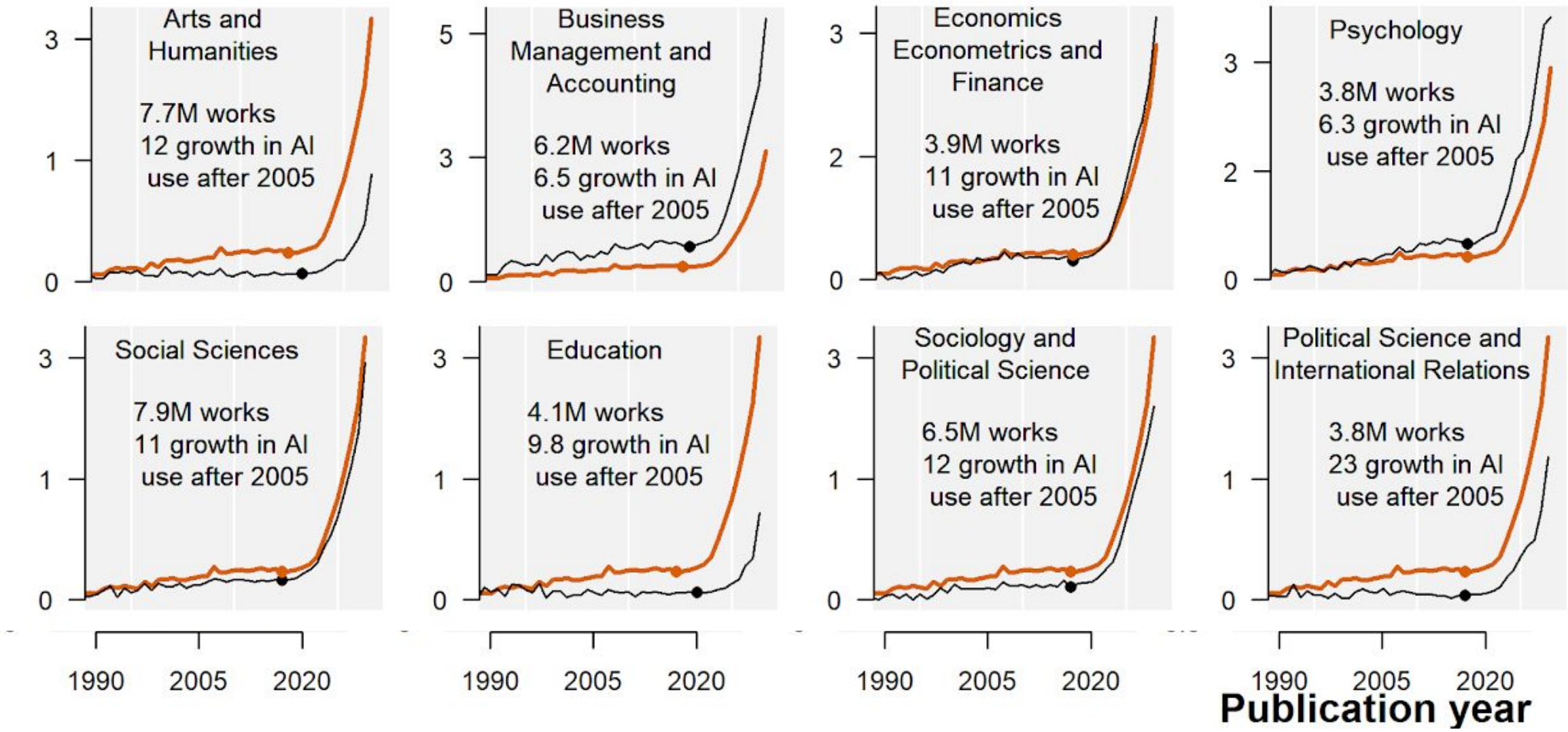
AI methods ●

Non-AI methods △

No methods □

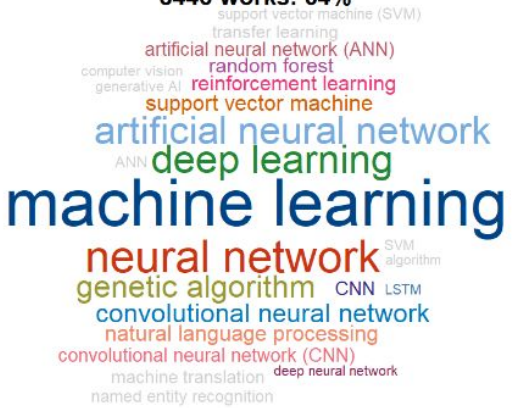


Percentage of works using AI-based methods across **Social Sciences** subfields and the **Humanities**, 2002-2024

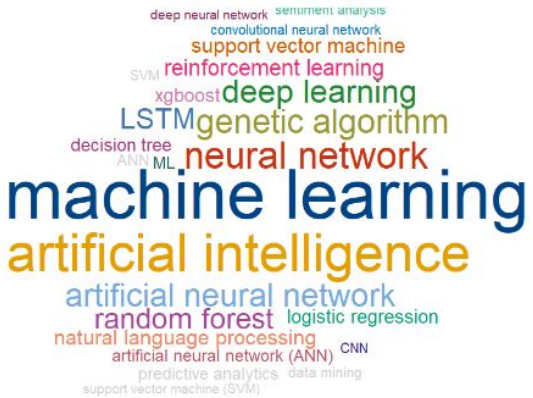


Extracted methods from SSH abstracts by fields

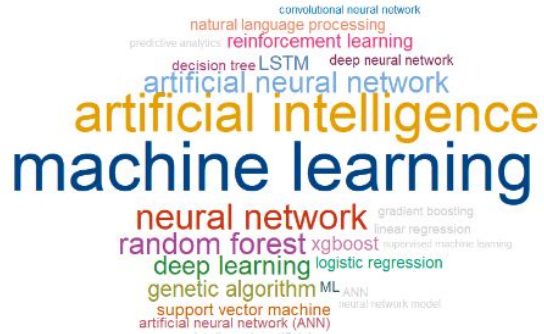
Arts and Humanities
8446 works: 64%



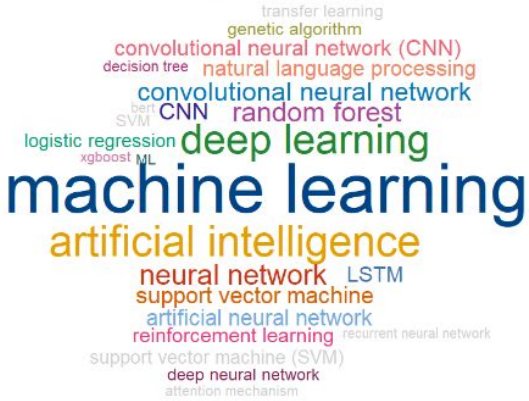
Business, Management and Accounting
63455 works: 60%



