artificial intelligence index

2018 annual report

Steering Committee

Yoav Shoham *(Chair)*
Stanford University

Raymond Perrault
SRI International

Erik Brynjolfsson
MIT

Jack Clark
OpenAI

James Manyika
McKinsey Global Institute

Juan Carlos Niebles
Stanford University

Terah Lyons
Partnership On AI

John Etchemendy
Stanford University

Barbara Grosz
Harvard University

Project Manager

Zoe Bauer

Welcome to the AI Index 2018 Report

Our Mission is to ground the conversation about AI in data.

The AI Indexは、人工知能に関連するデータの足跡をたどり、照合し、 意味を汲み取った上で視覚化した努力の結晶です。
政策立案者や研究者、経営者、ジャーナリストなど多くの方々に、 AIという複雑な分野についての直感的な理解を得るための包括的な資料として活用してもらいたいと考え、AI Indexを制作しました。

引用元としての利用

Yoav Shoham, Raymond Perrault, Erik Brynjolfsson, Jack Clark, James Manyika, Juan Carlos Niebles, Terah Lyons, John Etchemendy, Barbara Grosz and Zoe Bauer, "The AI Index 2018 Annual Report", AI Index Steering Committee, Human-Centered AI Initiative, Stanford University, Stanford, CA, December 2018.

(c) 2018 by Stanford University, "The AI Index 2018 Annual Report" is made available under a Creative Commons Attribution-NoDerivatives 4.0 License (International) https://creativecommons.org/licenses/by-nd/4.0/legalcode

Table of contents

Introduction to the AI Index 2018 Report 5
Overview 6
Volume of Activity 8
- Research
  - Published Papers 9
  - Course Enrollment 21
  - Participation 26
  - Robot Software 29
- Industry
  - Startups / Investment 31
  - Jobs 33
  - Patents 35
  - AI Adoption Survey 36
  - Earnings Calls 38
  - Robot Installations 41
- Open Source Software
  - GitHub Project Statistics 42
- Public Interest
  - Sentiment of Media Coverage 43
  - Government mentions 44

Technical Performance
- Vision 47
- Natural Language Understanding 50

Other Measures
- Derivative Measures 55
- Government Initiatives 57

Human-Level Performance Milestones 59
What's Missing? 63
Acknowledgements 66
Appendix 69

Introduction to the AI Index 2018 Report

昨年に引き続き、2018年も本レポートを発表できたことを光栄に思う。
2018年、昨年の評価指標の更新と AI分野に関する国際的な現状を伝えるため、本レポートを作成した。
AI Indexのミッションである「 AIに関する議論の基礎を築きあげること」とは、すなわち AIの技術的発展と量的発展を継続的に追い続けることを意味している。
そのため、評価指標の更新は AI Indexの重要な使命である。しかし、AI分野における国際的な見解を知らずして、AIの変遷を語ることはできないため、 AI分野の国際的な現状を伝えることもまた、 AI Indexの欠かせない使命となっている。
2017年版のAI Indexでは、グローバルパートナーが限られていたため、北米における AIに関する動向に焦点を当てざるをえなかった。本レポートをさらに包括的なものにするため、より多くの協力と、外部からの参加者が必要だと認識した上で、2018年版では昨年以上の国際的な動向を記している。

2017年、Scopus（世界最大の論文引用データベース）に登録された AI関連論文の83%はアメリカ国外からのものだった。AI関連論文の登録数で最も大きな割合を占めたのはヨーロッパだった。
大学におけるAI、機械学習分野の講義受講者数が世界的に増加を続ける中、特に目を引くのが中国の清華大学だ。清華大学では 2010年と比べ、AI、機械学習関連講義の受講者数が 16倍に増加している。中国は世界最大の産業用ロボット生産国であり、 2017年には、一国で世界全体の 30%もの産業用ロボットを出荷している。
AIの隆盛はアメリカ、中国、ヨーロッパに限られたものではなく、世界中で起こっている。 AI関連特許の数で見た場合、2014年には韓国がアメリカに次いで 2位、日本が3位の新規特許を生み出している。
世界規模でのAIの隆盛を裏付けるように、南アフリカでは、アフリカの 20ヶ国以上から500人を超える参加者を集め、世界最大級の機械学習イベント、第 2回Deep Learning Indaba会議を開催した。

AIの多様性は地理的なものだけではない。
ACLU、Amnesty International、Oxford’s Future of Humanitiy Institute、そしてUnited Nations Development Programmeなど、AI開発者の集う非営利団体は 50を超える。その上、AI分野における性別や人種の多様性を重要視する認識が高まっている。 AI4AllやWoman in Machine Learningといった団体は、AI分野に対する関わりの不十分なグループの AI分野への参入を促し、参加に必要な支援をおこなっている。

Overview

本レポートは大まかに以下の 4セクションに分けられる。

1. データ: Volume of Activity、Technical Performance
2. その他の基準: Recent Government Initiative、Derivative measures & Human-Level Performance Metrics
3. 議論: What's Missing?
4. Appendix

<u>データ</u>

Volume of Activityの指標では、学界、企業、起業家、世間一般の視点から見たAIに関する洞察を示す。本レポートで計測に用いられたデータは、AIについて学ぶ大学生の数やAI関連の求人に対する女性応募者の割合、各EU加盟国の産業用ロボットの生産量の成長率など、多岐にわたっている。

Technical Performanceでは、AIの性能の変遷を経時的に表している。例えば、本レポートでは、AIが質問に返答した際の回答の質や、物体検知の訓練にかかる時間の変遷を経時的にまとめている。

