Quantifying the impact of AI on productivity and labor demand: evidence from U.S. Census microdata

Dean Alderucci, Sagar Baviskar, Lee Branstetter, Nathan Goldschlag, Eduard Hovy, Andrew Runge, Prasana Table, and Nick Zolas

Carnegie Mellon University
Language Technologies Institute

Carnegie Mellon University
Heinz College

United States
Census
Bureau

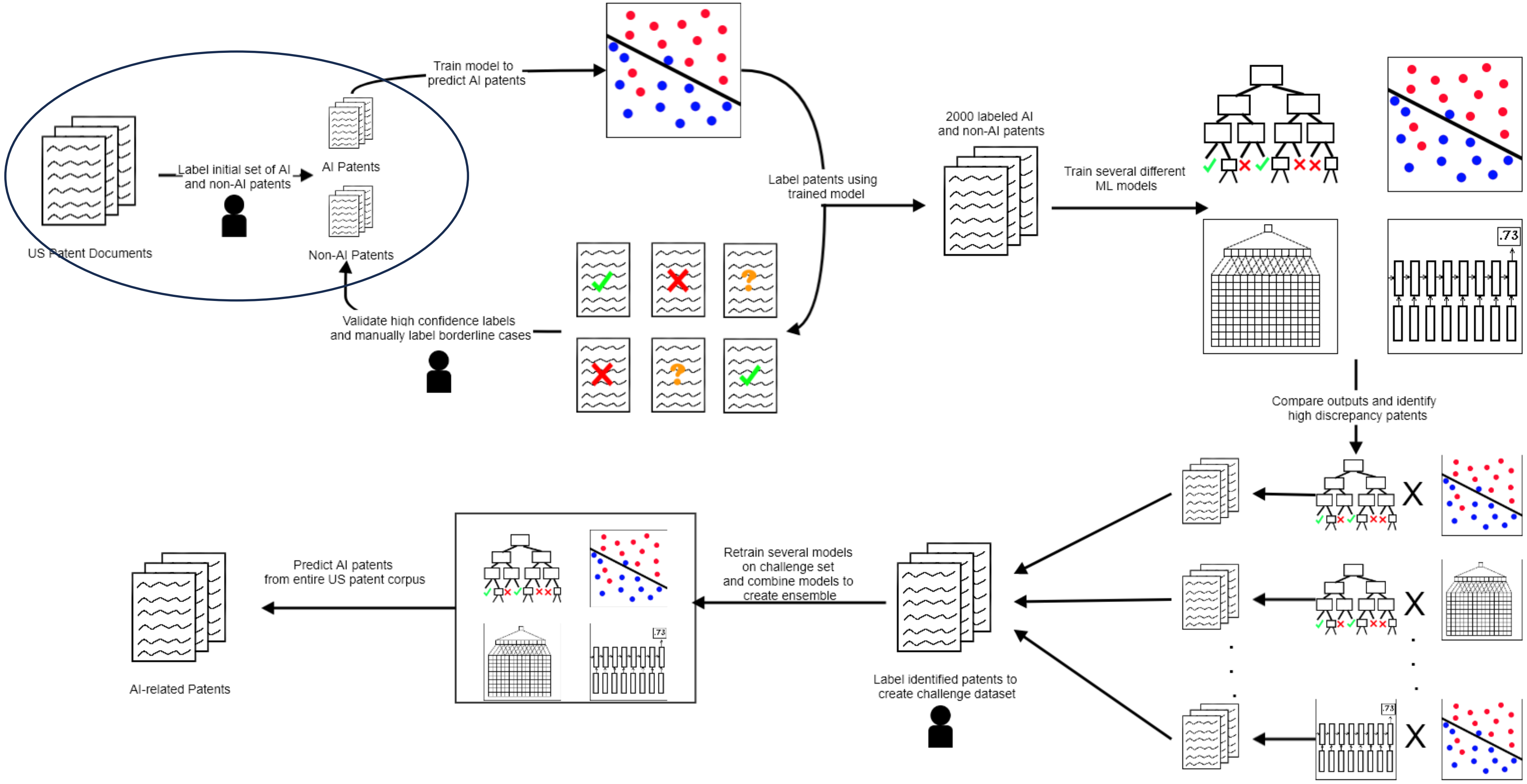
Any opinions and conclusions expressed herein are those of the author(s) and do not necessarily represent the views of the U.S. Census Bureau or the National Bureau of Economic Research. All results have been reviewed to ensure that no confidential information is disclosed. The DRB codes for this project are: DRB-B0027-CED-20190205, CBDRB-FY19-414, and CBDRB-FY20-105.

This paper provides new measures of AI-related innovation that complement existing studies

- There have been impressive efforts efforts to directly measure AI adoption (McElheran et al., 2024; Zolas et al, 2019)…
- And efforts to measure the impact of the hiring of AI workers as reflected in online job ad data (Babina et al., 2022, 2024)…
- And more recent experimental studies of the impact of particular kinds of AI in particular work contexts (Brynjolfsson et al., 2023; Noy and Zhang, 2023).
- We build on and extend the efforts of Cockburn et al. (2019), Webb et al. (2019), and Giczy et al. (2021) to identify AI-related patents…
- And we examine the impact of these AI inventions on the AI-inventing firms, using Census micro-data, finding evidence that suggests significant effects on productivity
- We are also beginning to measure the impact of elite academic scientists on AI-related invention, through collaboration and the employment of their doctoral students.