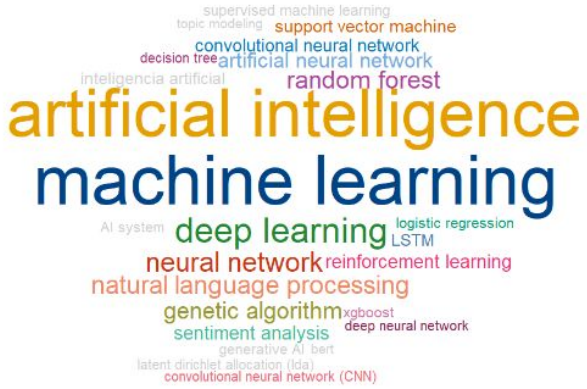
Economics, Econometrics and Finance
16011 works: 68%



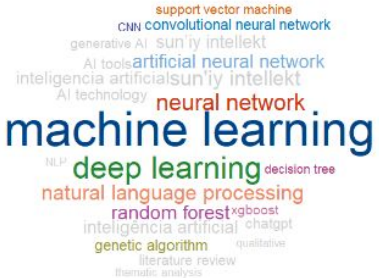
Psychology
25429 works: 61%



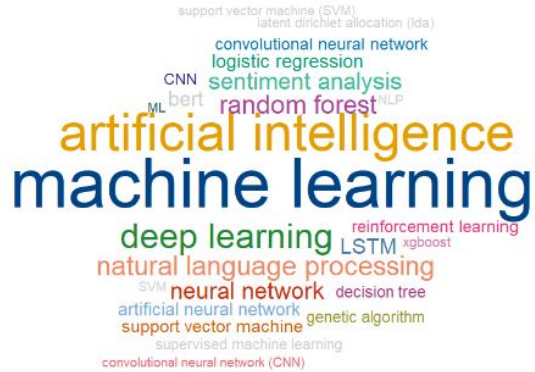
Social Sciences
22402 works: 61%



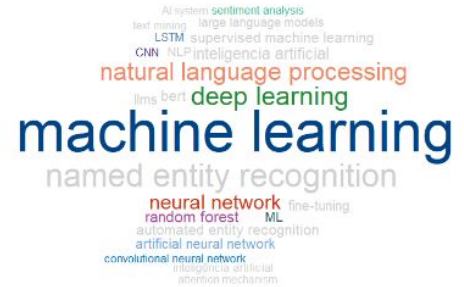
Education
5294 works: 64%



Sociology
14626 works: 63%



Political Science and International Relations
3471 works: 71%



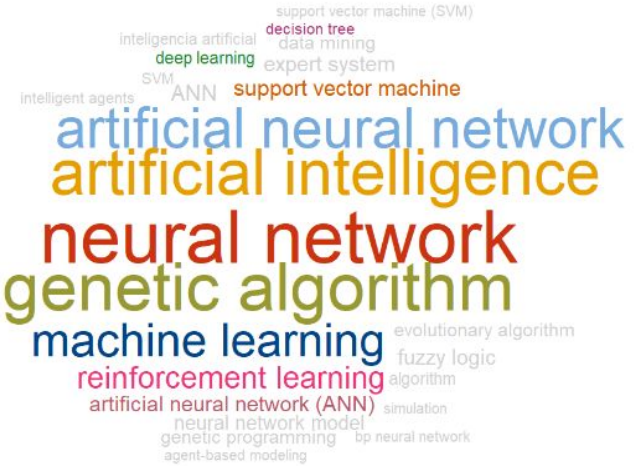
Data analysis and visualizations developed by Paloma Navarro under PI's supervision

Increased **diversity** and persistent **ambiguity**

1980-1989
244 works: 100%



1990-2009
9817 works: 67%



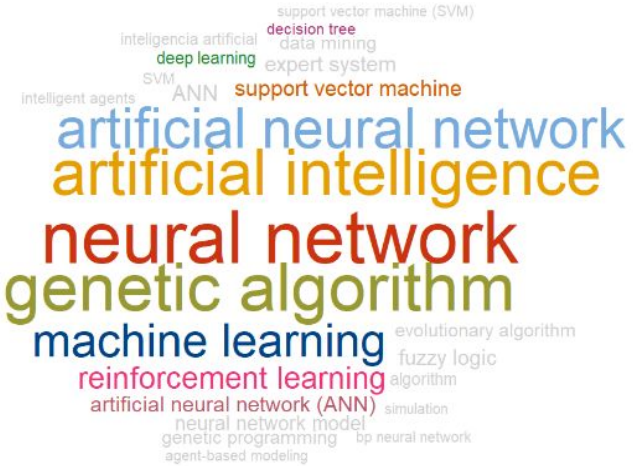
Data analysis and visualizations developed by Paloma Navarro under PI's supervision