2018年版の本レポートでは、昨年版で使われたロボットの導入数やAIカンファレンスへの出席者数など、多くの項目に国単位の粒度でみる指標が加えられた。さらに、2018年度版ではAI関連特許やロボットオペレーティングシステム（ROS）のダウンロード数、GLUE metric、COCO leaderboardなどの新たな指標も加わっている。
昨年と同様に、AI分野における活動はあらゆるところで活発化しており、AIの性能は全面的に向上し続けていると言えるが、他言語翻訳技術の驚異的な発展やAI関連講座受講者数における性別的な多様性の欠如は特筆に値する。

<u>その他の指標</u>

昨年と同様に、the ***Derivative Measures*** セクションでは潮流間の関係性を調査している。また、AI分野での活気を定量化するため、学問や産業の動向を組み合わせた、AI Vibrancy Indexという探索的指標を示す。

本レポートでは、政府の取り組みに関する指標、**Recent Government Initiatives**を新たに追加した。本項目では、アメリカ、中国、ヨーロッパにおける政府の人工知能分野に対する投資と発表をまとめた。来年以降、AI Indexでは、より多くの国々の取り組みに関するデータと分析をしていきたいと考えている。

Human-Level Performance Milestonesのセクションでは、AIが人のレベルに達した、あるいは人を超える能力を示した事例を時系列順に表す。2018年は4つの新たな成果が付け加えられた。

AI Index Report Overview (続)

What's Missing のセクションでは、来年以降の AI Indexがより包括的なレポートとなるために必要な指標を各分野の専門家からいただいたフィードバックをもとにまとめた。

APPENDIX

Appendixでは、完全な透明性のもと、読者に対してソース、手法、ニュアンスに関する解説をまとめている。また、AppendixにはAI Indexに記されているおよそ全てのグラフの基礎データを格納した。
AIコミュニティの各人がそれぞれの分野におけるデータや指標をもとに交流していくことを願う。

SYMBOLS

本レポートでは、AIの普遍性に関するデータを示す場合、地球型のシンボルを目印として設置している。各国の比較や、アメリカ国外の各地域に関して掘り下げた調査、 AI関連の共同体の多様性などを普遍性に関するデータとして扱う。

VOLUME OF ACTIVITY

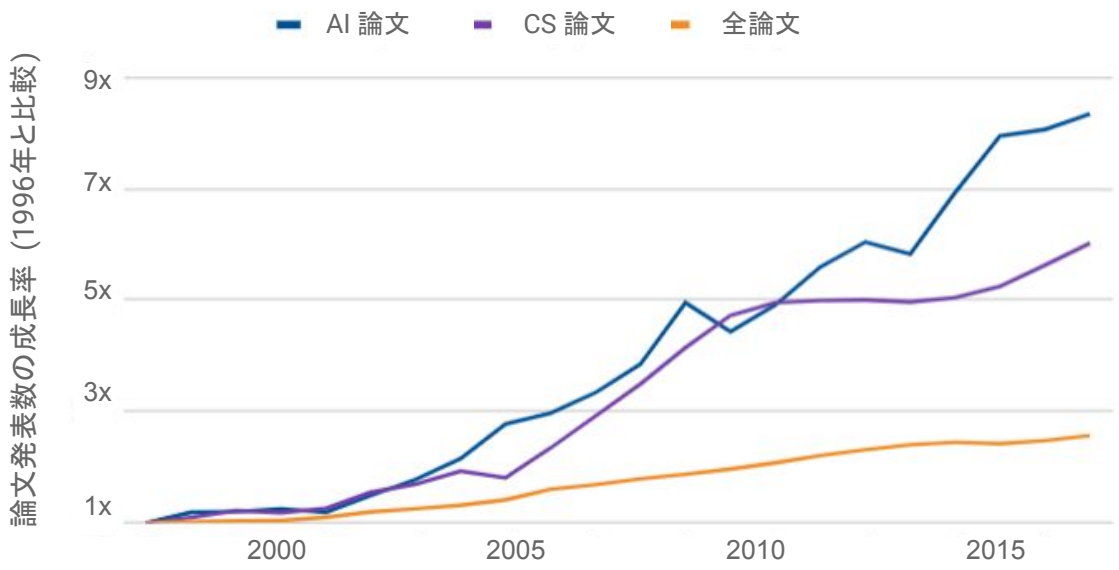
発表論文：研究内容ごとの論文増加率

下記のグラフは、1996年の年間学術論文発表数と2017年までの各年間学術論文発表数を比較し、相対的増加率を表したものである。グラフでは、全分野の論文数、コンピュータサイエンス（CS）分野の論文数、人工知能（ AI）分野の年間論文数の成長率を比較している。

人工知能分野における年間論文発表数は増加を続け、コンピュータサイエンス分野の年間論文発表数の増加率を上回っている。
このことから、AI分野での論文発表数の増加は、 Computer Science分野に対する関心の高まりによるものだけではないことがわかる。

データと手法の詳細は Appendix1に記載。

各分野における論文発表数の年間成長率 (1996−2017)
ソース: Scopus



備考: AI関連の論文数の増加率はElsview Keyword Approachではなく、Scopusで"Artificial Intelligence"というクエリ検索によって算出している。詳細はAppendixに記載。

AI outpaces CS

AI papers on Scopus have increased 8x since 1996.
CS papers increased 6x during the same timeframe.