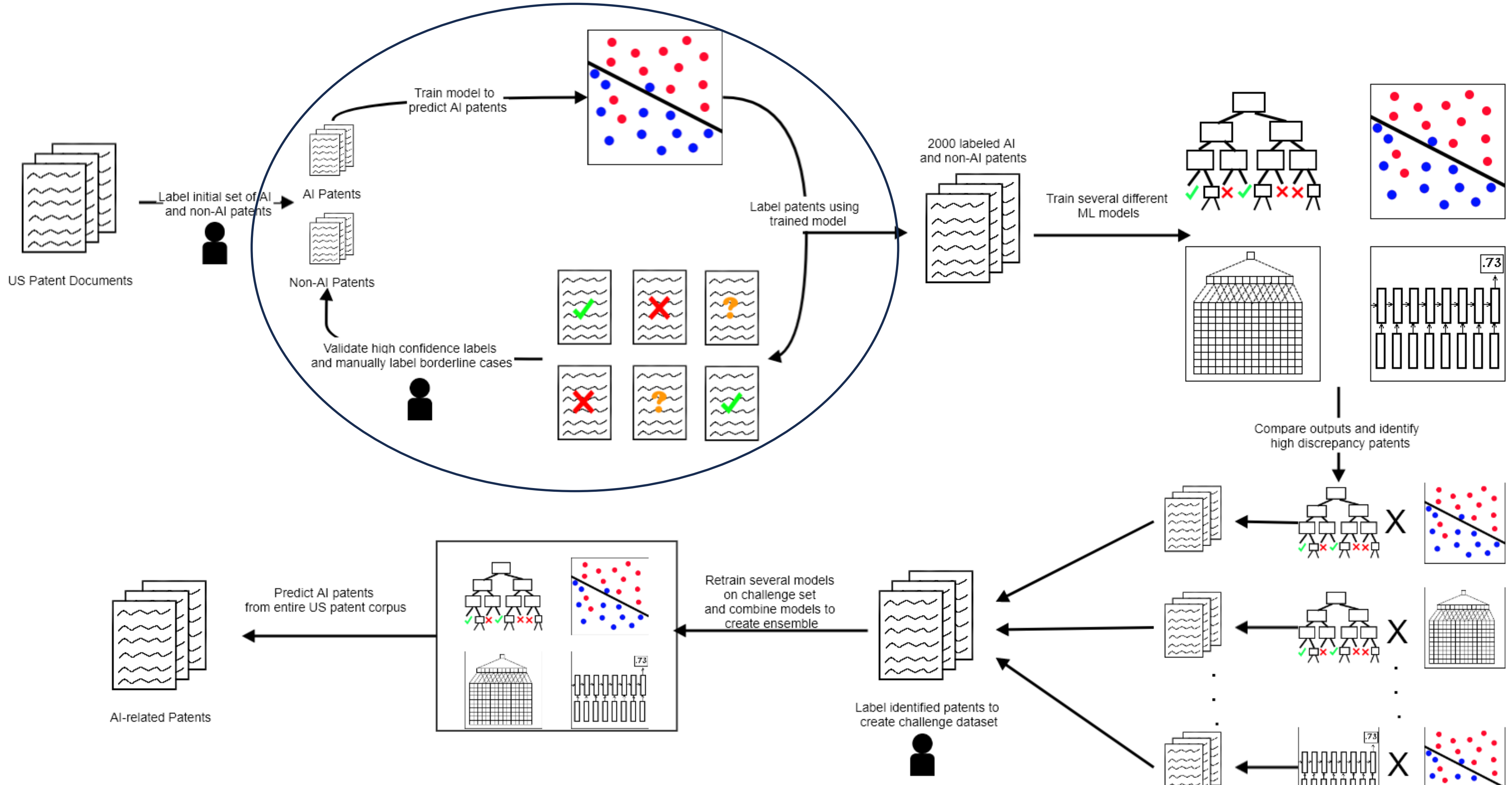
Using AI to find AI inventions (1)

Begin with "hand-curated" AI (and non-AI) patents



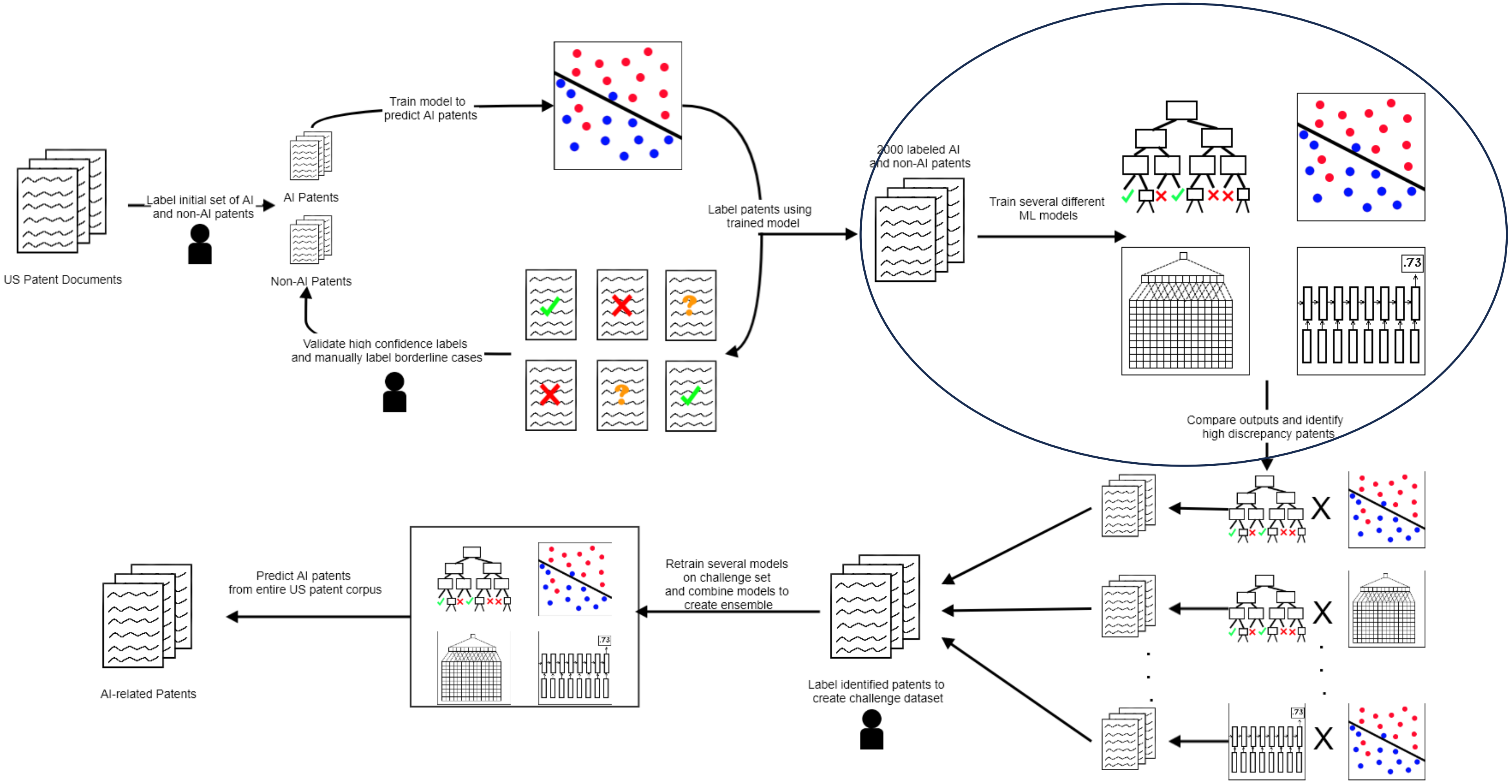
Using AI to find AI inventions (2)

Train a SVM to identify more AI (and non-AI) patents (with human input)