Increased **diversity** and persistent **ambiguity**

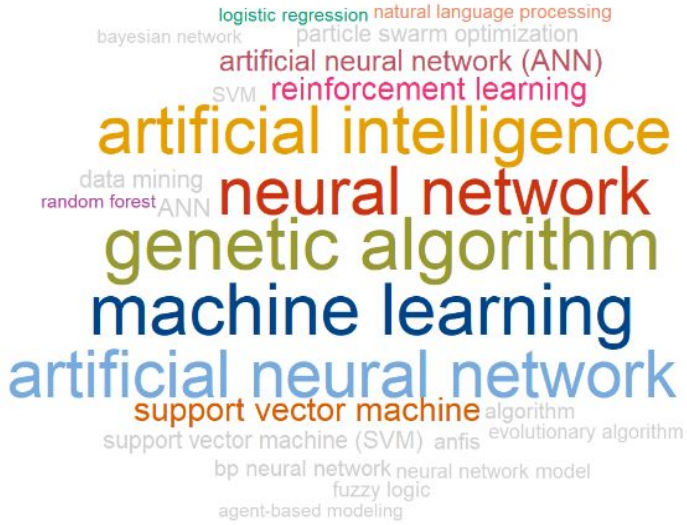
1980-1989
244 works: 100%



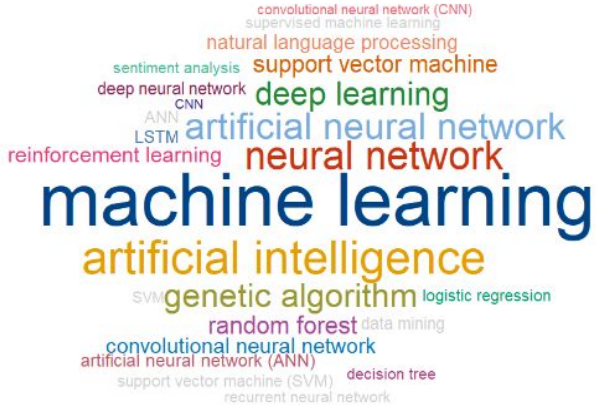
1990-2009
9817 works: 67%



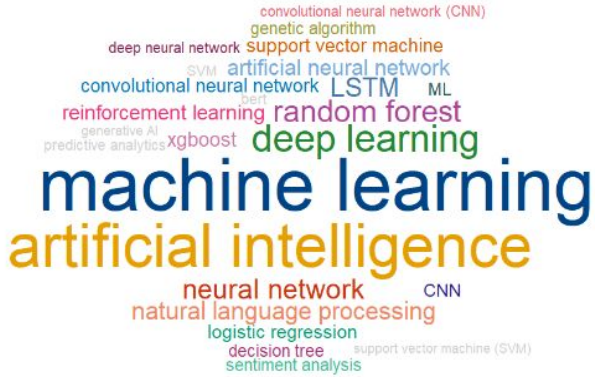
2010-2014
9261 works: 64%



2015-2019
23162 works: 63%



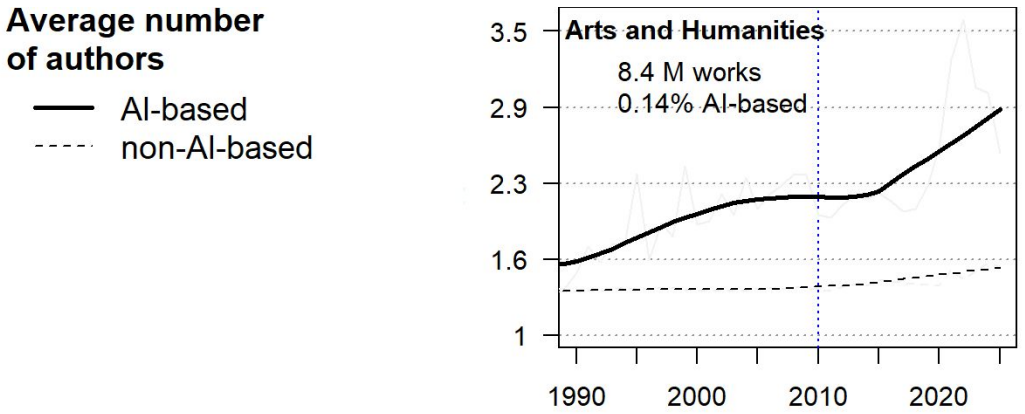
2020-2025
124387 works: 58%



Data analysis and visualizations developed by Paloma Navarro under PI's supervision

