Patents and the technological evolution of AI

GECON Conferene
Rome, 25-26 September 2024

Arianna Martinelli
(arianna.martinelli@santannapisa.it)

Sant'Anna
School of Advanced Studies – Pisa

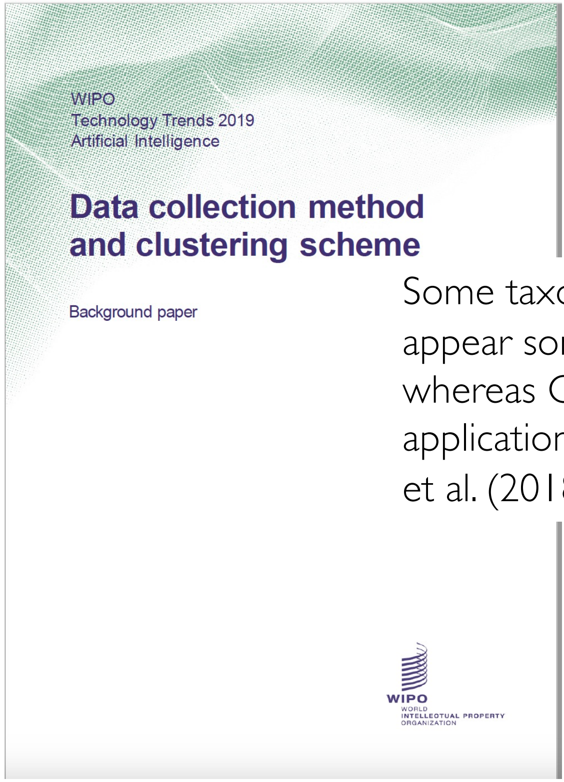
Road Map

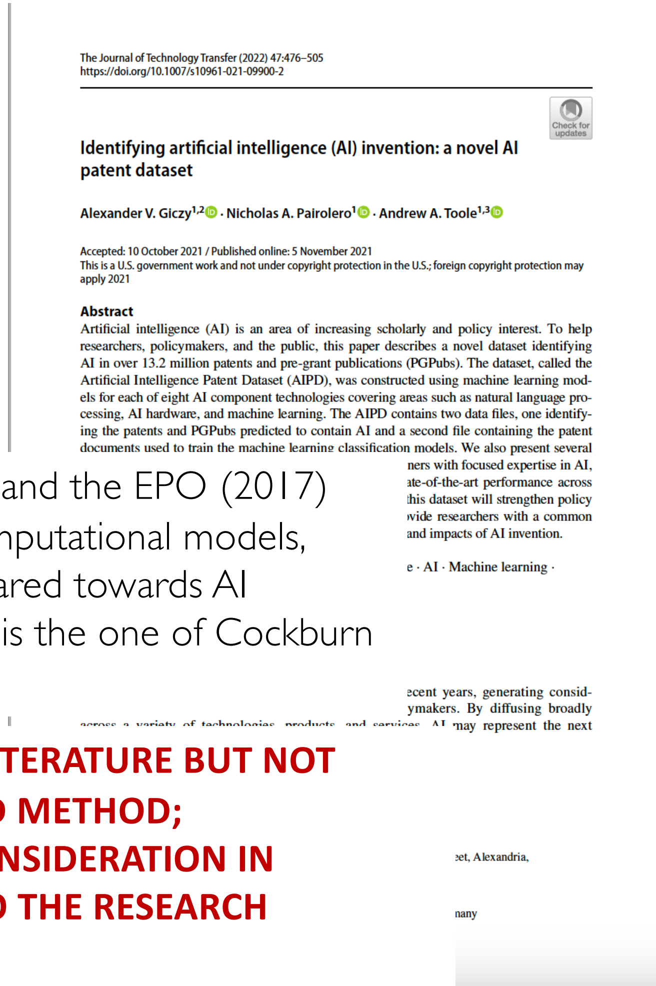
- To present a **new method** to trace technological evolution and use it to assess the effect of **government funding** on the development of **AI.**
- Three steps to cover some broader issues, such as:
  1. **Suitability of patents** to study AI developments
  2. The empirical challenges associated with assessing the **effect of government funding** on innovative activity, in general, and on AI development in particular
  3. Present the **new methodology** based capable of addressing these challenges and an application to the AI domain.

Patents and AI

- Despite well-known limitations, patents have been extensively used to measure inventive activities (Griliches, 1990)
  - "Research industry" dedicated to squeezing patent data to assess inventions characteristics and grasp the features of the underlying inventive process
- Attempts to design procedures to identify AI-related patents and address some major issues - "Making the impossible possible" (OECD, 2020):
  - Define the boundaries of AI and "decide" to what extent include methods specific to different application domains (e.g., industrial robots, autonomous vehicles, medical technologies) and to what extent include techniques that might relate to fundamental research
  - Role of complementary technologies in fuelling AI developments
  - Need to complement patent data to other sources

Identifying AI patents





Some taxonomies such as those of Fujii and Managi (2017) and the EPO (2017) appear somewhat conservative, as they focus mainly on computational models, whereas OECD (2020) experimental definition is rather geared towards AI applications, including image processing or digital devices, as is the one of Cockburn et al. (2018), especially with respect to robotics.

- **GROWING LITERATURE BUT NOT ESTABLISHED METHOD;**
- **CAREFUL CONSIDERATION IN RELATION TO THE RESEARCH QUESTIONS**

Sant'Anna
School of Advanced Studies – Pisa

Public funding and the direction of technical change

- Government and technical change:

  Due to market failures in the production of knowledge (Nelson, 1959; Arrow, 1962), governments play a crucial role in creating incentives and supporting R&D activities in the economy (Bloom et al., 2019).

- Growing interest on the role of public funding:

  While the economic literature focused on firm R&D investments and their spillover effects (Azoulay et al., 2019), the interest in the role of public funding is growing since uncoordinated private investments in new technologies might be insufficient to face complex societal challenges (Mazzucato, 2015; Van Reenen, 2020).

Public funding and the direction of technical change II

- Existing literature studied the impact of public funding on the rate of technical change:

  Rate of returns of R&D investments (Hall et al., 2010); Policy evaluations of the effects of R&D subsidies (Bloom et al., 2002; Wilson, 2009; Dechezlepretre et al., 2016; Akcigit et al., 2018) and government grants (Bronzini and Iachini, 2014; Howell, 2017; Santoleri et al., 2020) on private innovation outcomes.

- Relatively little (systematic, quantitative) evidence on the role of public funding on the direction of technical change:

  Attempt to evaluate the persistence of the government funding effect looking at a very specific historical case: the establishment of the Office of Scientific Research and Development (OSRD) during WWII (Gross and Sampat, 2023)

  Difficult to evaluate outcomes that are produced over the long-run especially by early interventions (Dosi, 1988; Griliches, 1992).

Public funding and the direction of technical change III

- AI likely to be a general purpose technology (GPT) of the coming era (Cockburn et al., 2018; Martinelli et al., 2021): it will favor deep transformations in economic systems and could generate waves of radical innovations leading to widespread economic disruption (Trajtenberg, 2019).
- AI may affect the economy in several ways: it might have a direct effect on growth and labor (Acemoglu and Restrepo, 2018; Korinek and Stiglitz, 2019) as well as on the innovation process itself (Cockburn et al., 2018) and the industrial structure (Varian, 2018).
- The role of government may be important for this kind of GPTs because their development is very risky and, therefore, either very costly or simply impossible to finance by means of private funds, given the uncertainty and time-horizons of returns.