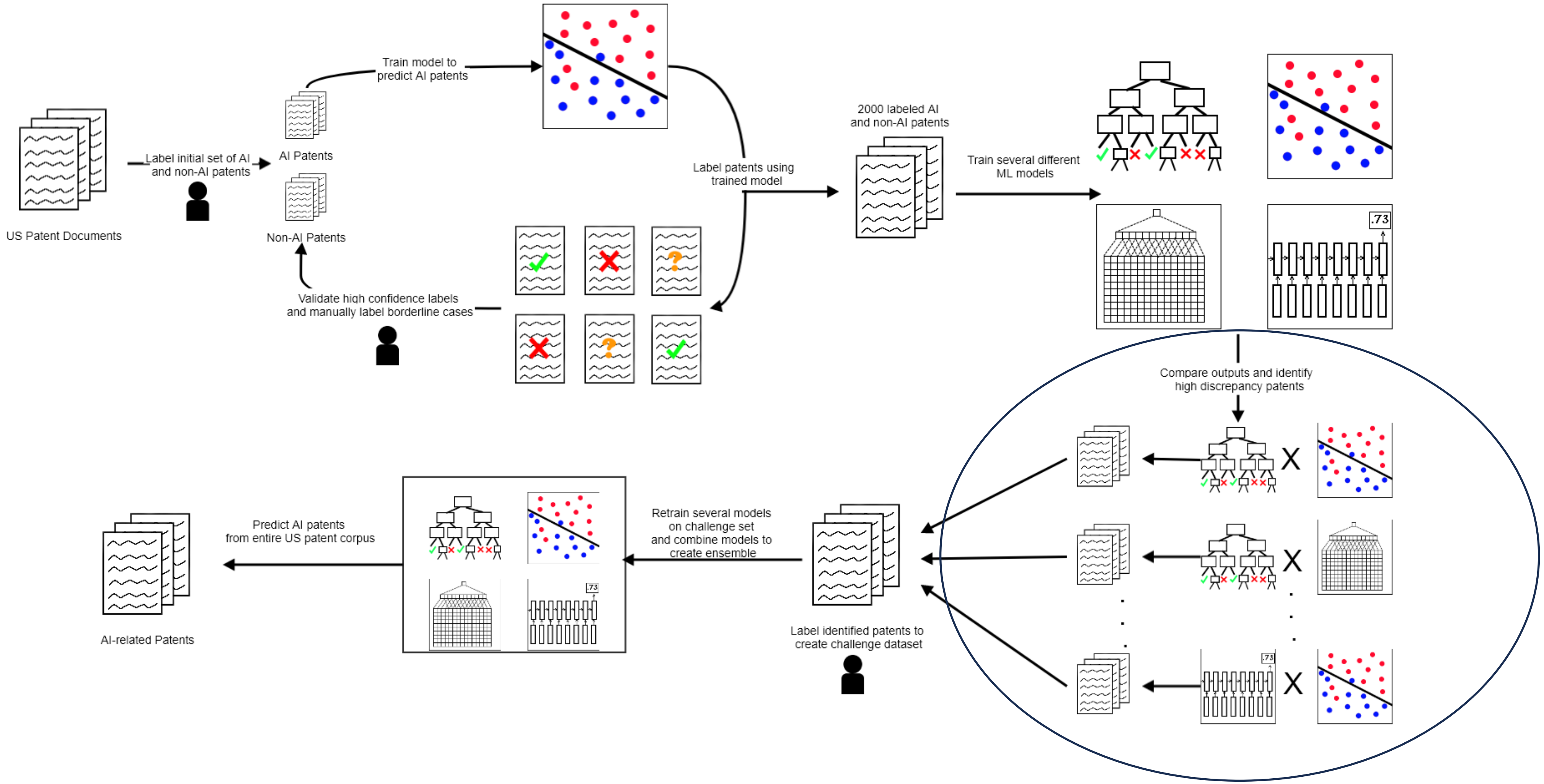
Using AI to find AI inventions (3)

Train several different ML models on the expanded training data set



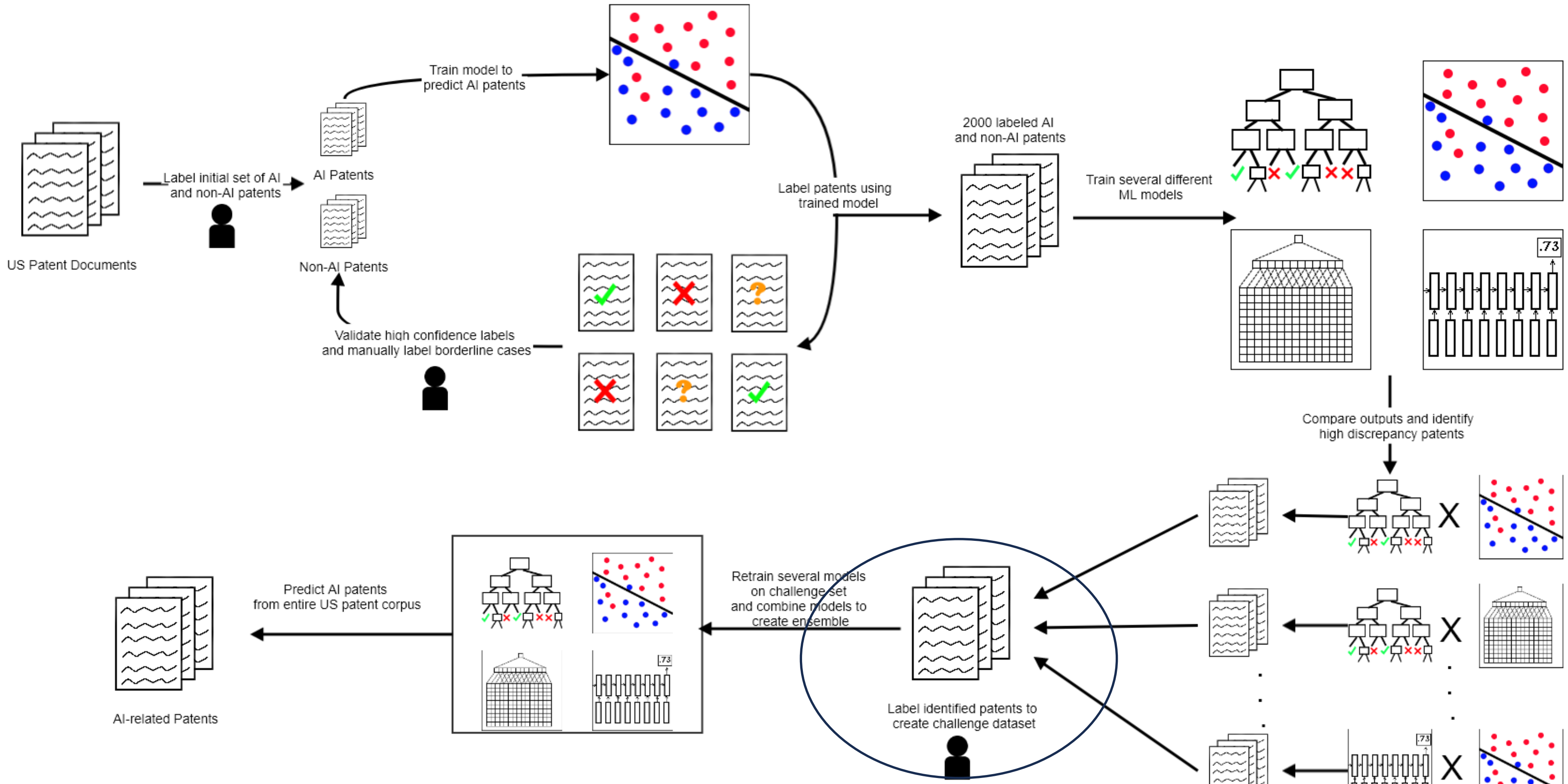
Using AI to find AI inventions (4)