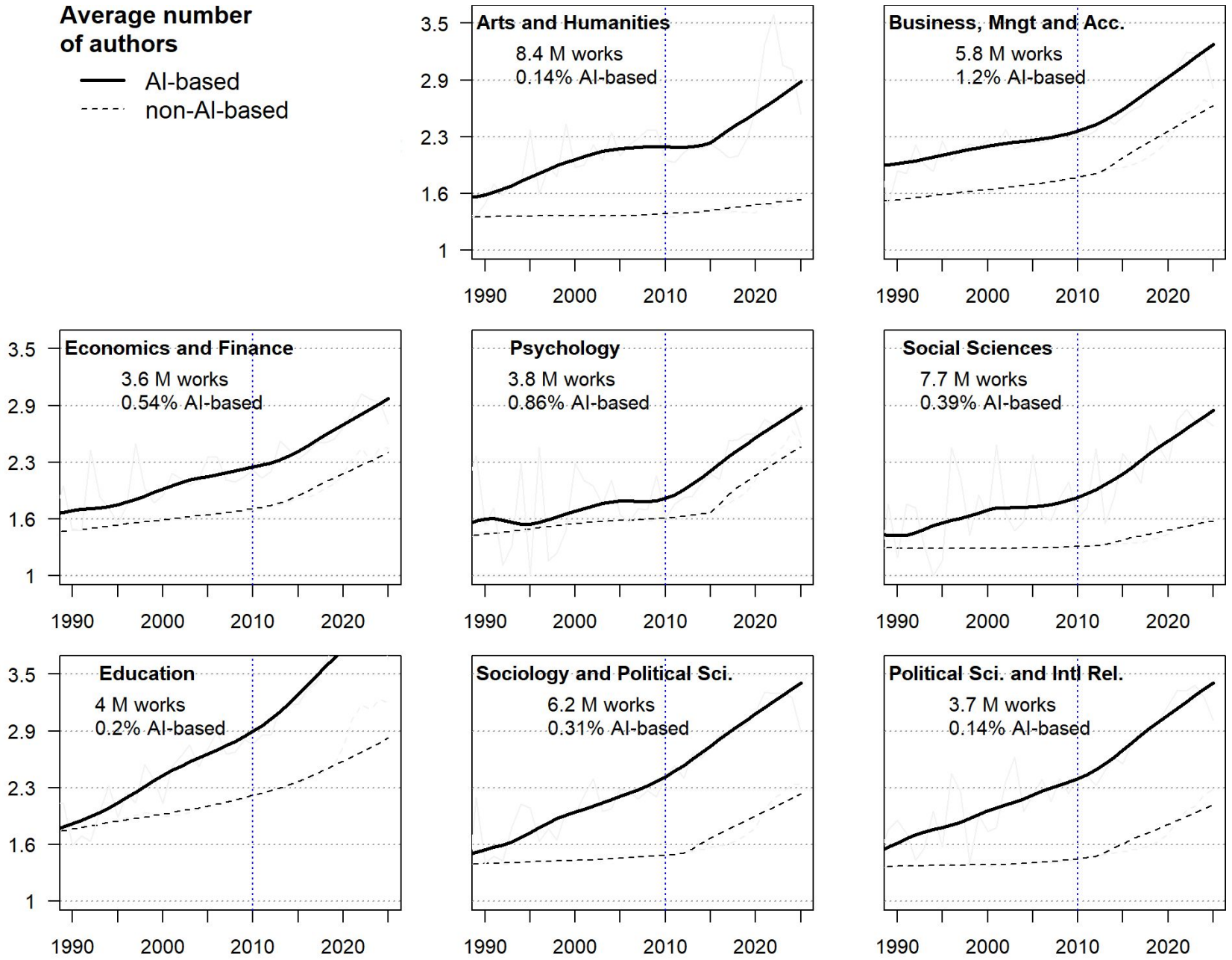
Changing publication patterns

(1) More collaborative work



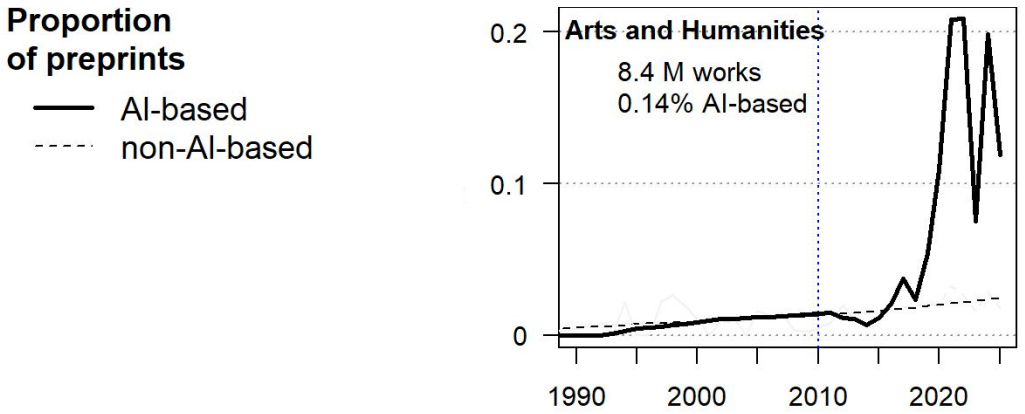
Changing publication patterns

(1) More collaborative work