Direction of technical change: technological trajectories

- Evolutionary process:

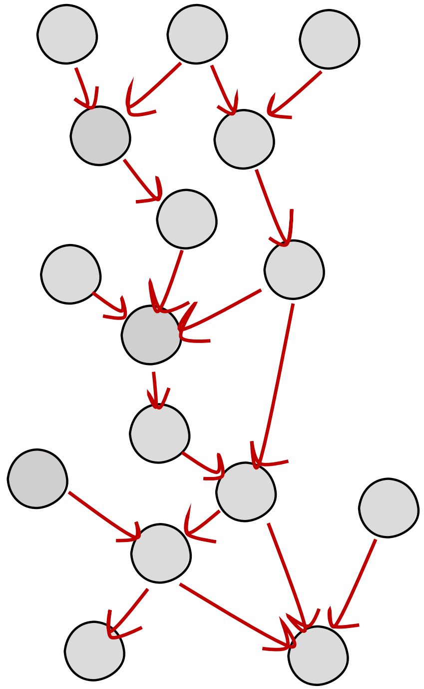
Over time, more useful and valuable knowledge is selected, on which further knowledge will be built, and select out less valuable or obsolete knowledge.

- Cumulativeness of technical change (Dosi, 1982; Dosi, 1988):

New knowledge builds on prior knowledge, often in a recombinatory way (Weitzman, 1998; Wuchty et al., 2007).

- Patterns of cumulative change:

Technological trajectories emerge over time. They can be viewed in retrospect as the path-dependent outcome of dispersed research efforts converging into particular ways of solving problems (Dosi, 1988).



Does government play a role in directing technical change and influencing the patterns of knowledge accumulation (in AI)?

Sant'Anna
School of Advanced Studies – Pisa

The effect of government funding on AI technological trajectory

- Based on "*The direction of technical change in AI and the trajectory effects of government funding*" co-authored with M. Iori and A. Mina
  - Available at: https://www.lem.sssup.it/WPLem/2021-41.html
- Aim of this paper:
  - To investigate the role of government funding on the direction of AI development.
  - To show an application of the connectivity analysis to trace technological trajectories in the case of AI
  - Provide quantitative evidence on the key financing pattern that supported AI development.

Data: AI-related patents granted by the USPTO

- We use patents granted by the USPTO from 1976 to 2019 (EPO-PATSTAT database, Autumn 2019 version).
- We identify AI patents combining specific technological classes (CPC) with a text-based search of technical keywords on patent titles and abstracts (WIPO, 2019; UKIPO, 2014).

- We identify government-funded patents (Fleming et al., 2019):
  - Government assignee patents: Patents assigned to federal agencies, national laboratories, and state departments (EPO-PATSTAT and Patentsview disambiguation of assignee and applicant categories)
  - Government interest patents: Inventions developed with federal funding (e.g. as per a Government Interest Statement)

USPTO patents in AI

We select 114,670 USPTO patents

Technology name	Number of patents	%
Electrical engineering - Computer technology	74192	64.72
Instruments - Control	8513	7.43
Mechanical engineering - Transport	5378	4.69
Instruments - Measurement	4346	3.79
Electrical engineering - Audio-visual technology	4306	3.76
Instruments - Medical technology	3587	3.13
Electrical engineering - Digital communication	2804	2.45
Electrical engineering - Telecommunications	2110	1.84
Electrical engineering - IT methods for management	1808	1.58
Mechanical engineering - Mechanical elements	1094	0.95

Main technologies in AI patents

Assignee	Number of patents	%
International Business Machines Corporation	6710	5.85
Microsoft Corporation	3927	3.42
Google Inc.	3094	2.70
Canon Kabushiki Kaisha	1834	1.60
Samsung Electronics Co., Ltd.	1655	1.44
Sony Corporation	1602	1.40
AT&T Corporation	1191	1.04
Amazon Technologies, Inc.	1169	1.02
Xerox Corporation	1087	0.95
Fujitsu Limited	1068	0.93

Main assignees of AI patents

Government funded patents in AI: the role of the Department of Defense

929 government assignee patents

Assignee	Number of patents	%
Secretary of the Navy	370	39.83
National Aeronautics and Space Administration	153	16.47
Secretary of the Army	109	11.73
Secretary of the Air Force	106	11.41
Department of Energy	33	3.55
National Security Agency	29	3.12
Department of Health and Human Services	29	3.12
United States Postal Service	22	2.37
Lawrence Livermore National Security	19	2.05
Department of Commerce	10	1.08

Federal agency	Number of patents	%
Department of Defense	1670	46.43
United States Government	703	19.54
Department of Health and Human Services	627	17.43
National Science Foundation	478	13.29
Department of Energy	462	12.84
National Aeronautics and Space Administration	166	4.61
Small Business Administration	42	1.17
Department of Transportation	36	1.00
Department of Commerce	36	1.00
Department of Homeland Security	35	0.97

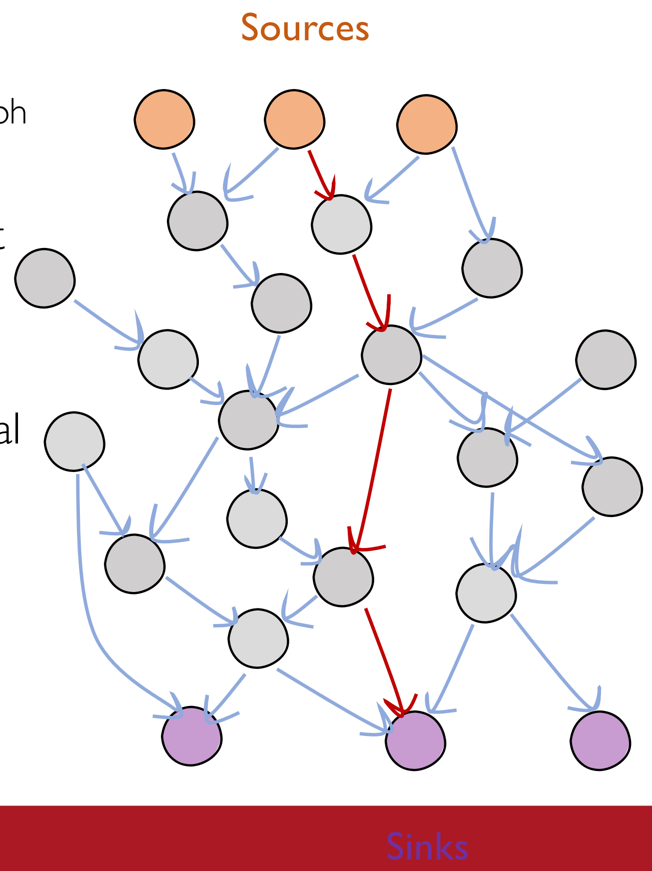
3597 government interest patents

Mapping technological trajectories using patent citation networks

- Patents disclose (and therefore embed) information about a new solution to a specific technical problem
- If patent B cites patent A, there is a knowledge flow between the two patent: A→B
- Identified all the patents related to a technology and all their citations it is possible to build a binary and directed network that represents the available/possible technological space