Compare outputs and identify high-discrepancy patents



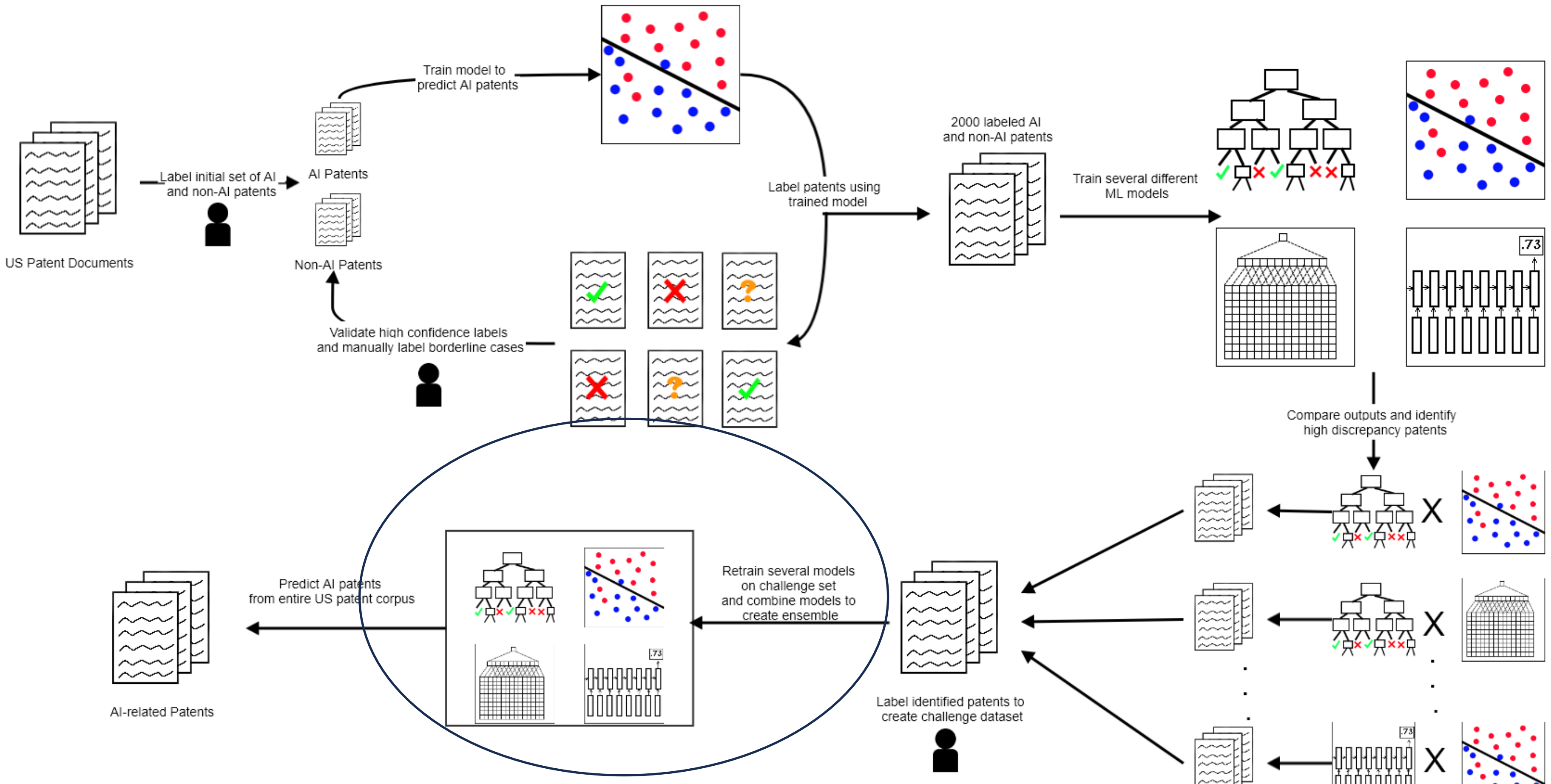
Using AI to find AI inventions (5)

Create a "challenge" data set



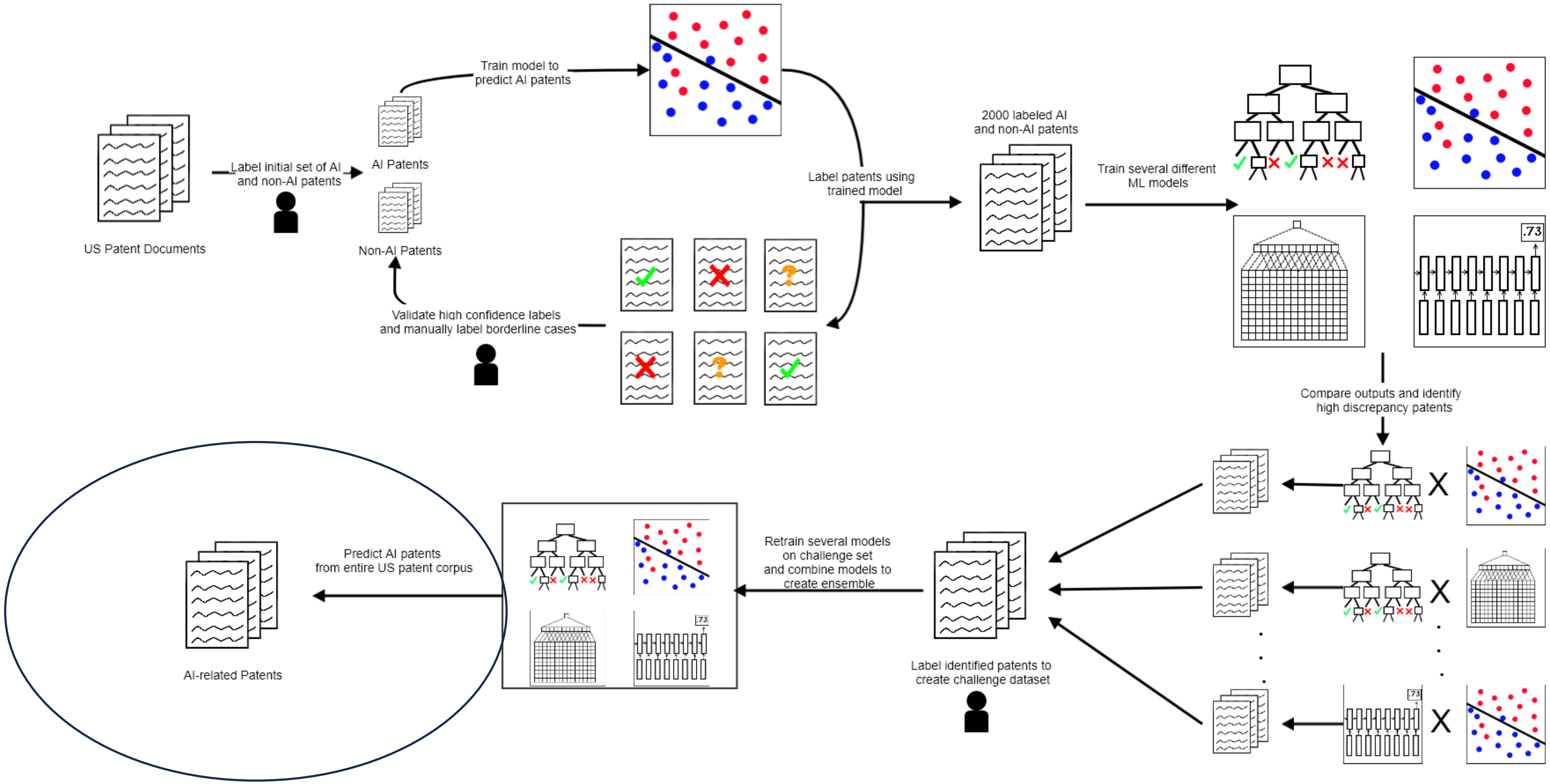
Using AI to find AI inventions (6)