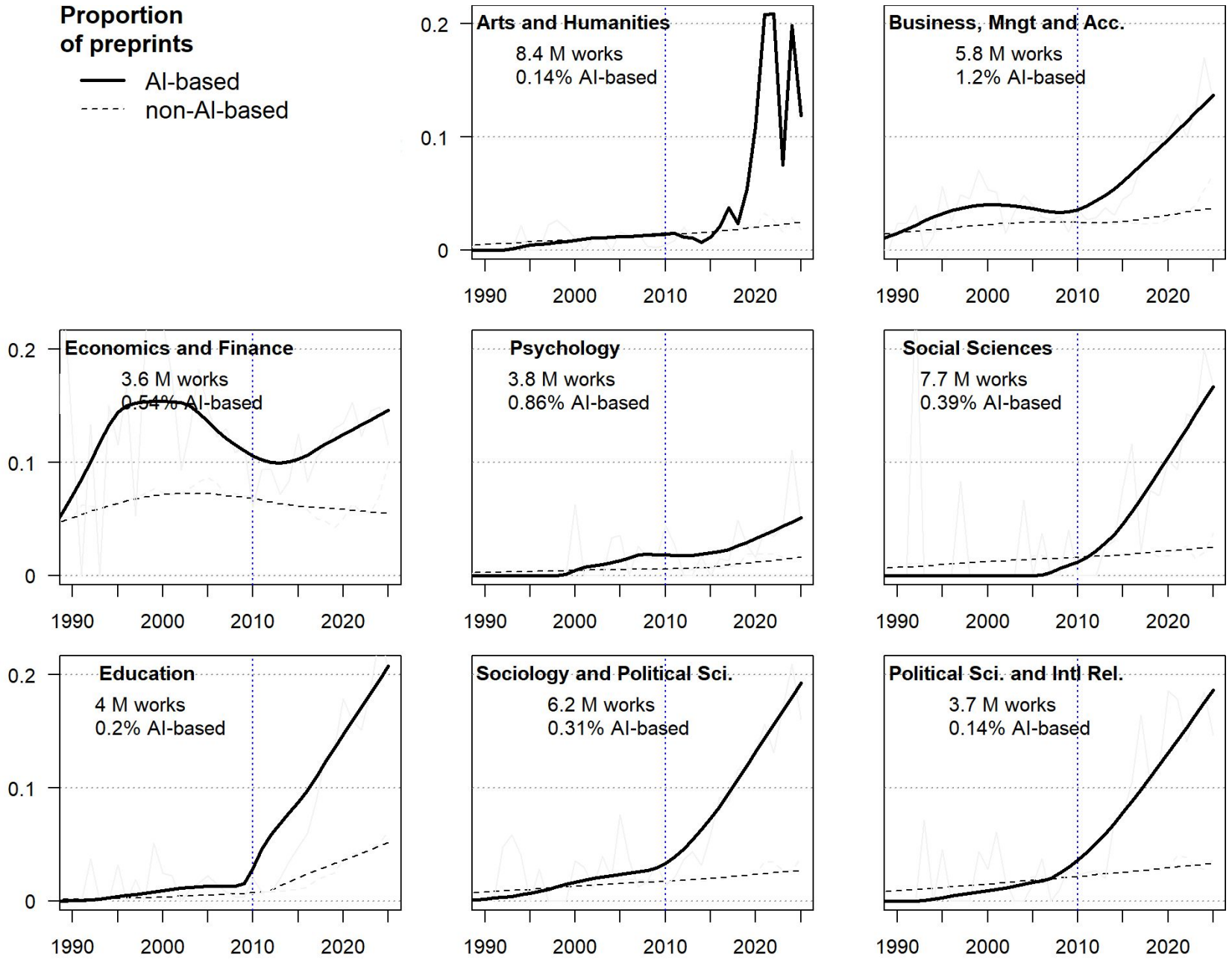
Changing publication patterns

(2) More likely to be a preprint



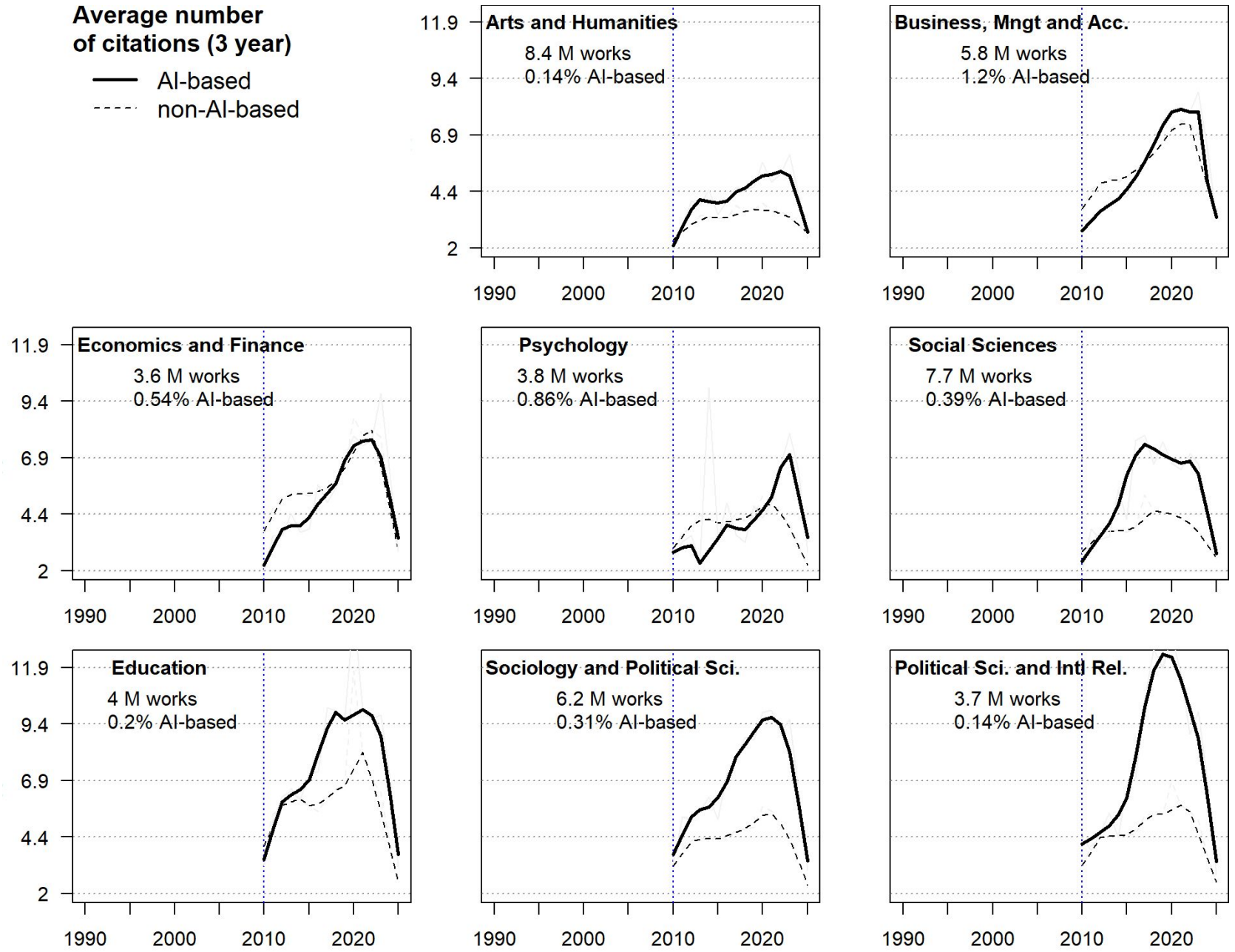
Changing publication patterns

(2) More likely to be a preprint