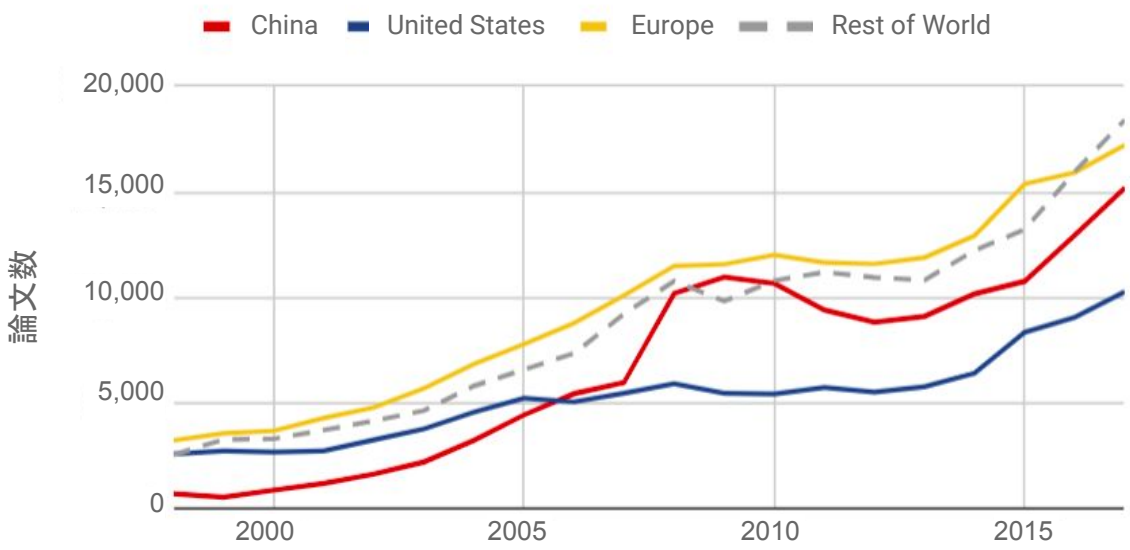
発表論文：研究内容ごとの論文増加率

Elsvierが発表したAIに関する研究で使われた手法と、Scopus上のデータに基づき、各地域の年間論文発表数を算出した。

AI関連論文の発表数では、ヨーロッパが継続的に最多論文発表数を誇っており、 2017年にScopusへ登録されたAI論文の28%がヨーロッパで記されたものである。
一方、中国では2008年前後の急激な論文数の増加と減少があったにも関わらず、 2017年に発表されたAI関連論文数は、2007年から1.5倍に増加している。

データと手法の詳細は Appendix2に記載。

Scopus上における各地域ごとの年間AI関連論文発表数 (1998−2017)
ソース: Elsevier



備考: 2008年前後の中国における AI関連論文発表数の急激な増加は、国家中長期科学技術発展計画 などの政府計画による AI研究への投資と奨励制度の充実に起因するものだと推測している。 2008年から2010年のヨーロッパにおける FP7などの研究計画も、中国の政府計画と同様にヨーロッパの論文発表数に対して上昇効果をもたらした可能性がある。

Europe is the largest publisher of AI papers

In 2017, 28% of AI papers on Scopus were affiliated with European authors, followed by China (25%) and the U.S. (17%).

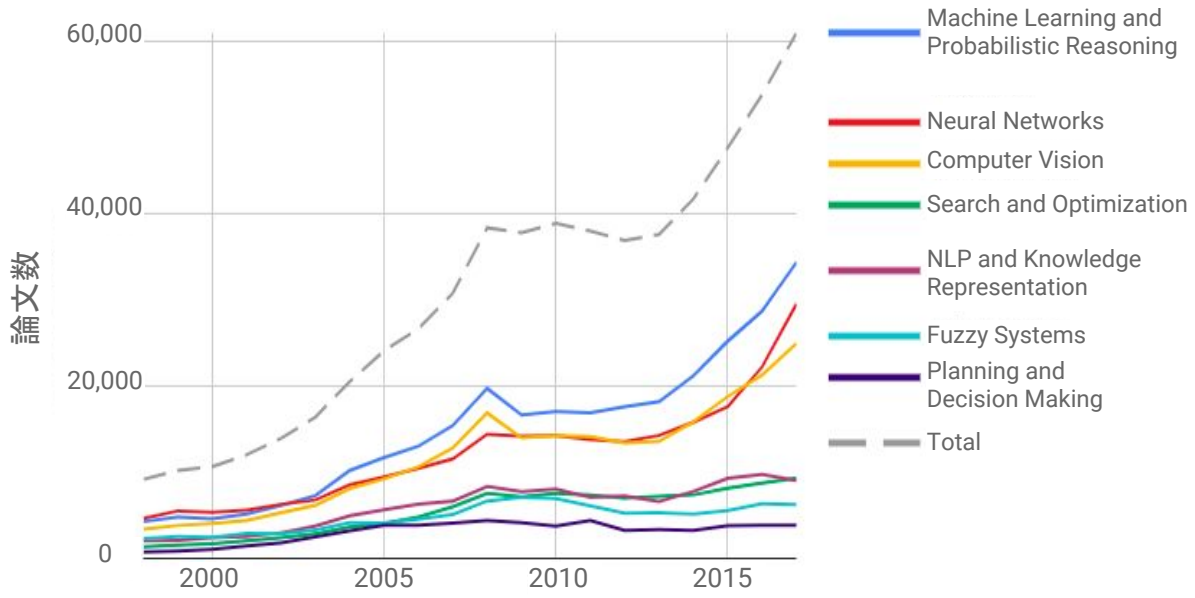
発表論文：サブカテゴリごとのAI関連論文数

サブカテゴリごとに見る Scopus上のAI関連論文数。
カテゴリの一部は重なり合っている。

2010年の論文では28%が機械学習と確率的推論に分類される中、2017年に発表された論文を振り分けた場合、56%が同サブカテゴリに分類される。2010年-2014年と2014年-2017年の期間を比較したとき、2014-2017年の期間において、図にまとめられたサブカテゴリの大半で論文発表のペースが早まっていることが明らかになった。特筆すべき点として、ニューラルネットワーク関連の論文発表数の年平均成長率が 2010年-2014年期間の3%から、2014年-2017年には37%へと大幅に上昇していることが挙げられる。