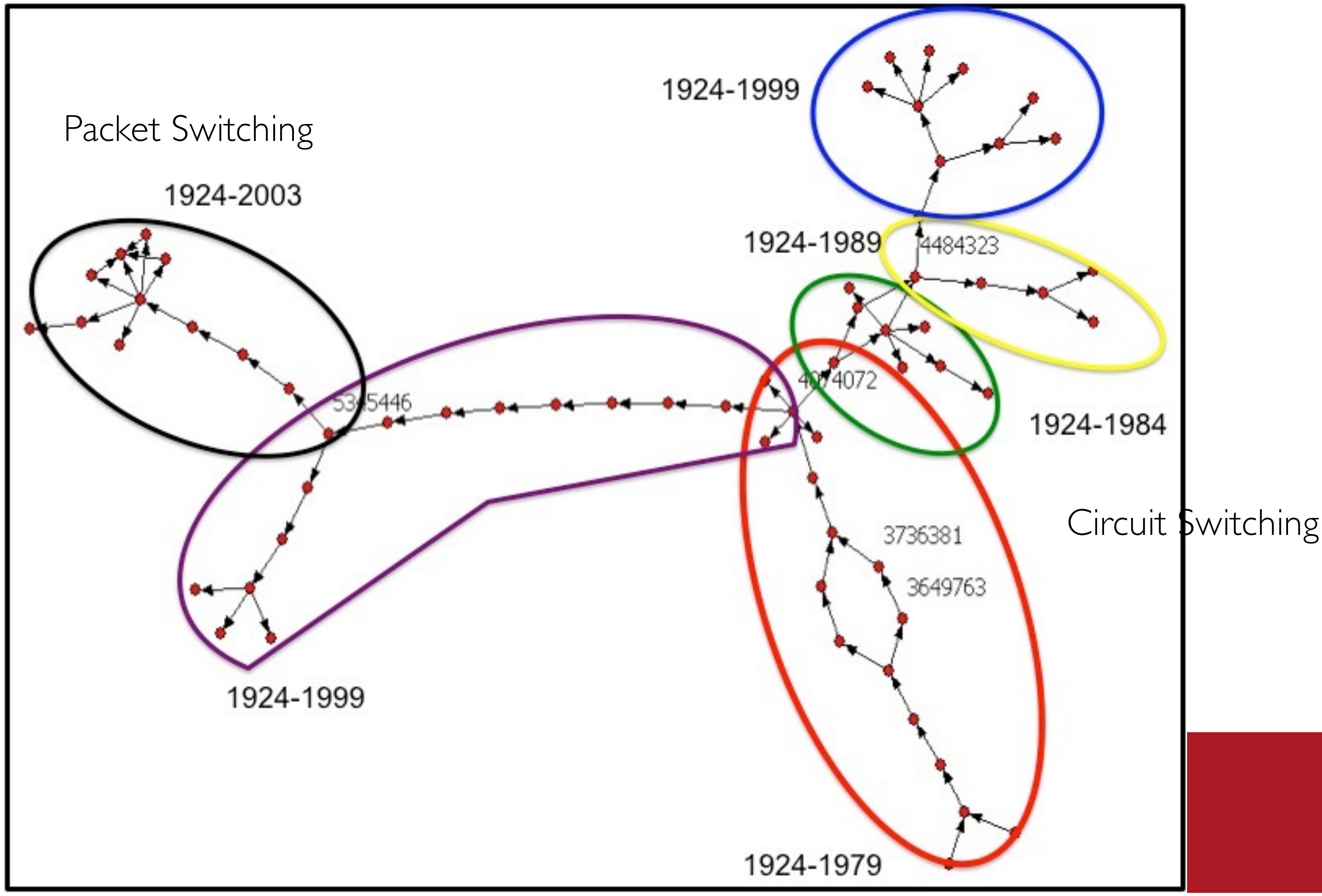
New indicators to measure trajectory effect

- We create a citation network of 514,599 nodes and 2,661,528 edges from AI inventions and their references.
  - Citations respect the time flow and there are no loops: Directed Acyclic Graph (DAG).
- In DAG, it is possible to define paths from sources to sinks without encountering each node more than once.
- Connectivity indicators (Hummon and Doreian, 1989; Batagelj, 2003): Search Path Count assigns to each edge (u; v) a weight equal to the number of paths from s to t through (u; v).
  - The higher the weight, the more important the edge is for network connectivity and the development of the entire technological domain.
- Paths with the largest sum of SPC identify the most relevant trajectories.
  - Early explorations of this methodology (Mina et al., 2007; Martinelli, 2012) used traversal counts to in small technological domains.



Sant'Anna
School of Advanced Studies – Pisa

Technological trajectories for the telecommunication switches



Martinelli, 2012

Technological trajectories in the 2G and 3G (CDMA)