Retrain the models on the challenge data set, combine models to create an "ensemble"



Using AI to find AI inventions (7)

Predict the "AI-ness" of every U.S. patent, 1990-2018



Our methods find far more AI patents than previous approaches taken by some economists

- Cockburn et al. (2019) take a “standard approach,” focusing on a relatively small set of key words and patent classes.
- This approach identifies fewer than 14,000 patents between 1990 and 2014, and it includes large numbers of ”robotics hardware” patents.
- Webb et al. (2019) take a similar, more focused approach, identifying 2,000+ patents related to ”machine learning” and 4,000+ related to “neural networks.”
- Our approach identifies 52,896 patents that are AI related with 95% confidence and 146,952 patents that are AI related with 70% confidence (through 2018).
- We identify most of the AI patents tagged by other economists as ”AI patents” but also capture a very large number that many earlier approaches omit.
- However, our methods find far *fewer* patents than do recent efforts by the USPTO to apply machine learning methods to patent data (Giczy et al., 2021).

AI patenting has grown rapidly over our sample period

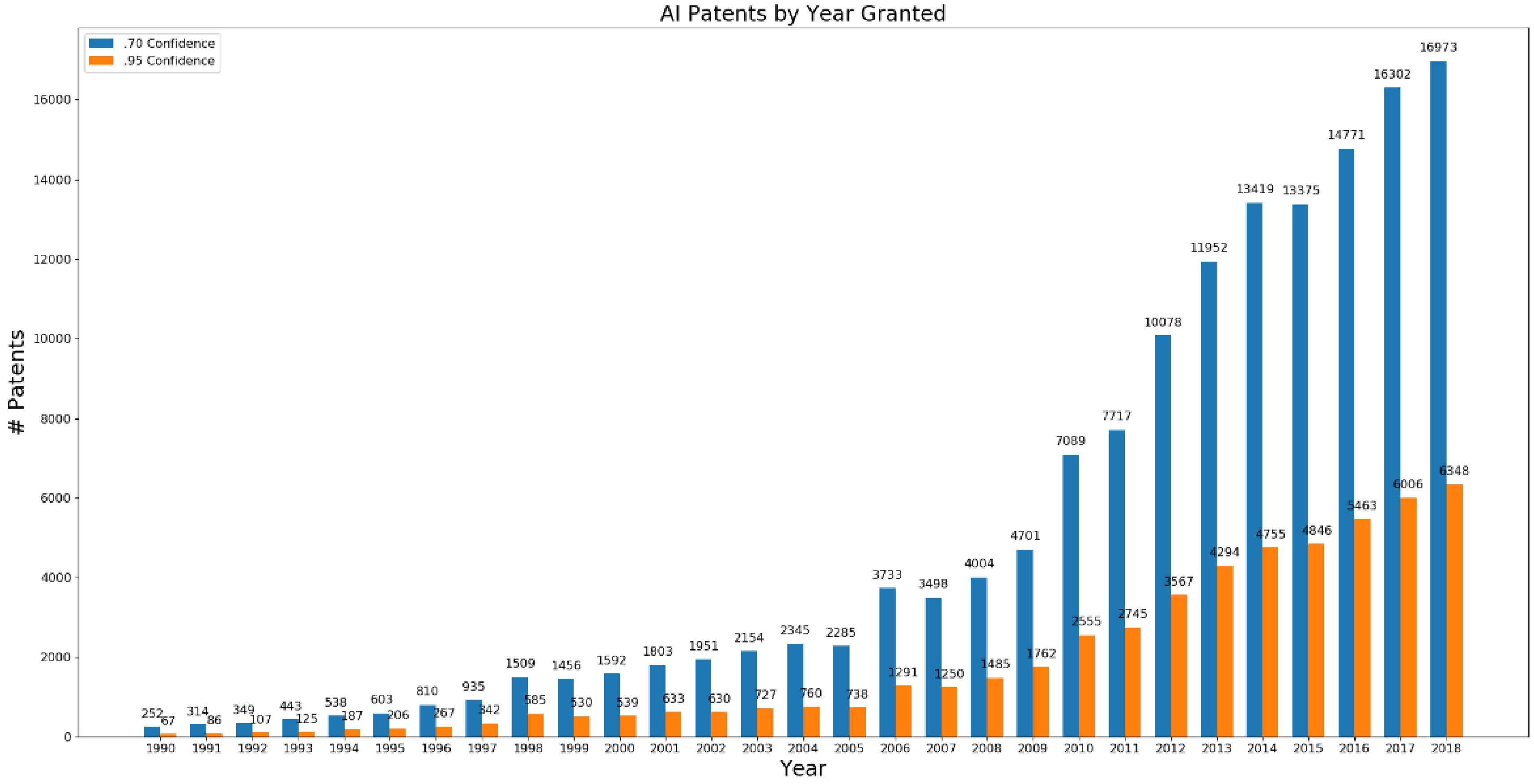


Figure 2. AI Patents by Grant Year

AI patenting is widely distributed across patent classes…

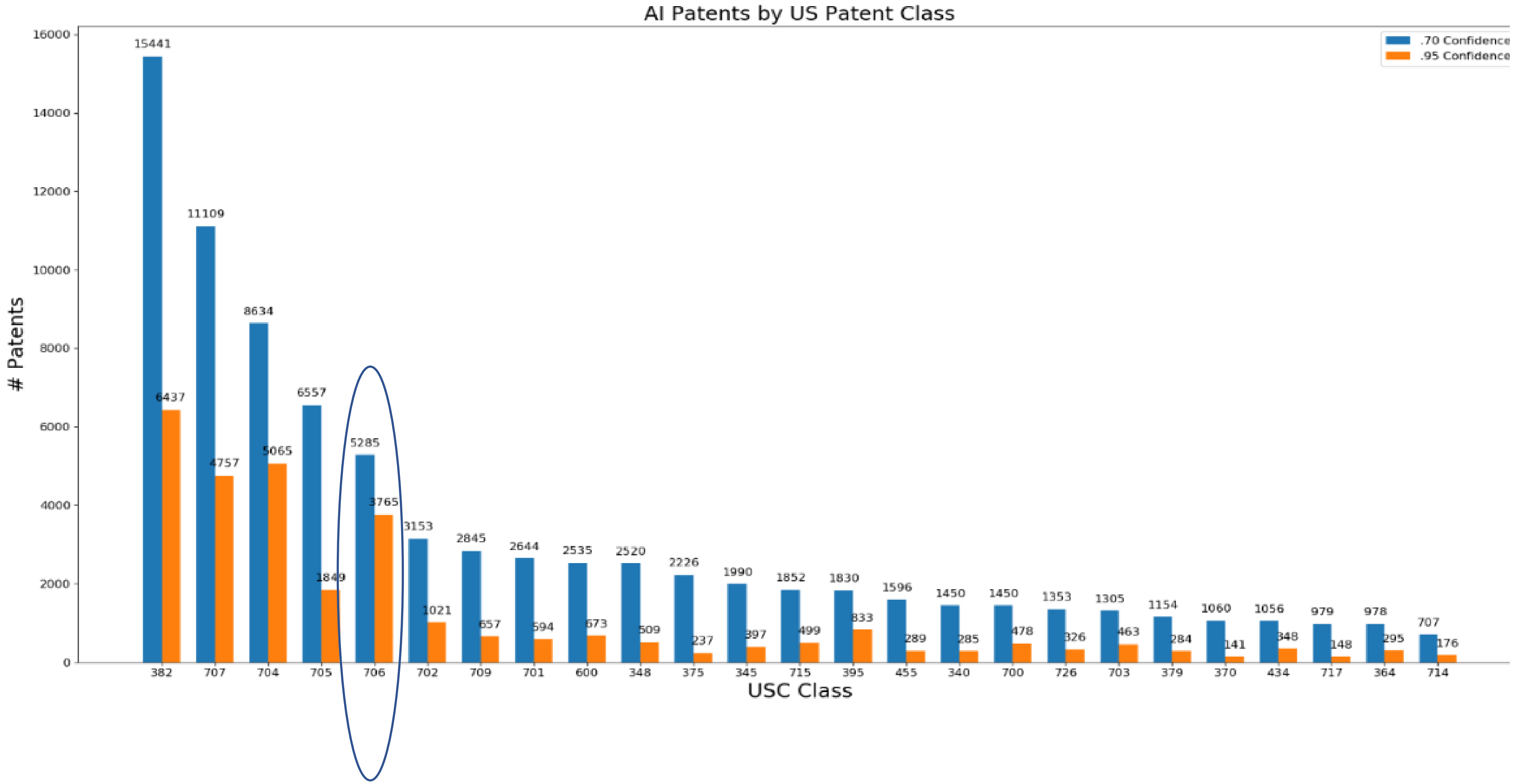
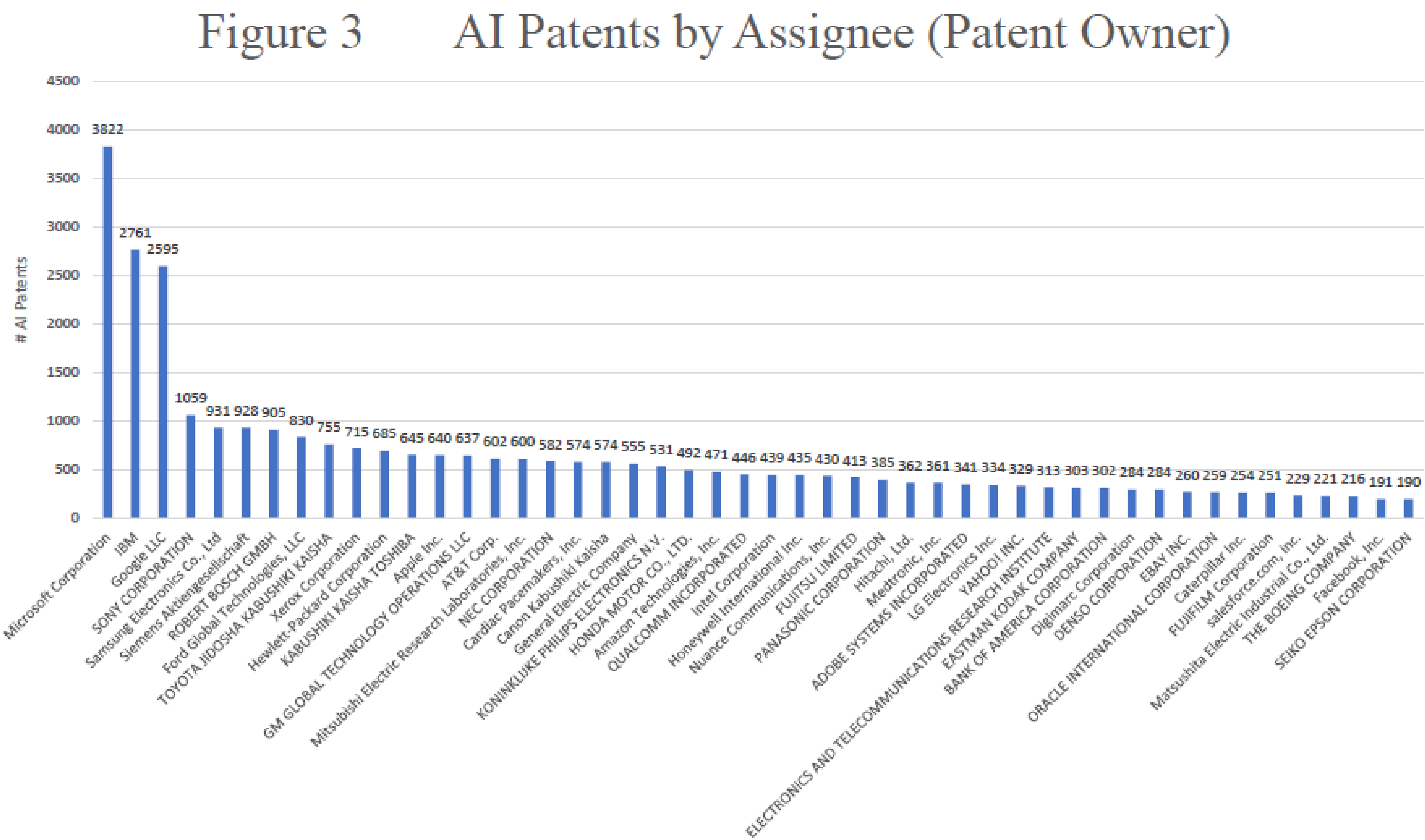


Figure 5. AI Patents by USPC Class

And across firms….