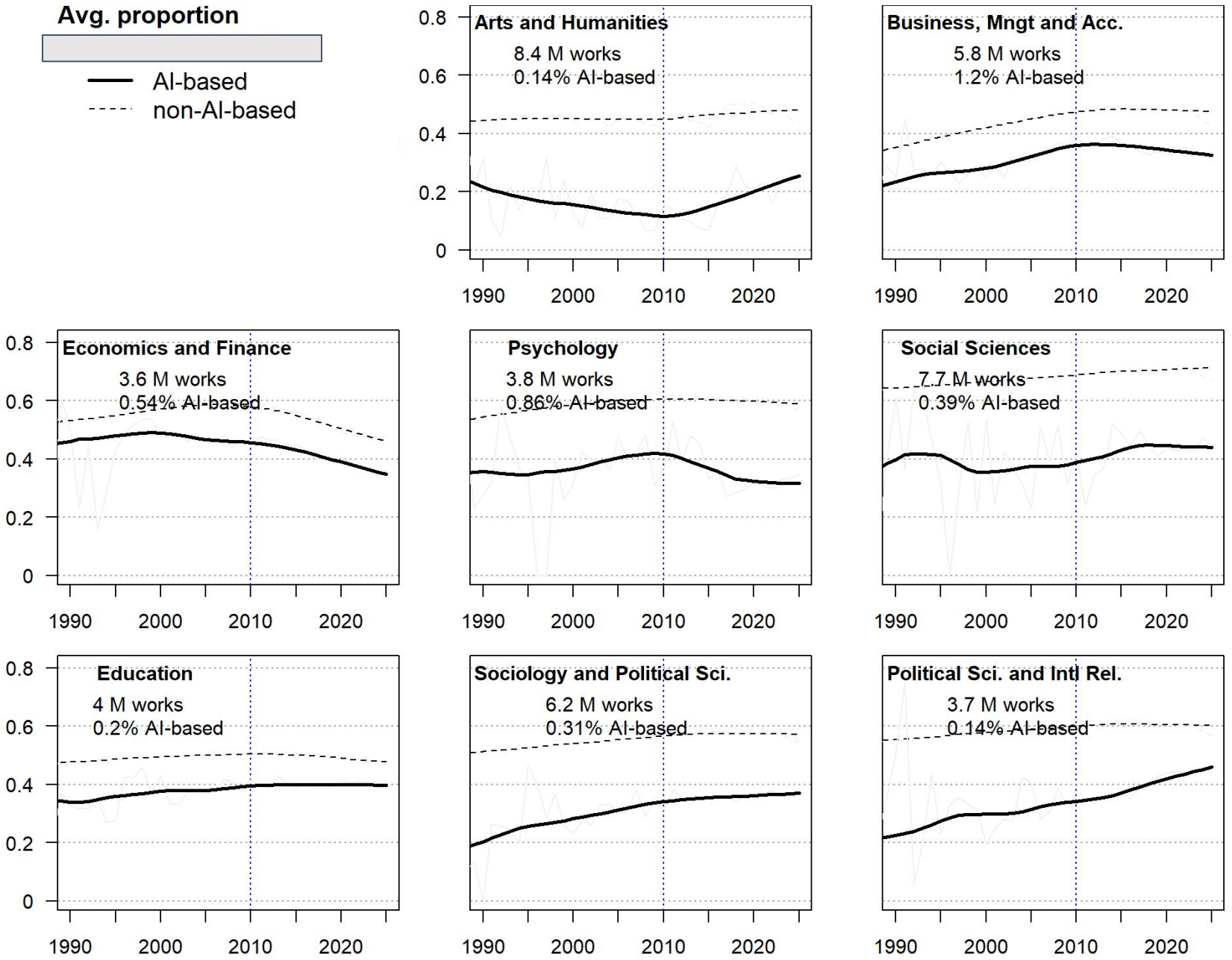
Changing publication patterns

(3) More likely to attract citations



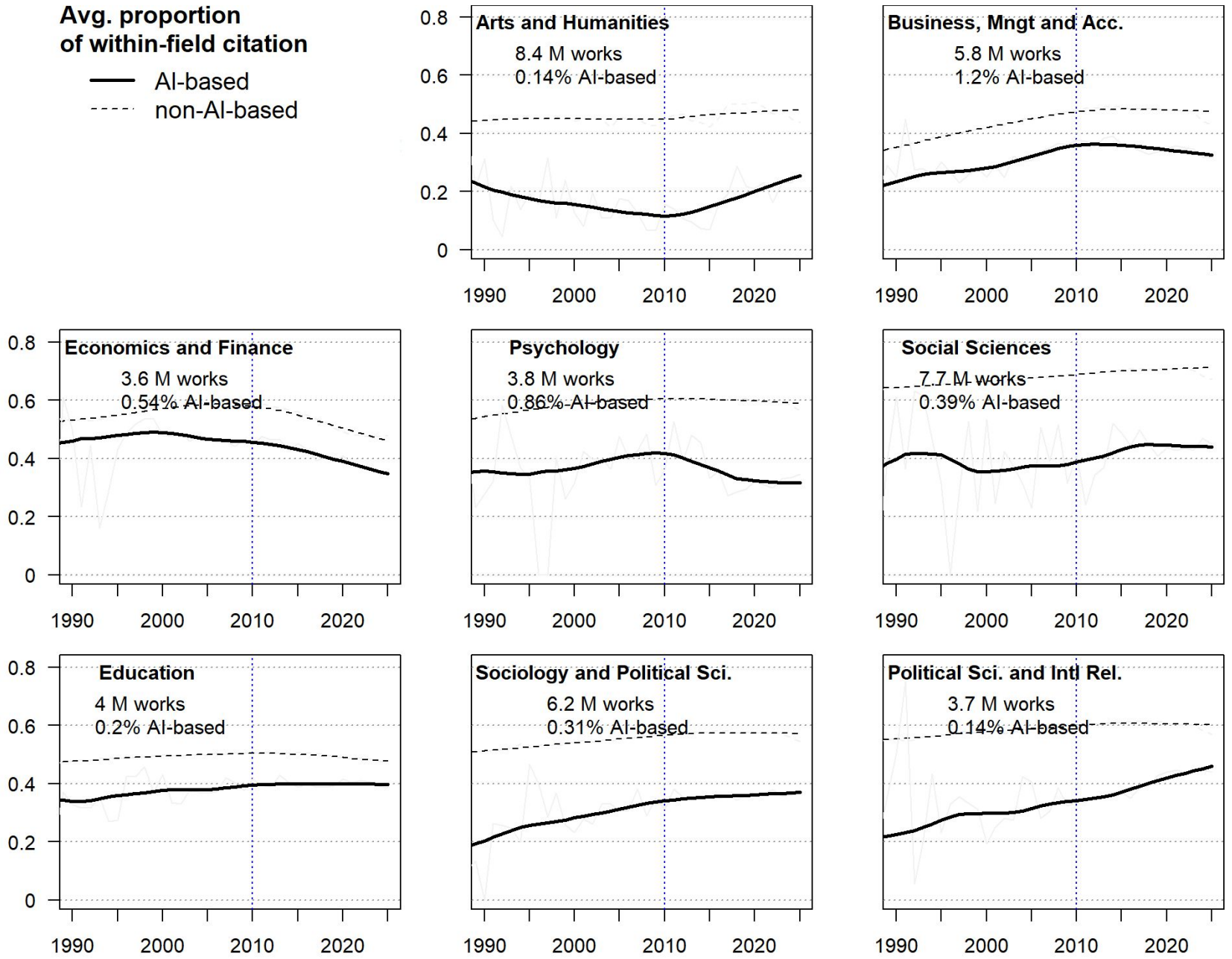
Changing citation patterns

(4)