Bekkers and Martinelli, 2012

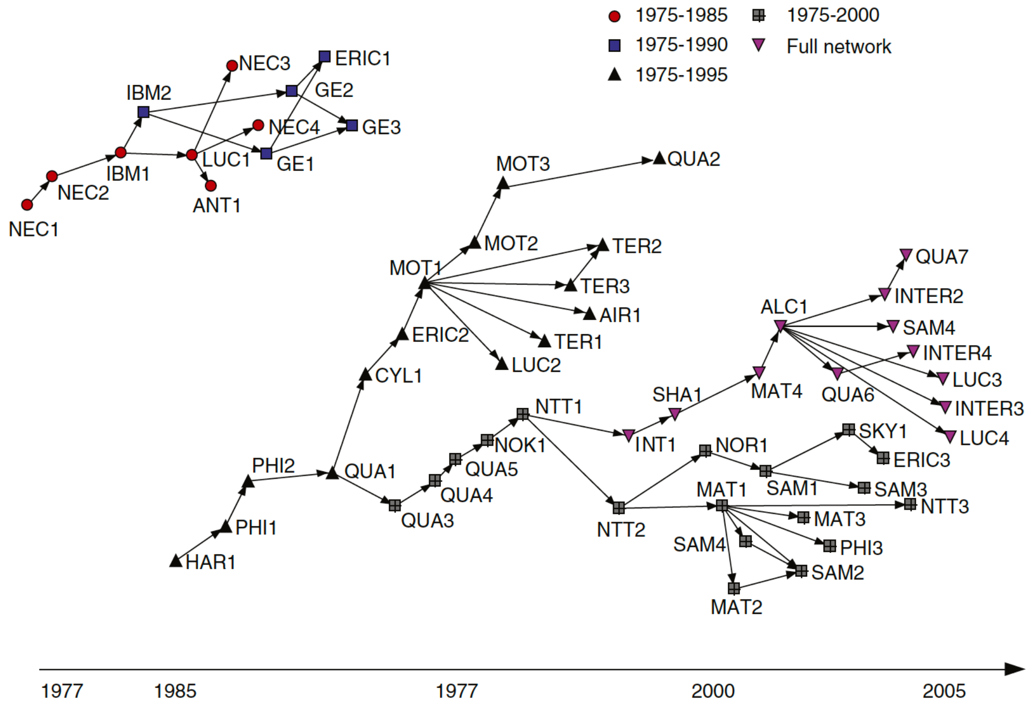
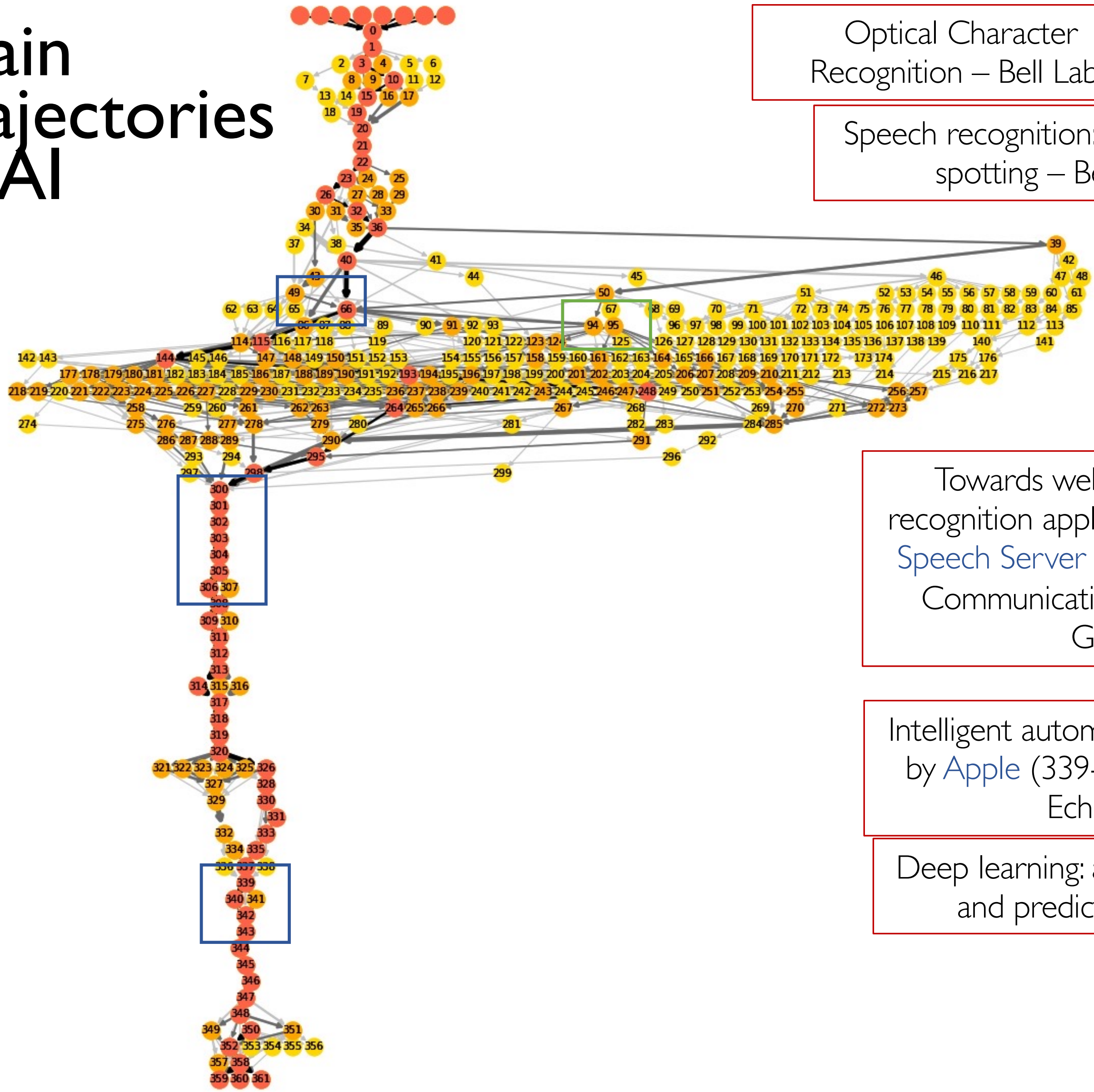


Fig. 7. Technological trajectories and the patent's assignees, for five time periods.

Main trajectories in AI



Optical Character Recognition – Bell Labs

Early 1950s

Speech recognition: templates and keyword spotting – Bell Labs and NEC

1950s – 1970s

Speech recognition: probabilistic learning – Hidden Markov Models at Bell Labs (49) and IBM (66) – and first commercial software by Dragon Systems (88,94,95)

1980s

Speech recognition: towards Natural Language Processing

1990s

Towards web-based speech-recognition applications – Microsoft Speech Server (300-307), Nuance Communications, Amazon and Google

2000

Intelligent automatic assistants – Siri by Apple (339-343) and Amazon Echo (344)

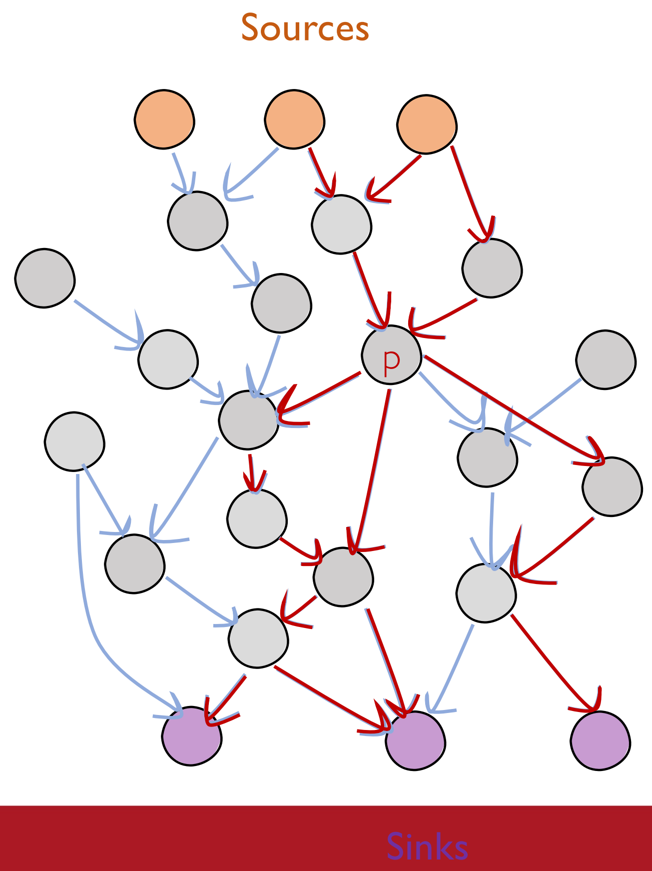
2010

Deep learning: applications in multimedia language context and predictions of future translations – Facebook

2015

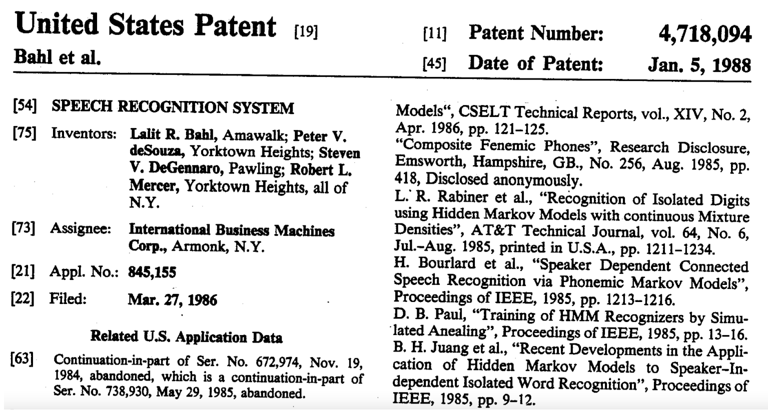
New indicators to measure trajectory effect (II)

- We can extend the methodology to the nodes of the network to measure the relevance of a single patent from a trajectory perspective.
- Trajectory indicator: this measure captures the number of paths from s to t through the patent p.
- A patent with a high weight is a patent that channels and "cumulates" large knowledge flows within the network.
- Global measure of relevance
  - Different from local measure such as number of citations -> number of outward arrows (i.e. outdegree centrality)



New indicators to measure trajectory effect (III)

- A network approach allows to consider a long-run perspective on follow-on innovation
  - Citations consider only short-run development
- Evidence of patents covering technological breakthroughs that receive a relatively low number of citations but that are on the trajectory
  - Not in the top 1% of citations distribution (Ahuja and Lampert, 2001)
- Example from our data: patents disclosing system of probabilistic learning in speech recognition research



United States Patent [19] — Bahl et al.