AI patenting is concentrated in a few metro areas within the United States…

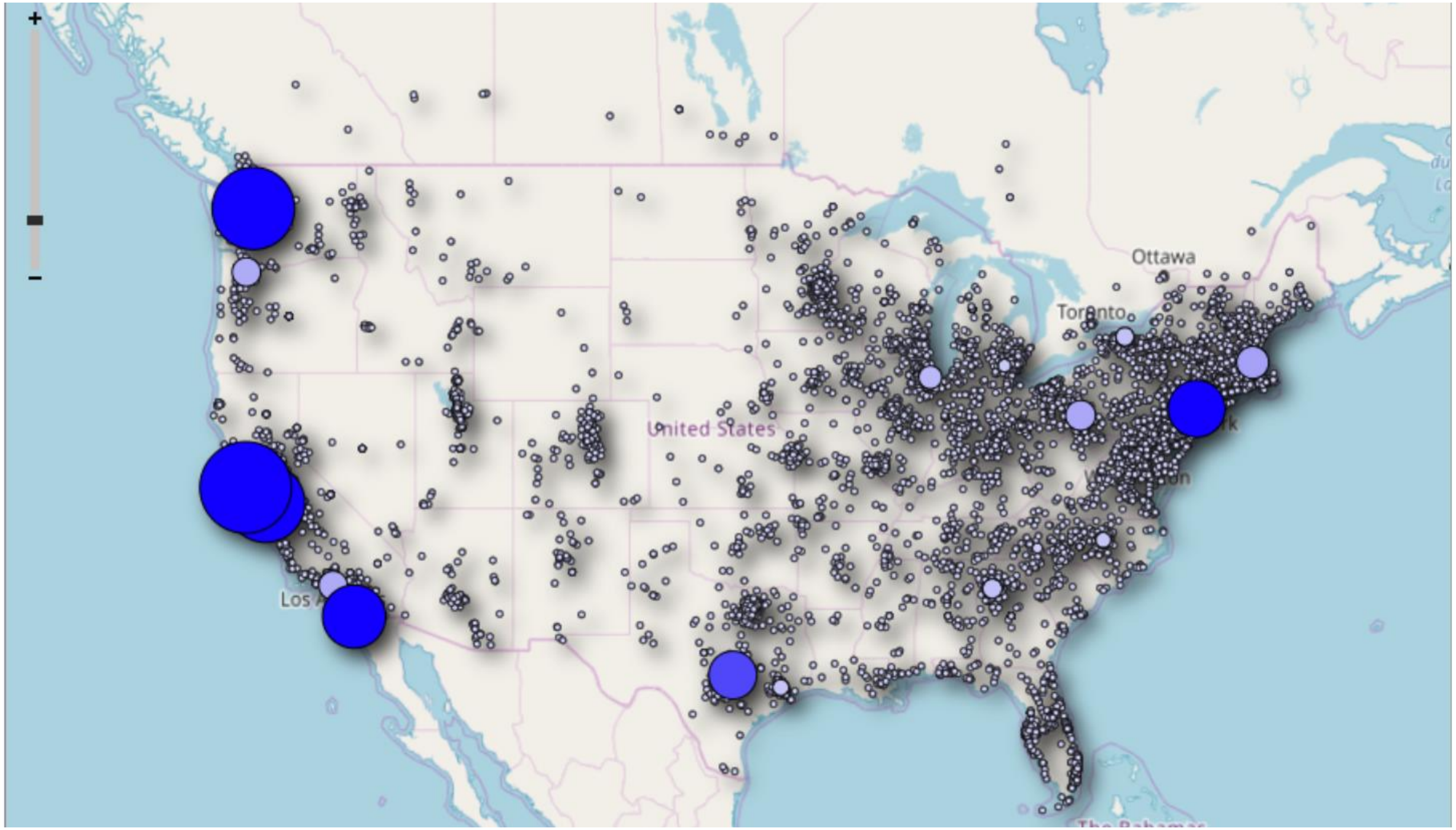


Figure 4 Inventor Heat Map of AI Patents in U.S

Methodology for impact assessment

- Use existing USPTO-Census patent-to-firm crosswalk (Graham et al. 2018), augmented by our own work to link the AI patents with a Census FirmID
  - Incorporate various outcome measures including:
    - Employment, Revenue and Revenue per Employee (Revenue-enhanced LBD)
    - Value-added, Total Factor Productivity, Production Worker Share (ASM and CMF)
    - 90-10, 90-50 and 50-10 earnings ratio by firm (LEHD)

- Measure within-firm changes from AI innovations at the *extensive* and *intensive* margins.

- Construct a “control” group of firms to which AI-inventing firms can be compared and perform an event-study analysis
  - Use propensity score/exact matching to construct a control group using size, age, industry (4-digit NAICS), payroll, and patenting activity as predictors
  - Event study centered around the timing of the first AI patent filed by the AI-inventing firm.

AI Invention and firm productivity

Table 5: Impact of AI Innovations on Firm Productivity, 1997-2018 (manufacturing only)

	(1)	(2)	(3)	(4)	(5)	(6)
	Ln Total Value of Shipments per Employee		Ln Value Added Per Employee		Ln Total Factor Productivity (TFP)	
AI Treatment (1/0)	0.272*** (0.0322)		0.225*** (0.0337)		0.0830*** (0.0225)	
IHS AI Patents		0.148*** (0.0161)		0.104*** (0.0166)		0.0565*** (0.0119)
Ln Capital Stock	0.310*** (0.00195)	0.310*** (0.00195)	0.301*** (0.00203)	0.301*** (0.00203)	-0.0633*** (0.000888)	-0.0633*** (0.000888)
Age Bins	Yes	Yes	Yes	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,124,000	1,124,000	1,124,000	1,124,000	1,124,000	1,124,000
R-squared	0.674	0.674	0.921	0.921	0.706	0.706

Robust Standard Errors clustered at the 4-digit NAICS industry level. *, ** and *** denotes significance at the 5%, 1% and 0.1% respectively. Each regression includes a constant, age bins, multi-unit and multi-national indicator controls, which are not displayed here. Note that our Multi-Unit regressor drops out from the within-firm specification as the firm-identifier for multi-unit status does not change. Across firm effects are listed in the appendix.

Event study on employment and labor productivity (all firms, not just manufacturing)

Table 10: Impact of AI Innovations on Employment and Revenue, 1997-2018 (matched only)

	(1)	(2)	(3)	(4)
	Ln Employment		Ln Revenue per Employee	
AI Treatment (1/0)	Dropped	Dropped	Dropped	Dropped
Post AI Year	0.0206 (0.0144)		-0.124*** (0.0192)	
AI Treatment x Post AI Year	0.138*** (0.0194)		0.164*** (0.0263)	
Age Bins	Yes	Yes	Yes	Yes
Industry-Year Fixed Effects	Yes	Yes	Yes	Yes
Firm Fixed Effects	Yes	Yes	Yes	Yes
Observations	36,000	36,000	36,000	36,000
R-squared	0.975	0.976	0.787	0.787

Robust Standard Errors clustered at the 4-digit NAICS industry level. *, ** and *** denotes significance at the 5%, 1% and 0.1% respectively. Each regression includes a constant, which is not displayed here.

Does AI widen income inequality?

- Earlier generations of IT innovation dramatically expanded income inequality by raising demand for skill in the U.S. labor market.
- Will AI continue – or even worsen – this longstanding trend?
- We can examine the impact of AI on demand for production workers.
- By linking data on our AI-inventing firms to the LEHD, we can also examine the the association between AI invention and actual changes in within-firm earnings inequality.
- We measure earnings inequality using the 90th percentile – 10th percentile, 90th-50th, and 50th-10th income ratios.