データと手法の詳細は Appendix2に記載。

Scopusにおける各AI関連サブカテゴリの論文数（1998−2017）
ソース: Elsevier



The number of Scopus papers on Neural Networks had a CAGR of 37% from 2014 to 2017

発表論文: arXiv上のAI関連論文

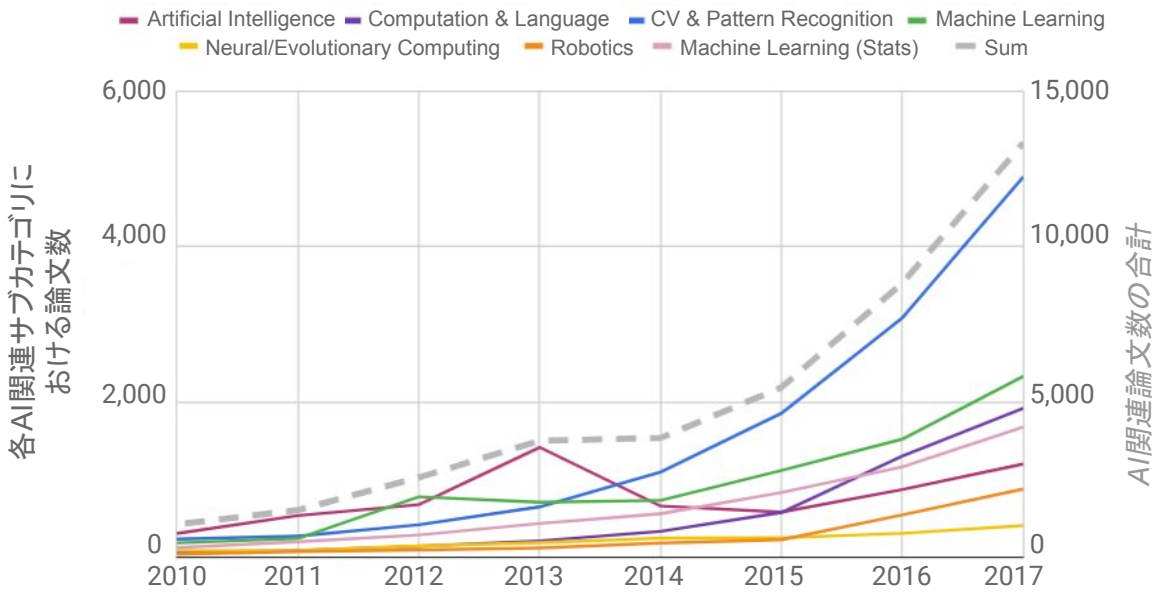
arXiv上に登録されているAI関連論文の主要サブカテゴリ分類。右軸は arXiv上の全AI論文の合計数を表す。論文数は累計で算出。

arXiv上のAI関連論文数は、多くのサブカテゴリで増加し続けている。この数値は、ピアレビュージャーナルか、あるいは著名な会議で採択されたものかに関わらず、 AI論文著者が研究をどの分野で普及させようとしているのか表している。また、この数値は各分野の競争性も示している。

2014年以前、コンピュータビジョンとパターン認識は、人工知能あるいは機械学習といったサブカテゴリと近しい成長曲線を描いていたが、 2014年以降、arXiv上でのAI関連分野における最大のサブカテゴリとなっている。この現状は、コンピュータビジョン（およびその一般的なアプリケーション）への関心の高まりに加え、コンピュテーション、または言語、ロボティクスのようなAI導入分野の成長も示唆している。

データと手法の詳細は Appendix3に記載。

arXiv上の各AI関連サブカテゴリにおける論文数 (2010−2017)
ソース: arXiv



"...Aside from the increase in publications, it's important to note the adoption of arXiv by these communities for disseminating results. We've seen many times how establishing some critical mass then catalyzes ever higher levels of participation within a community..."
Paul Ginsparg, Cornell

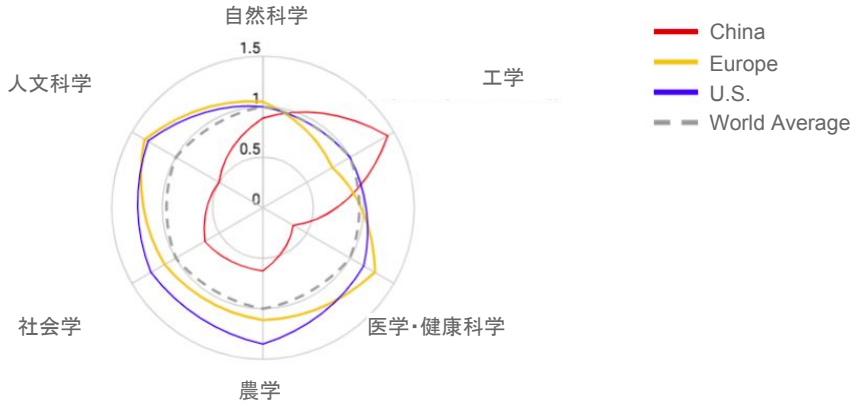
発表論文: 各地域の専門領域

相対的活動指標は、世界的な AI分野の研究活動と比較し、特定地域の専門領域を求める。相対的活動指標の定義は、AI分野における出版物の世界シェアと比較した特定国の出版物のシェアとする。
相対的活動指標において、1.0の値は、特定国の AI研究活動が世界の AI研究活動と同程度の水準で行われているということになる。 1.0より大きな値は、特定国がその分野を重視していることを意味し、1.0未満の場合はその分野への注力が弱いことを示している。