[11] Patent Number: **4,718,094**
[45] Date of Patent: **Jan. 5, 1988**

[54] SPEECH RECOGNITION SYSTEM

[75] Inventors: Lalit R. Bahl, Amawalk; Peter V. deSouza, Yorktown Heights; Steven V. DeGennaro, Pawling; Robert L. Mercer, Yorktown Heights, all of N.Y.

[73] Assignee: International Business Machines Corp., Armonk, N.Y.

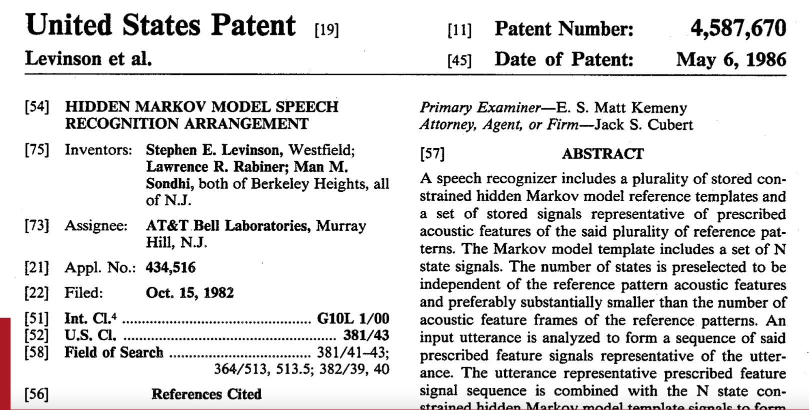
[21] Appl. No.: 845,155

[22] Filed: Mar. 27, 1986

Related U.S. Application Data

[63] Continuation-in-part of Ser. No. 672,974, Nov. 19, 1984, abandoned, which is a continuation-in-part of Ser. No. 738,930, May 29, 1985, abandoned.

Models", CSELT Technical Reports, vol., XIV, No. 2, Apr. 1986, pp. 121–125.
"Composite Fenemic Phones", Research Disclosure, Emsworth, Hampshire, GB., No. 256, Aug. 1985, pp. 418, Disclosed anonymously.
L. R. Rabiner et al., "Recognition of Isolated Digits using Hidden Markov Models with continuous Mixture Densities", AT&T Technical Journal, vol. 64, No. 6, Jul.–Aug. 1985, printed in U.S.A., pp. 1211–1234.
H. Bourlard et al., "Speaker Dependent Connected Speech Recognition via Phonemic Markov Models", Proceedings of IEEE, 1985, pp. 1213–1216.
D. B. Paul, "Training of HMM Recognizers by Simulated Anealing", Proceedings of IEEE, 1985, pp. 13–16.
B. H. Juang et al., "Recent Developments in the Application of Hidden Markov Models to Speaker-Independent Isolated Word Recognition", Proceedings of IEEE, 1985, pp. 9–12.



United States Patent [19] — Levinson et al.

[11] Patent Number: **4,587,670**
[45] Date of Patent: **May 6, 1986**

[54] HIDDEN MARKOV MODEL SPEECH RECOGNITION ARRANGEMENT

[75] Inventors: Stephen E. Levinson, Westfield; Lawrence R. Rabiner; Man M. Sondhi, both of Berkeley Heights, all of N.J.

[73] Assignee: AT&T Bell Laboratories, Murray Hill, N.J.

[21] Appl. No.: 434,516

[22] Filed: Oct. 15, 1982

[51] Int. Cl.⁴ G10L 1/00
[52] U.S. Cl. 381/43
[58] Field of Search 381/41–43; 364/513, 513.5; 382/39, 40

[56] References Cited

Primary Examiner—E. S. Matt Kemeny
Attorney, Agent, or Firm—Jack S. Cubert

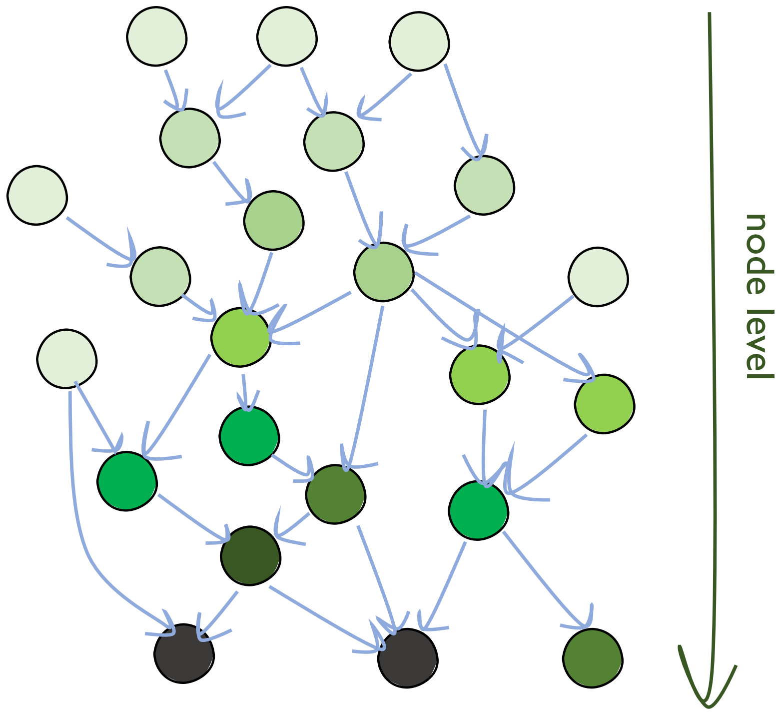
[57] ABSTRACT

A speech recognizer includes a plurality of stored constrained hidden Markov model reference templates and a set of stored signals representative of prescribed acoustic features of the said plurality of reference patterns. The Markov model template includes a set of N state signals. The number of states is preselected to be independent of the reference pattern acoustic features and preferably substantially smaller than the number of acoustic feature frames of the reference patterns. An input utterance is analyzed to form a sequence of said prescribed feature signals representative of the utterance. The utterance representative prescribed feature signal sequence is combined with the N state constrained hidden Markov model template signals to form

Sant'Anna
School of Advanced Studies – Pisa

Time in citation networks: the node level indicator

- Indicator of timing in directed citation network: node position (distance from the sources) in the graph.
- Node level indicator marks time in terms of the patent citation network and overall evolution of the field.
- Node level takes value 0 for network sources and, for all the other patents, it is equal to 1 plus the maximum node level of their cited patents.
- Low values refer to the early stages of the technology (i.e., closer to sources).
- High values indicate innovations in a mature phase (i.e., closer to sinks).



Sant'Anna
School of Advanced Studies – Pisa

Estimating the role of government funding on the trajectory effect

We estimate, for patent p in technological class i, the following baseline specification:

$$\begin{aligned} Ln(\text{trajectory effect}_{pi}) = & \beta_0 + \beta_1 \,\text{government funding}_p \\ & + \beta_2 \,\text{government funding}_p \times \text{timing}_p \\ & + \beta_3 \,\text{timing}_p + \gamma_p + \delta_i + \epsilon_{pi}, \end{aligned}$$

We add:

- Controls at the patent level: number of claims, inventors' team size, US university as assignee (dummy).
- Sub-field (3-digits CPC) fixed effects to control for diverse citation behavior in different fields.