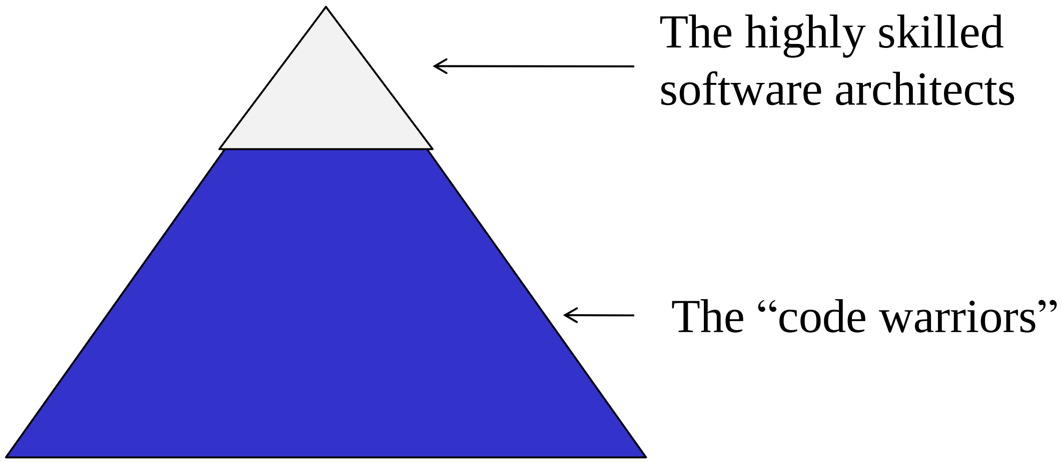
Does AI invention increase within-firm income inequality (in event study models)?

Table 12: Impact of AI Innovations on 90-10, 90-50 and 50-10 Earnings Ratio, 1997-2019 (full matched set of firms)

	(1)	(2)	(3)	(4)	(5)	(6)
	90-10 Earnings Ratio		90-50 Earnings Ratio		50-10 Earnings Ratio	
AI Firm (1/0)	Dropped	Dropped	Dropped	Dropped	Dropped	Dropped
Post AI (1/0)	-0.0170		-0.00563		-0.0101	
	(0.0159)		(0.00825)		(0.0117)	
AI x Post (1/0)	-0.00464		-0.0135		0.00920	
	(0.0211)		(0.0109)		(0.0149)	
AI x Year = -2		-0.00633		-0.00442		0.00107
		(0.0274)		(0.0149)		(0.0193)
AI x Year = -1		0.00795		0.00647		-0.00394
		(0.0216)		(0.0114)		(0.0160)
AI x Year = 0	Dropped	Dropped	Dropped	Dropped	Dropped	Dropped
AI x Year = +1		-0.000185		0.00146		-0.00348
		(0.0181)		(0.00978)		(0.0134)
AI x Year = +2		0.00475		-0.0143		0.0161
		(0.0223)		(0.0123)		(0.0161)
AI x Year = +3		-0.0249		-0.0236		-0.00268
		(0.0242)		(0.0137)		(0.0171)
AI x Year = +4		-0.00885		-0.0306*		0.0177
		(0.0267)		(0.0148)		(0.0187)
AI x Year = +5		0.000276		-0.0217		0.0249
		(0.0290)		(0.0160)		(0.0201)
ln (Emp)	0.104***	0.105***	0.0422***	0.0432***	0.0646***	0.0646***
	(0.0130)	(0.0131)	(0.00743)	(0.00746)	(0.00784)	(0.00787)
Age Bins	Yes	Yes	Yes	Yes	Yes	Yes
Industry-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	33,500	33,500	33,500	33,500	33,500	33,500
R-squared	0.727	0.727	0.728	0.729	0.679	0.679

Robust Standard Errors clustered at the 4-digit NAICS industry level. *, ** and *** denotes significance at the 5%, 1% and 0.1% respectively. Each regression includes a constant, which is not displayed here.

AI knowledge: Measuring AI transfer to industry by tracking the movement of experts



- Any firm seeking to apply frontier AI technology to a particular problem requires a “pyramid” consisting of workers with different levels of AI skills.
- At the moment, the high-level “software architects” who can guide the application of frontier technology are in especially short supply.
- The allocation of these elite software architects across firms, industries, and projects may be especially predictive of success.
- Can one track the high-level architects as they move across firms that employ them?
- We find the leading AI scientists using Elsevier publication data and measure their direct collaboration with firms using coauthored papers. We use faculty websites and other sources to identify their (doctoral/post-doc) students.
- We are then using Revelio data to trace the movements of these students across organizational boundaries over time.
- We can use U.S. Census data to test the hypothesis that the emergence of a “critical mass” of frontier AI researchers within a firm leads to increases in output, employment, and productivity.

Tracing the impact generated by AI experts across time and organizational boundaries

Carnegie Mellon University
School of Computer Science

Google DeepMind

SIEMENS
Healthineers

ARTERYS

Dr. Who

CMU Ph.D.

We can measure the impact of an accumulating stock of AI experts on firm productivity and other outcomes

Event study models

$$y_{it} = \alpha + \beta_1 AI_{it}(1|0) + \beta_2 TIME + \beta_3 AI_{it} \times TIME + X_{it} + \varepsilon$$

Production functions (with fixed effects)

$$q_i - l_i = \alpha k_i + \varphi a_i + \varepsilon_i$$

A report on our (ongoing) progress…

No.	Data point \ Domain	NLP	ML	Robotics	Agents	HCI	Speech	KR	CV	IR
1.	# of Immortals	208	320	160	29	17	140	4	79	459
2.	# of journals/conference venues	722	79	17	58	18	12	32	80	146
3.	# of papers associated with immortals	50,414	74,165	56,476	9,362	4,451	36,396	873	33,509	89,490
4.	# of children of immortals identified/information tagged (so far)	1272	2938	342	297	86	126	No NA based immortal	Yet to be found	Yet to be tagged
5.	Balance # of children of immortals (yet to be tagged)	124	420	5622	101	19	251	-	-	6102
6.	# of direct Immortal-Corporate collaborative papers	8,553	16,026	4,582	665	987	6,859	30	4,301	12,047

Some (very!) preliminary evidence on collaboration with star scientists…

Table 14 The Impact of Collaboration with Elite AI Scientists

	Employment	**Payroll per Employee**	**Revenue**	**Revenue per Employee**	**AI Patents**
Coauthored Publications	+***	+***	+***	+***	+***
Industry-Year FE	Y	Y	Y	Y	Y
Firm FE	N	N	N	N	N
Publication>0 Firms	N	N	N	N	N
Cumulative Coauthored Publications	+***	+***	+***	-	+***
Industry-Year FE	N	N	N	N	N
Firm FE	Y	Y	Y	Y	Y
Publication>0 Firms	N	N	N	N	N
Cumulative Coauthored Publications	-	+***	+	-	+***
Industry-Year FE	N	N	N	N	N
Firm FE	N	N	N	N	N
Publication>0 Firms	Y	Y	Y	Y	Y

(Preliminary) Conclusions

- ML-based approaches can identify AI-related patents.
- We find strong relationships between AI invention and the subsequent (labor) productivity growth of AI-inventing firms!
- Current results omit the impact of recent advances in LLMs and generative AI; currently updating our AI patents to include these recent advances.

Next steps

- Measure direct collaboration between elite AI scientists and U.S. firms.
- Obtain data on the movement of the students of elite AI scientist into U.S. firms.
- Examine the impact of this elite human capital (if any) on U.S. firm output, employment, and productivity.

Thanks!