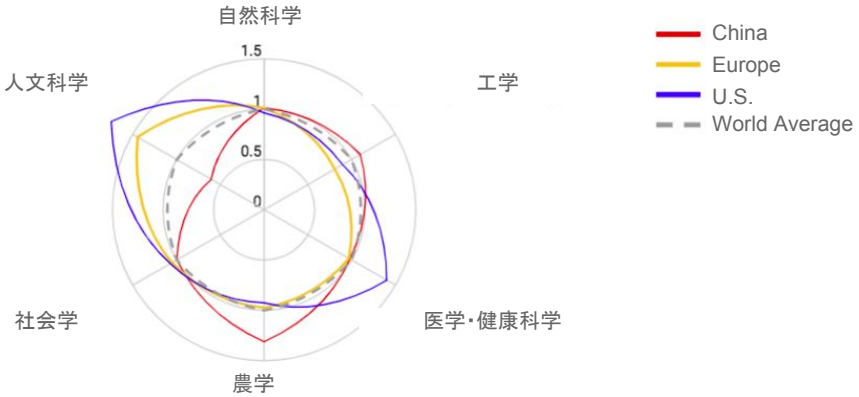
中国は、人文科学や医学、健康科学に力を入れているアメリカやヨーロッパと比べ、工学と農学に注力している。2000年のデータと比較すると、2017年のデータからは 3地域を通じて、専門化が進んでいることと、中国が農業重視の方向性に転換していることがわかる。中国が世界最大の食物生産国であり、AIの応用に焦点を当てているため、上記の結果は我々の予想通りだと言える。

データと手法の詳細は Appendix2に記載。

各地域の相対的専門領域とAIの研究分野 (2000)
ソース: Elsevier



各地域の相対的専門領域とAIの研究分野 (2017)
ソース: Elsevier



発表論文: セクターごとの論文数

Scopusに登録されている政府、企業、医療団体と連携している論文数のデータ。中国のデータは本ページ、ヨーロッパとアメリカのデータは次ページに記載している。
16ページでは3地域を直接比較し、企業と連携した論文数のグラフと政府と連携した論文数のグラフを作成した。

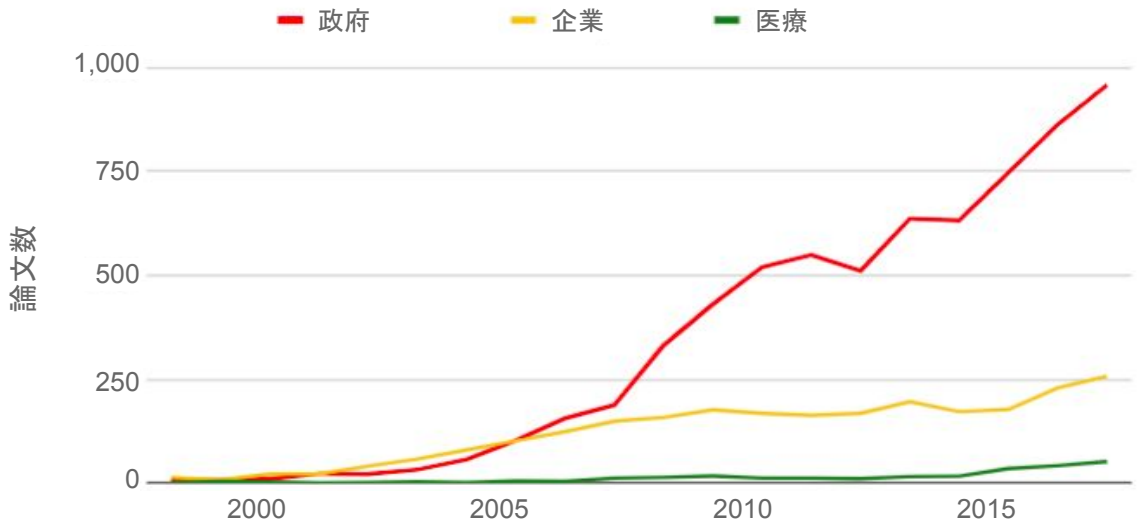
2017年中国政府は中国系企業の 4倍ものAI関連論文を作成。さらに、中国政府が 2007年に発表したAI関連論文数の400%に当たる論文を2017年に発表している。一方、同期間中に中国系企業の発表した AI関連の論文数は、2007年から73%の増加にとどまっている。

中国やヨーロッパとは相対的に、アメリカでは企業による論文が大きな割合を占めている。2017年、アメリカの企業から発表された AI関連論文数は中国の 6.6倍、ヨーロッパの 4.1倍だった。

3つの地域すべてにおいて、学術論文数（図示せず）が政府、企業、および医学論文を大幅に上回っている。学術論文数は各図の下にあるコメントに記載。

データと手法の詳細は Appendix2に記載。

連携セクターごとの論文数 – 中国 (1998–2017)
ソース: Elsevier



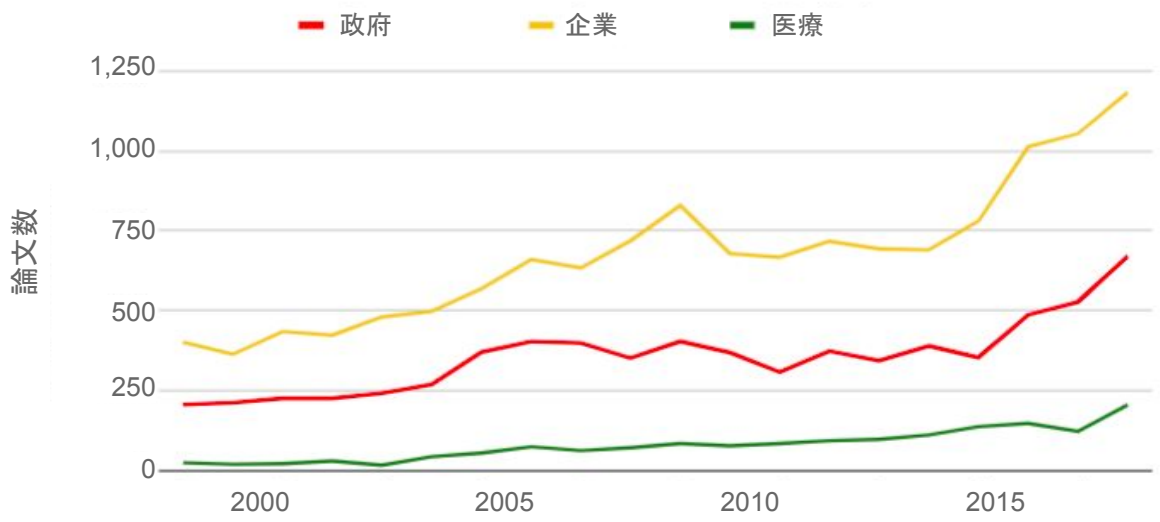
備考: 2017年、中国の学術部門と連携して発表されたAI関連論文の数（図示せず）は、約14,000本、つまり中国の全AI関連論文の約92%だった。