Note that this analysis is run on the set of 114670 patents related to AI

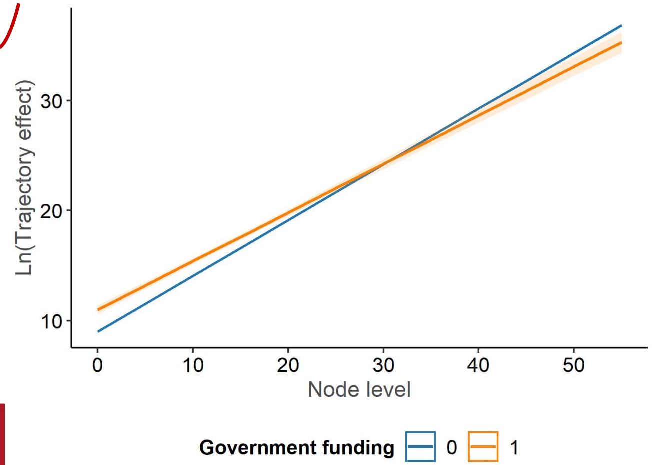
Sant'Anna
School of Advanced Studies – Pisa

The role of government funding - Results

	Dependent variable:		
	log(Trajectory effect)		
	(1)	(2)	(3)
Government funding	1.184***	1.096***	1.959***
	(0.132)	(0.147)	(0.263)
Government funding*Timing			-0.064***
			(0.011)
US university		0.272	0.282*
		(0.166)	(0.166)
Timing	0.503***	0.503***	0.505***
	(0.002)	(0.002)	(0.002)
Number of claims	0.043***	0.043***	0.043***
	(0.002)	(0.002)	(0.002)
Number of inventors	-0.106***	-0.106***	-0.106***
	(0.011)	(0.011)	(0.011)
Intercept	8.594***	8.592***	8.562***
	(0.078)	(0.078)	(0.078)
3-digit CPC	Yes	Yes	Yes
Observations	114,670	114,670	114,670
R^2	0.435	0.435	0.435
Adjusted R^2	0.435	0.435	0.435
Residual Std. Error	7.292	7.292	7.291
F Statistic	3078.115***	3008.006***	2951.426***

Note: All the models are estimated using OLS.
Robust standard errors are reported in parenthesis.
Legend: *p<0.1; **p<0.05; ***p<0.01

Patents receiving government funding have, on average, a trajectory effect 223.9% higher than other patents



Government grants vs. government inventions

Patents receiving government grants have, on average, a trajectory effect 164.9% higher than other patents

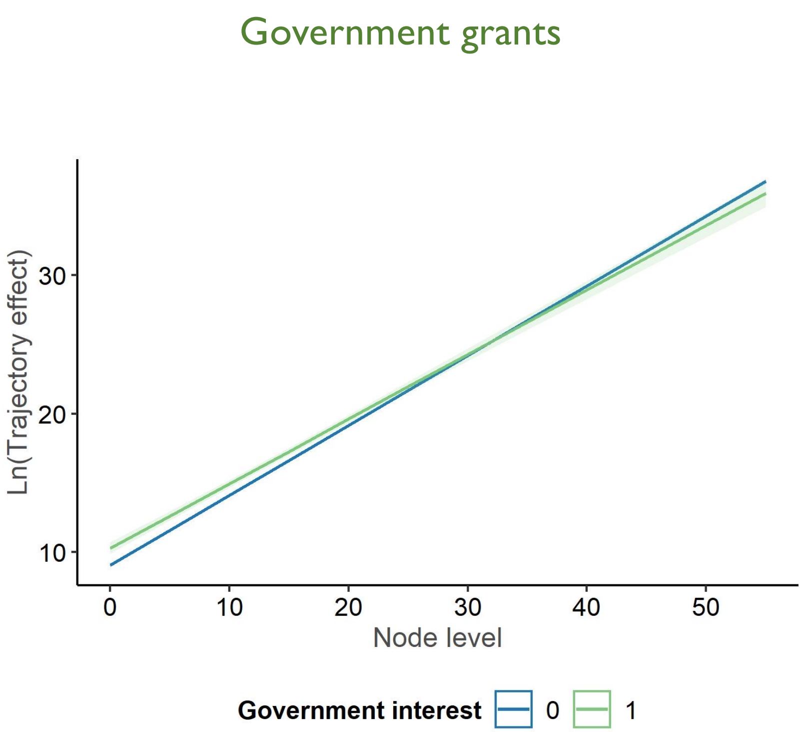
	Dependent variable:					
	log(Trajectory effect)					
	(1)	(2)	(3)	(4)	(5)	(6)
Government interest	0.983***		0.460***	0.999***	0.481***	0.622**
	(0.134)		(0.157)	(0.281)	(0.156)	(0.288)
Government interest*Timing				-0.037***		-0.010
				(0.012)		(0.012)
Government assignee		2.322***	2.050***	1.959***	4.323***	4.233***
		(0.321)	(0.338)	(0.340)	(0.537)	(0.562)
Government assignee*Timing					-0.230***	-0.224***
					(0.030)	(0.031)
US university			0.551***	0.541***	0.551***	0.548***
			(0.168)	(0.168)	(0.167)	(0.168)
Timing	0.503***	0.503***	0.504***	0.505***	0.505***	0.505***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Number of claims	0.043***	0.043***	0.043***	0.043***	0.043***	0.043***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Number of inventors	-0.106***	-0.102***	-0.104***	-0.104***	-0.105***	-0.105***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
Intercept	8.610***	8.597***	8.574***	8.560***	8.555***	8.551***
	(0.078)	(0.078)	(0.078)	(0.078)	(0.078)	(0.078)
3-digit CPC	Yes	Yes	Yes	Yes	Yes	Yes
Observations	114,670	114,670	114,670	114,670	114,670	114,670
R^2	0.435	0.435	0.435	0.435	0.435	0.435
Adjusted R^2	0.435	0.435	0.435	0.435	0.435	0.435
Residual Std. Error	7.294	7.293	7.291	7.291	7.289	7.289
F Statistic	3074.472***	3078.966***	2944.038***	2886.062***	2891.010***	2831.363***

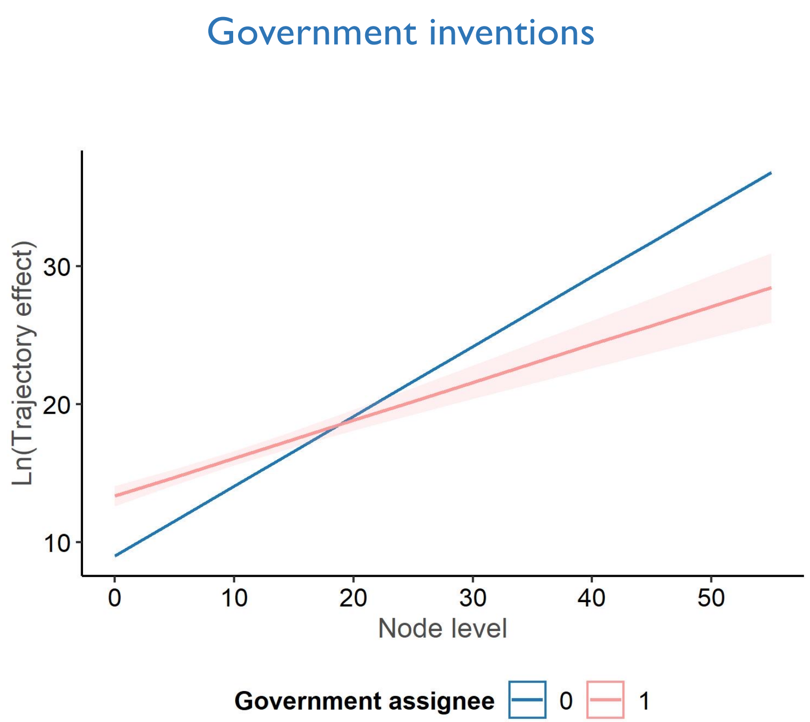
Note: All the models are estimated using OLS.
Robust standard errors are reported in parenthesis.
Legend: *p<0.1; **p<0.05; ***p<0.01