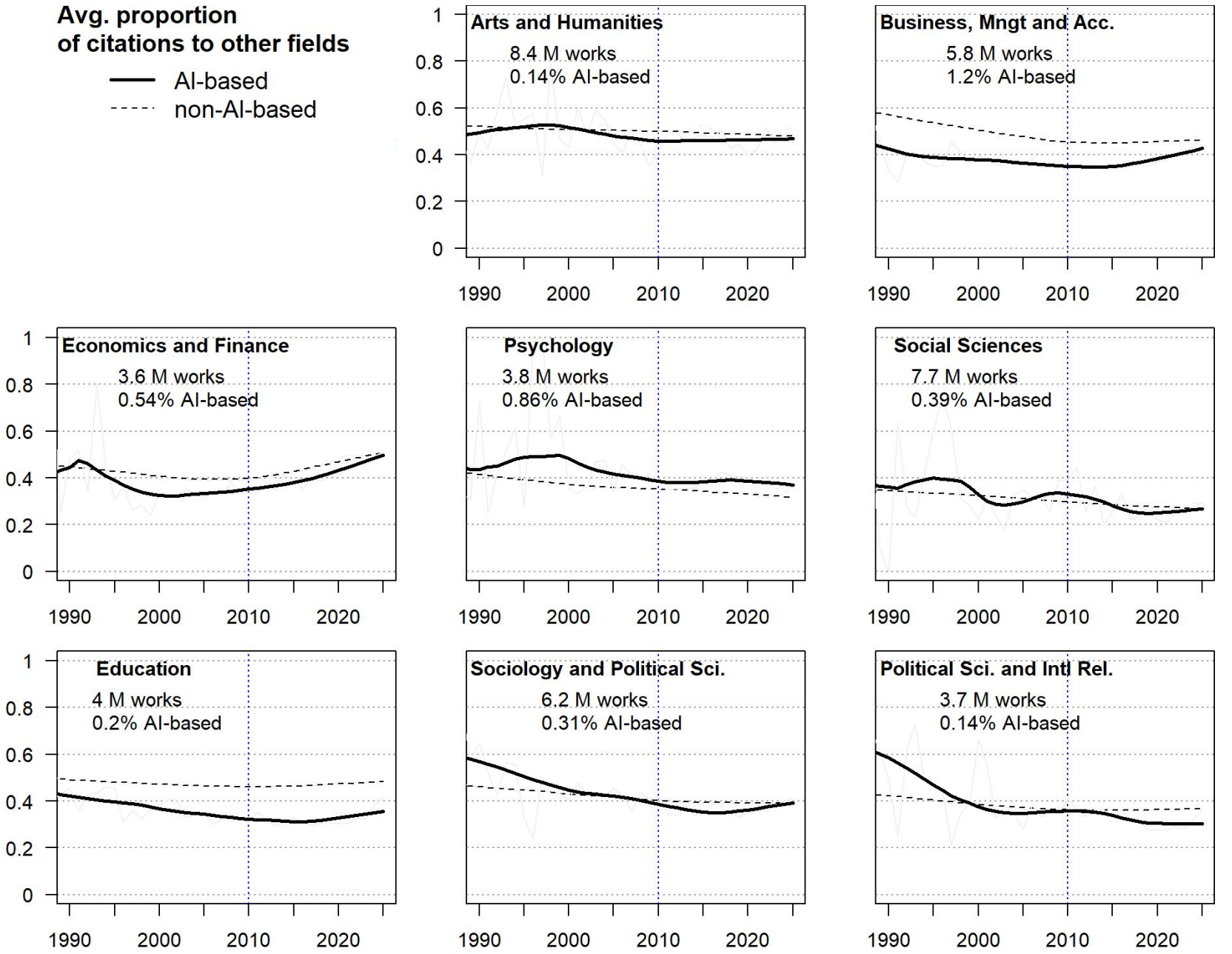
Changing citation patterns

(4) Less likely to cite works within the field



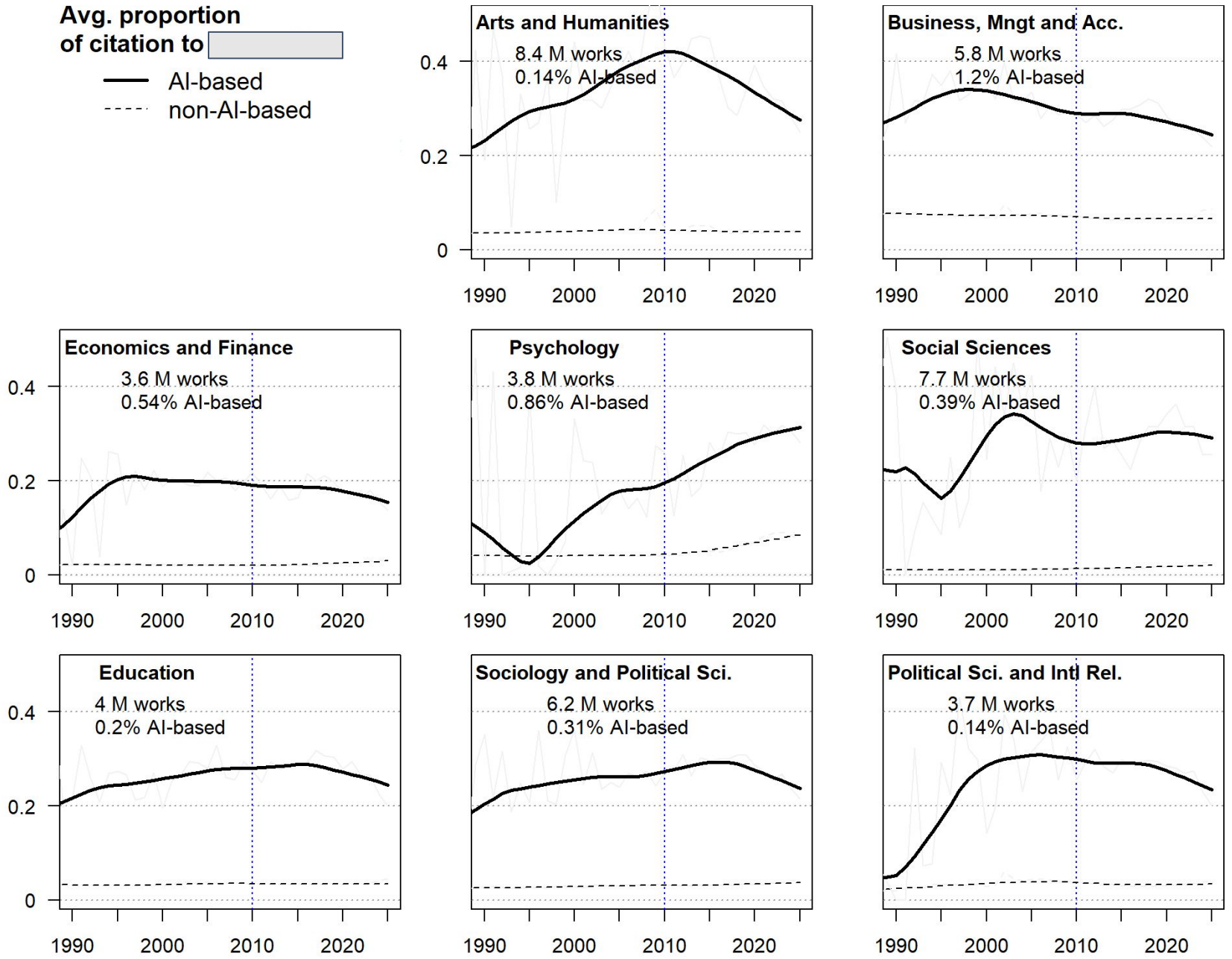
Changing citation patterns

(5) Equally likely to cite works from other fields…. except?