Government-affiliated AI papers in China have increased 400%

From 2007–2017, government AI papers in China increase 400% while corporate AI papers in China increased by 73%.

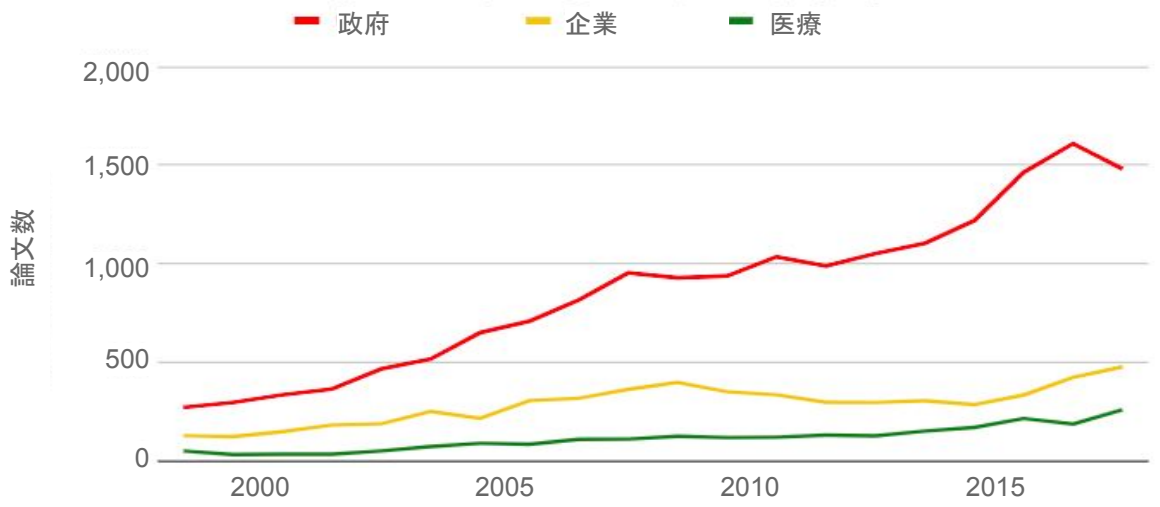
発表論文: セクターごとの論文数

連携セクターごとの論文数 – アメリカ (1998–2017)
ソース: Elsevier



備考: 2017年、アメリカの学術部門と連携して発表されたAI関連論文の数（図示せず）は約9,000本、つまり、アメリカ全AI関連論文のうち約81%だった。

連携セクターごとの論文数 – ヨーロッパ (1998–2017)
ソース: Elsevier



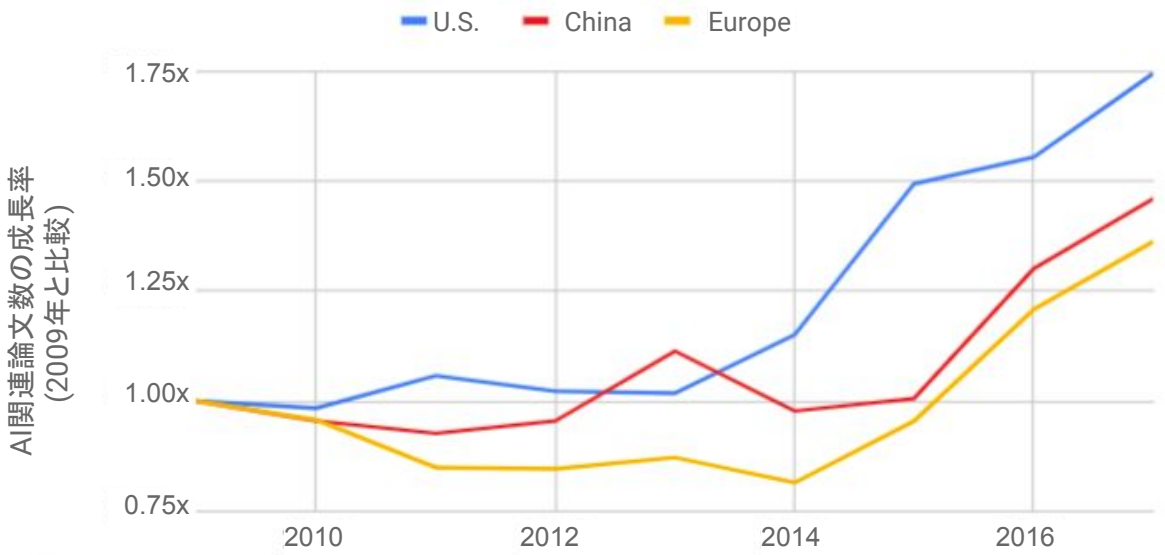
備考: 備考: 2017年、ヨーロッパの学術部門と連携して発表されたAI関連論文の数（図示せず）は約16,000本、つまり、ヨーロッパの全AI関連論文のうち約88%だった。