Federal agencies or state departments patents have, on average, a trajectory effect 868.4% higher than other patents

Sant'Anna
School of Advanced Studies – Pisa

The effect is stronger in early phases of development





Robustness checks: potential sources of endogeneity

- 1-1 matching without replacement: propensity score matching on technology classes (3-digits CPC) and node levels (timing)

	Dependent variable:					
	log(Trajectory)					
	(1)	(2)	(3)	(4)	(5)	(6)
Government funding	1.372***	4.849***				
	(0.191)	(0.317)				
Government funding*Timing		−0.259***				
		(0.014)				
Government interest			1.072***	4.433***		
			(0.198)	(0.332)		
Government interest*Timing				−0.242***		
				(0.015)		
Government assignee					2.977***	7.897***
					(0.384)	(0.605)
Government assignee*Timing						−0.493***
						(0.037)
US university	−1.485***	−1.424***	−1.258***	−1.260***	0.852	1.407
	(0.255)	(0.254)	(0.258)	(0.257)	(1.115)	(1.124)
Timing	0.573***	0.702***	0.578***	0.698***	0.535***	0.783***
	(0.008)	(0.009)	(0.008)	(0.009)	(0.022)	(0.023)
Number of claims	0.039***	0.041***	0.045***	0.046***	0.001	0.009
	(0.006)	(0.006)	(0.007)	(0.006)	(0.014)	(0.013)
Number of inventors	−0.162***	−0.157***	−0.129***	−0.126***	−0.231**	−0.223**
	(0.043)	(0.043)	(0.044)	(0.044)	(0.106)	(0.101)
Constant	8.083***	6.305***	7.846***	6.143***	8.567***	5.909***
	(0.245)	(0.248)	(0.251)	(0.256)	(0.525)	(0.506)
3-digit CPC	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7,864	7,864	7,194	7,194	1,858	1,858
R^2	0.449	0.468	0.470	0.487	0.301	0.348
Adjusted R^2	0.446	0.465	0.467	0.484	0.287	0.335
Residual Std. Error	7.244	7.118	7.089	6.975	8.216	7.936
F Statistic	155.571***	164.013***	154.672***	161.664***	21.185***	25.575***

Note: All the models are estimated using OLS on data matched through propensity score matching (1-1) without replacement) Robust standard errors are reported in parenthesis.
Legend: *p<0.1; **p<0.05; ***p<0.01

Sant'Anna
School of Advanced Studies – Pisa

Robustness checks: potential sources of endogeneity

- Instrumental variable: the predicted number of patents related to defense R&D in the CPC classes associated to each patent, (following Moretti et., 2019)

	Dependent variable:					
	log(Trajectory)					
	(1)	(2)	(3)	(4)	(5)	(6)
Government funding	45.915***	70.948***				
	(3.446)	(6.403)				
Government funding*Timing		−1.899***				
		(0.211)				
Government interest			53.566***	92.888***		
			(4.318)	(9.173)		
Government interest*Timing				−2.625***		
				(0.298)		
Government assignee					102.234***	109.005***
					(11.167)	(13.467)
Government assignee*Timing						−1.190
						(0.866)
US university	−18.951***	−18.412***	−22.288***	−23.599***	1.072***	1.103***
	(1.550)	(1.744)	(1.932)	(2.431)	(0.213)	(0.206)
Timing	0.521***	0.577***	0.517***	0.591***	0.539***	0.543***
	(0.003)	(0.007)	(0.003)	(0.009)	(0.004)	(0.005)
Number of claims	0.046***	0.048***	0.044***	0.046***	0.060***	0.060***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
Number of inventors	−0.156***	−0.145***	−0.174***	−0.166***	−0.038**	−0.042**
	(0.017)	(0.017)	(0.018)	(0.020)	(0.018)	(0.018)
Intercept	7.062***	6.179***	7.207***	6.000***	6.484***	6.490***
	(0.144)	(0.227)	(0.144)	(0.263)	(0.238)	(0.240)
3-digit CPC	Yes	Yes	Yes	Yes	Yes	Yes
Observations	114,670	114,670	114,670	114,670	114,670	114,670
F-test	181.6***	114.83***	154.4***	106.61***	88.16***	40.98***
F-test (interaction)		91.43***		88.04***		24.65***

Note: All the models are estimated using 2SLS.
Robust standard errors are reported in parenthesis.
Legend: *p<0.1; **p<0.05; ***p<0.01

Other robustness checks

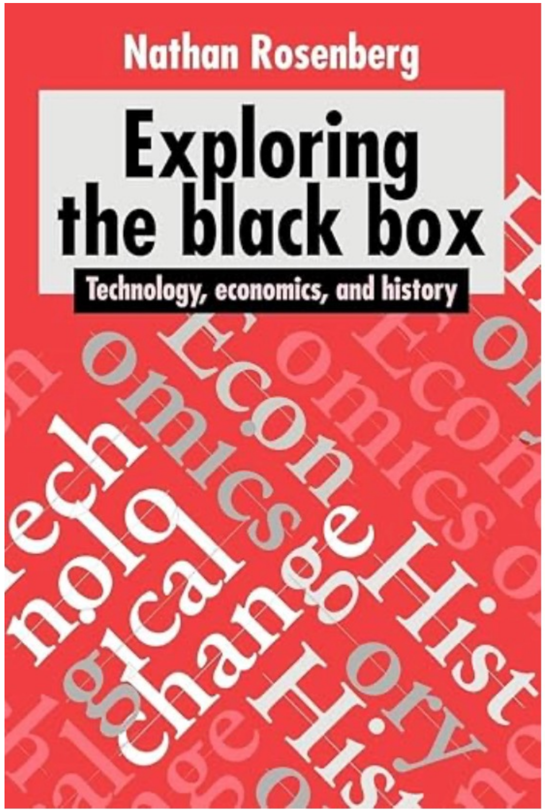
- Other robustness checks:
  - Other indicators of trajectory effect: longest path length
  - Time effect: forward trajectory effect
  - Indirect citations of government funding (following Fleming et al., 2019)
  - Sample composition: only WIPO (2019) patents and patents after 1980
  - Additional controls: all world universities, backward citations, and average growth rate of CPC classes (lagged)

- Patent relevance: tests on the effects of key variables on standard indicators (number of citations) give very different results (generally negative effects!).