Changing citation patterns

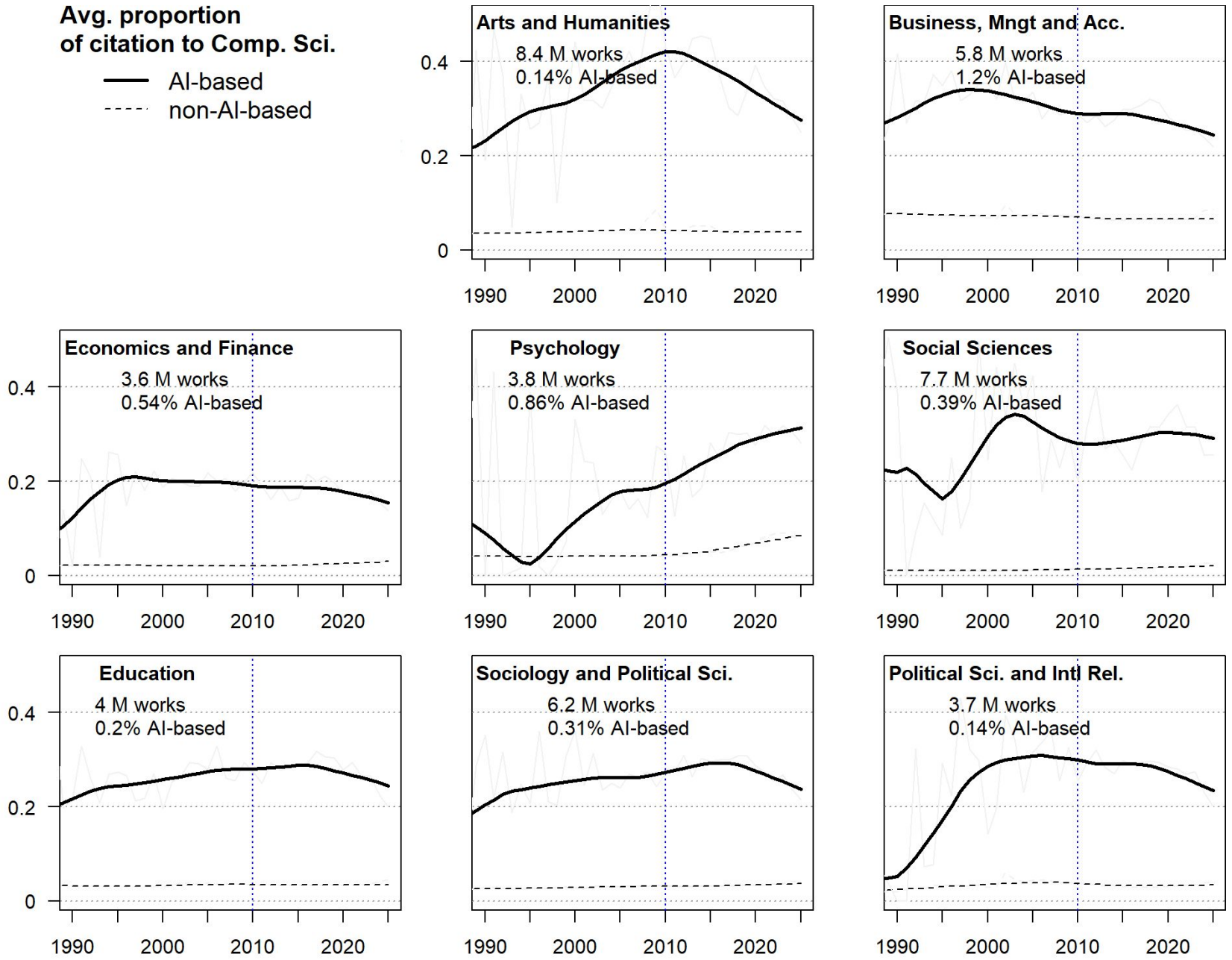
(3) …except



Changing publication patterns

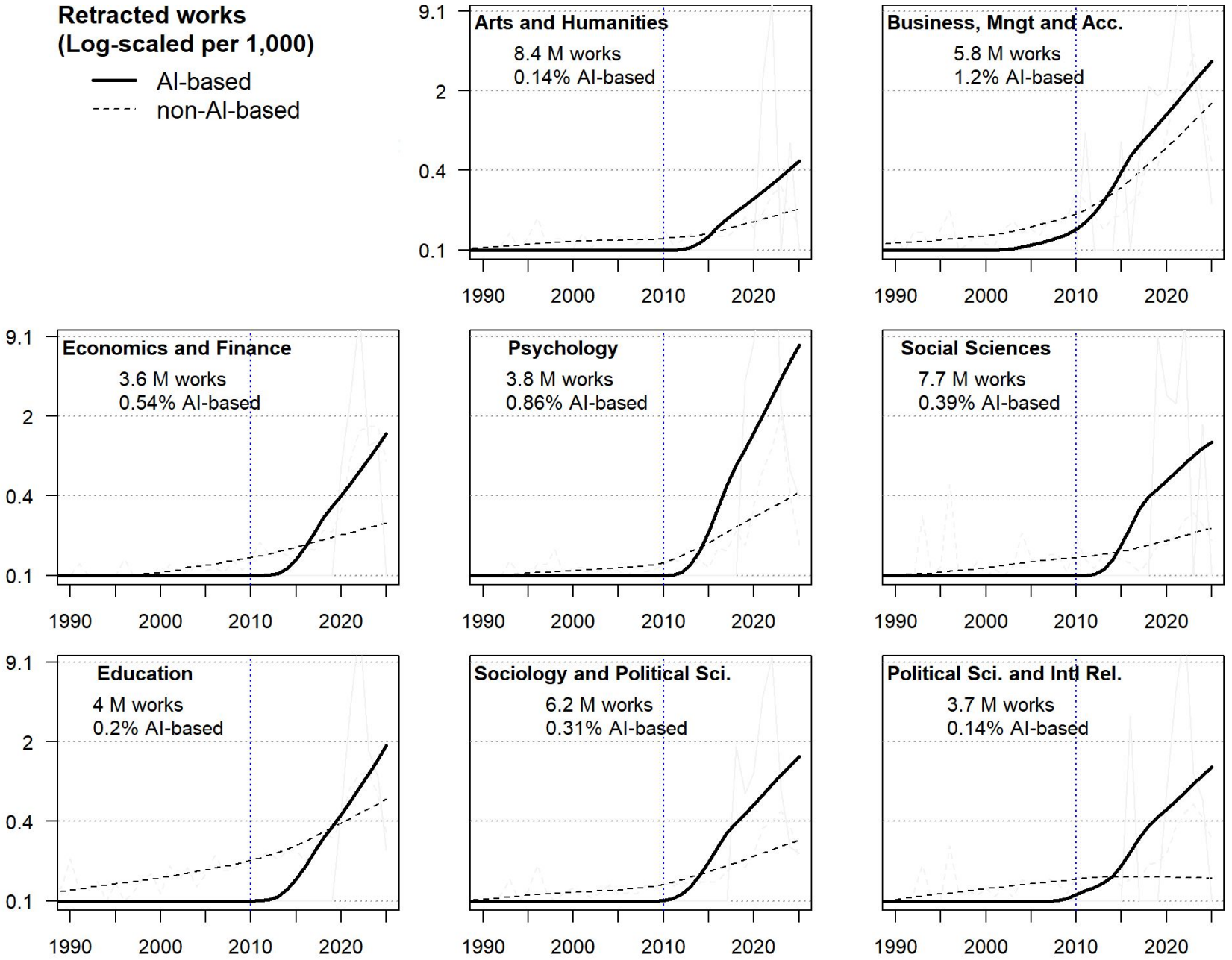
(3) …except

Computer Science!



Changing publication patterns

(6) why are AI-based papers more likely to be retracted?



Conclusions

1. Rapid adoption of AI methods pushed by research and developments from Computer Science

Conclusions

1. Rapid adoption of AI methods pushed by research and developments from Computer Science
2. Distinctive and puzzling patterns of AI adoption make unclear its general transformational potential:
    i. Strong topical concentration
    ii. Reliance on classic statistical methods
    iii. Diminished relevance of within-field knowledge
    iv. No clear signs of interdisciplinarity
    v. Larger retraction rates

Gracias por su atención.

Contact: andres.castro@bsc.es

Esta investigación ha sido financiada través de la ayuda **RyC RYC2023-042730-I** financiada por

MINISTERIO
DE CIENCIA, INNOVACIÓN
Y UNIVERSIDADES

Cofinanciado por
la Unión Europea

AGENCIA
ESTATAL DE
INVESTIGACIÓN

Amb el suport de la Generalitat de Catalunya

Generalitat de Catalunya