The proportion of corporate papers in the U.S. is 6.6x greater than that in China

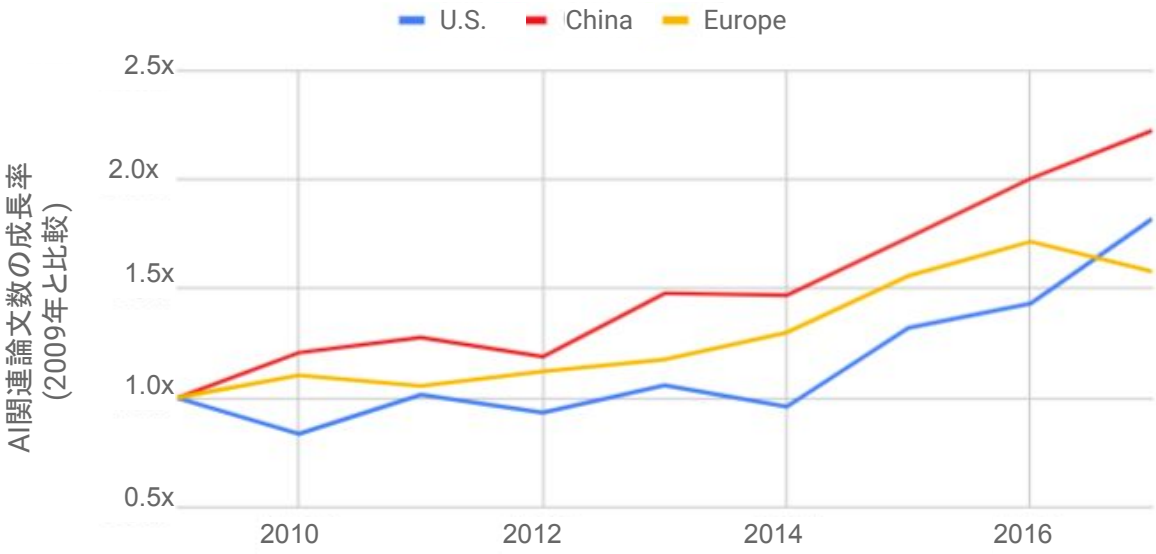
In 2017, the proportion of corporate AI papers in the U.S. was 6.6x greater than the proportion of corporate AI papers in China.

発表論文: Sector focus (continued)

企業と連携したAI関連論文数の成長率 (2009−2017)
ソース: Elsevier



政府と連携したAI関連論文の成長率 (2009−2017)
ソース: Elsevier



The number of Chinese government-affiliated AI papers has more than doubled since 2009

Meanwhile, the U.S. shows the greatest increase in corporate-affiliated AI papers. There were 1.7x as many corporate AI papers in 2017 as there were in 2009.

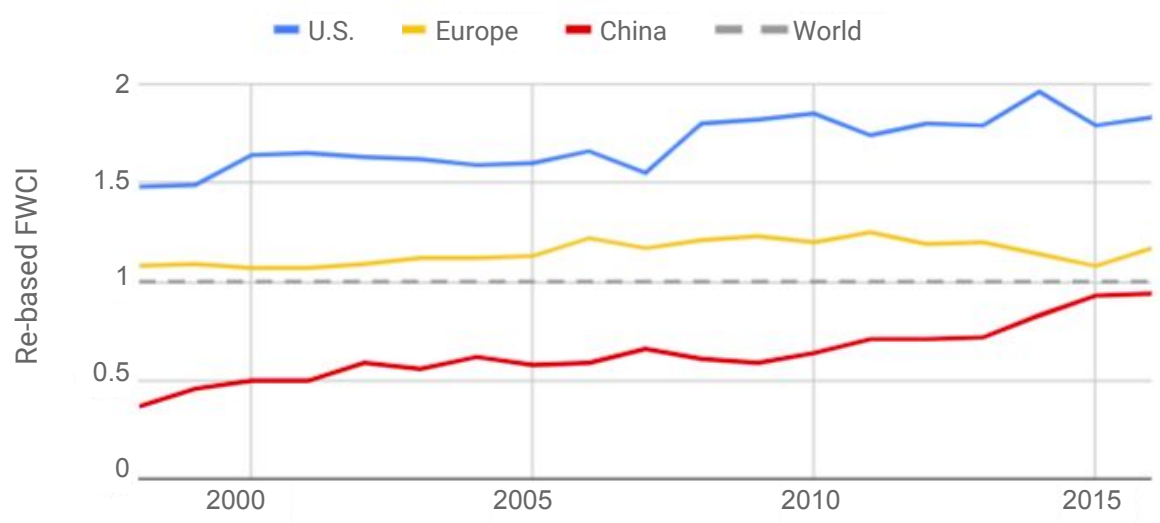
発表論文: Citation impact by region

地域ごとのField-Weighted Citation Impact（FWCI）。
地域ごとのFWCIは各地域のAI論文著者の引用された回数平均を世界全体の AI論文著者が受けた引用回数の平均で割ることで算出される。
この図では、各地の平均を世界平均と比較し、数値を算出している。 1.0の値は、世界平均と同数の引用回数で、0.85は、世界平均に比べ、15%引用回数が少ないことになる。

ヨーロッパは、世界最大の AI関連論文発表数を誇る一方、世界平均と比較し、比較的平坦なFWCIの値を保っていることがわかる。対照的に、中国では FWCIの値が向上している。中国のAI論文が引用される回数は、2000年から2016年の間で44%向上した。しかし、他地域と比べ、アメリカの論文引用回数が依然として高いままである。

データと手法の詳細は Appendix2に記載。

各地域のAI関連論文著者の Field-Weighted Citation Impact (1998—2016)
ソース: Elsevier



U.S. AI authors are cited 83% more than the global average.