Sant'Anna
School of Advanced Studies – Pisa

Conclusions

- Combining patent data and network analysis techniques provides the ground for empirically grasping technology dynamics even for difficult-to-identify technology such as AI
- ``Toolbox'' to study long-term technological development useful to frame specific research questions (i.e. technological catching up, building specific technological capabilities)
  - Empirically explore and open up the black box of innovation



Sant'Anna
School of Advanced Studies – Pisa

Conclusions

- US government grants and, especially, patents filed by federal agencies and government departments had profound effects on AI innovation. Their impact was stronger in early phase of technological development, while it weakened over time to leave room to privately funded research.
  - Novel quantitative evidence of key financing patterns that have supported the development of these technologies over the last few decades

Sant'Anna
School of Advanced Studies – Pisa

Thank you!

arianna.martinelli@santannapisa.it

Back-up slides

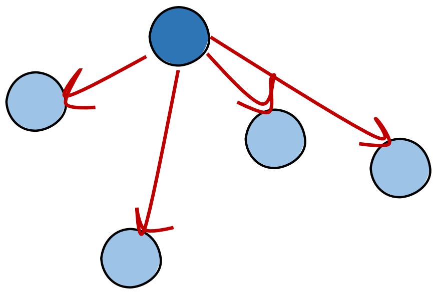
Detecting technological trajectories: methodology

1. CALCULATION OF SPx FOR EACH EDGE;
2. SEARCH OF MAIN PATH: from each starting point selects the sequence of edges with the highest SPx;
3. IDENTIFICATION OF TOP-PATH: select a list of connected patents and citations whose sum of SPxs is the highest

-> POSSIBILITY OF REPLICATE THESE STEPS FOR DIFFERENT TIME-PERIOD

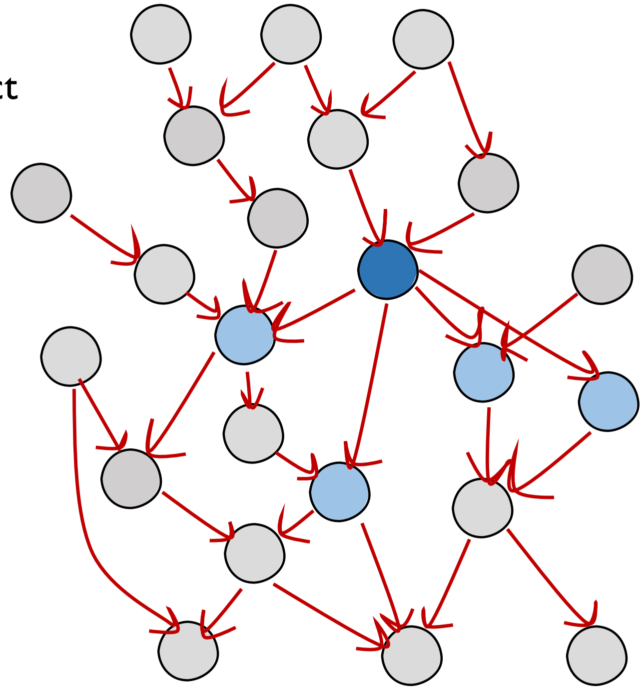
The need for new indicators : short-run vs. long-run impact of inventions

Number of citations



Long-term cumulative impact of new knowledge is not captured by standard patent citation measures

Trajectory effect



Citation networks: chains of local, cumulative, and irreversible technological developments, consistently with the definition of technological trajectories (Dosi, 1982; Verspagen, 2007)

Sant'Anna
School of Advanced Studies – Pisa

Descriptive statistics

Variable name	Min	Mean	Max	Std
Trajectory	1	$8.25 \cdot 10^{15}$	$1.87 \cdot 10^{19}$	$2.20 \cdot 10^{17}$
Longest path length	1	$2.27 \cdot 10^{19}$	$1.73 \cdot 10^{20}$	$3.74 \cdot 10^{19}$
Timing	0	16.17	55	11.27
Number of claims	1	18.92	522	12.35
Number of inventors	1	2.74	27	18.80
Application year	1952	2007.75	2019	8.49
Grant year	1976	2010.80	2019	8.41
Number of references	0	31.07	3951	90.61
Number of citations (network)	0	10.92	805	24.08
Number of citations (all)	0	25.07	2288	53.32
Number of citations up to 5 years (network)	0	4.55	240	8.63
Number of citations up to 5 years (all)	0	10.21	1156	18.28
Number of citations in 5 years (network)	0	5.94	240	9.50
Number of citations in 5 years (all)	0	14.16	1156	20.74

Descriptive statistics II

Variable name	Number of patents	%
Government funding	3932	3.43
Government interest	3597	3.14
Government assignee	929	0.81
US university	2947	2.57
University	4588	4.00
Citing government funding	34692	30.25
Citing government interest	31837	27.76
Citing government assignee	14075	12.27
Citing US university	31491	27.46

Construction of the instrument (Moretti et. Al. 2020)

1. We identify patents related to defense R&D by selecting USPTO patents that received government funding from the US Department of Defense or have this department (or one of its divisions, such as Army, Navy, or Air Force) as assignee.
2. Each patent related to defense R&D is then associated to 4-digit CPC classes. Since each patent may be associated with more than one CPC class, we introduce weights proportional to the importance of these classes in the patent. Then, we compute the weighted number of patents related to the US Department of Defense for each 4-digit CPC class.
3. To obtain results that are comparable over time, we normalized the number of patents associated to defense R&D in each CPC class by the total number of patents in that class.
4. The resulting indicator can be interpreted as a measure of the importance of defense R&D in each 4-digit CPC class. Moreover, since we are interested to capturing the predicted number of patents, we introduce a one-year lag. Therefore, for each 4-digit CPC class i at the time t, we compute:

$$\text{Predicted defense patents in CPC}_{i,t} = \frac{\text{Number of defense-related patents}_{i,t-1}}{\text{Number of patents}_{i,t-1}}.$$

Sant'Anna
School of Advanced Studies – Pisa

Construction of the instrument (Moretti et. Al. 2020) II

5. Then, we define the instrumental variable Predicted defense patents$_{p;t}$ for each patent p with application year t as the weighted average of Predicted defense patents in CPC$_{i;t}$ over the collection CPC$_p$ of 4-digit CPC classes related to the patent:

$$\text{Predicted defense patents}_{p,t} = \sum_{i \in CPC_p} \text{share}_i \cdot \text{Predicted defense patents in CPC}_{i,t},$$

IV first stage

	Dependent variable:		
	Government funding	Government interest	Government assignee
	(1)	(2)	(3)
Predicted defense patents	1.359***	1.165***	0.610***
	(0.101)	(0.094)	(0.065)
US university	0.427***	0.428***	−0.004***
	(0.009)	(0.009)	(0.001)
Timing	−0.0003***	−0.0002***	−0.0003***
	(0.00004)	(0.00004)	(0.00002)
Number of claims	−0.0001	−0.00002	−0.0002***
	(0.00004)	(0.00004)	(0.00002)
Number of inventors	0.001***	0.001***	−0.001***
	(0.0003)	(0.0002)	(0.0001)
Intercept	0.015***	0.011***	0.013***
	(0.002)	(0.002)	(0.001)
3-digit CPC	Yes	Yes	Yes
Observations	114,670	114,670	114,670
R^2	0.160	0.170	0.016
Adjusted R^2	0.159	0.169	0.016
Residual Std. Error	0.167	0.159	0.089
F Statistic	495.199***	532.877***	43.222***

Note: All the models are estimated using OLS.
Robust standard errors are reported in parenthesis.
Legend: *p<0.1; **p<0.05; ***p<0.01

Sant'Anna
School of Advanced Studies – Pisa