発表論文: Author mobility by region

次ページの図は、AI論文の著者の生産性と引用の影響力に対し、国際的な流動性がもたらす効果を示す。

本項目では、論文著者の流動性を *Sedentary*、*Transitory*、*Migratory Inflow*、*Migratory Outflow*の4グループに分けて見ていく。本レポートにおいて *Sedentary*は、自国を含む地域において活動的に論文発表を重ねる研究者を指す。 *Transitory*は、2年以下、自国を含む地域外において論文を発表し帰国した研究者を指す。 *Migratory inflow*は自国外において、2年以上論文発表したうえで自国へ戻った研究者を指す。 *Migratory Outflow*は自国・地域外から2年以上論文を発表し、そのまま自国外にとどまっている研究者を指す。
本項目では*Migratory Outflow*のみが自国・地域外で活動を続けるグループとなっている。

次ページの図は、X軸は各地域、各グループにおける *著者1人当たりの平均論文発表数を相対的な割合で示したもの* となっている。Y軸は各地域におけるグループの *FWCI*(論文の被引用数を、同じ出版年・分野・文献種の論文の被引用数と照らし合わせ、基準値化した指標)を示している。

3地域全てにおいて、Sedentaryグループの平均論文発表数が最低の割合となっている。加えて、3地域総じて、Migratory inflow, outflowグループが最高の FWCIを記録している。このことから、研究拠点の移動が多い論文著者は、被引用数、論文発表数ともに、移動の少ないグループと比べ、高い傾向にあることがわかる。

アメリカ、中国、ヨーロッパの 3地域では、中国は76%のAI関連論文の著者が Sedantaryグループに属し、3地域間で最も大きな Sedantaryグループを有している。Sedantaryグループの割合はヨーロッパで 52%、アメリカでは38%である。Sedentaryグループの割合が大きい中国だが、中国国内における他グループの論文著者は、アメリカ、ヨーロッパに比べ、論文発表率が高い傾向にある。つまり、中国は地理的な流動性の高い研究者が比較的少ないが、アメリカ、ヨーロッパの 2地域で研究を行う流動性の高い研究者と比べてもより生産性が高いということである。

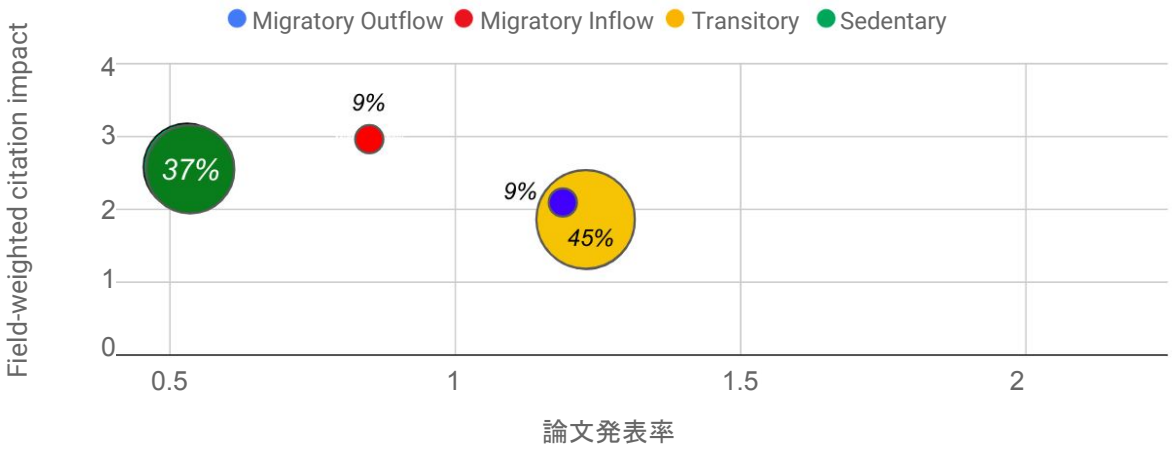
データと手法の詳細は Appendix 2に記載。

Mobile authors have a greater number of citations and publish more frequently

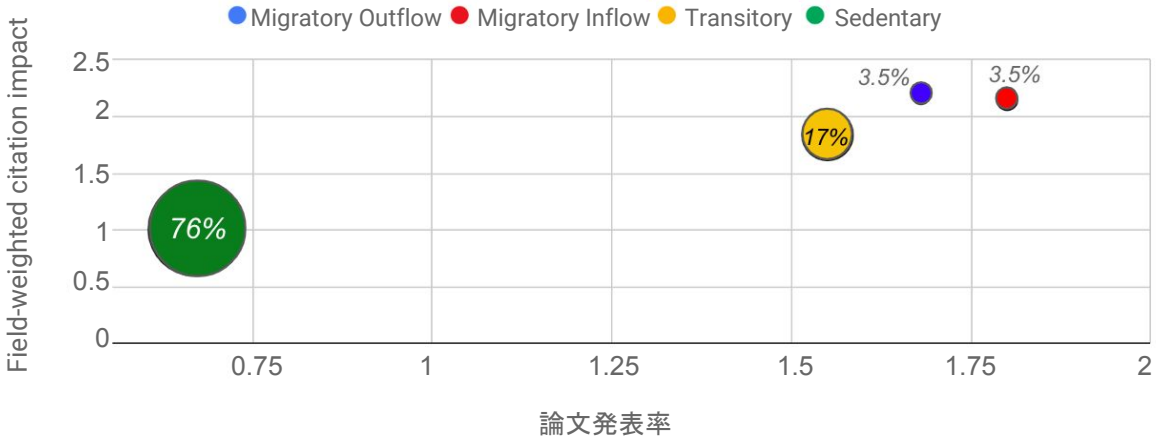
発表論文: Author mobility by region

解説は前のページを参照。

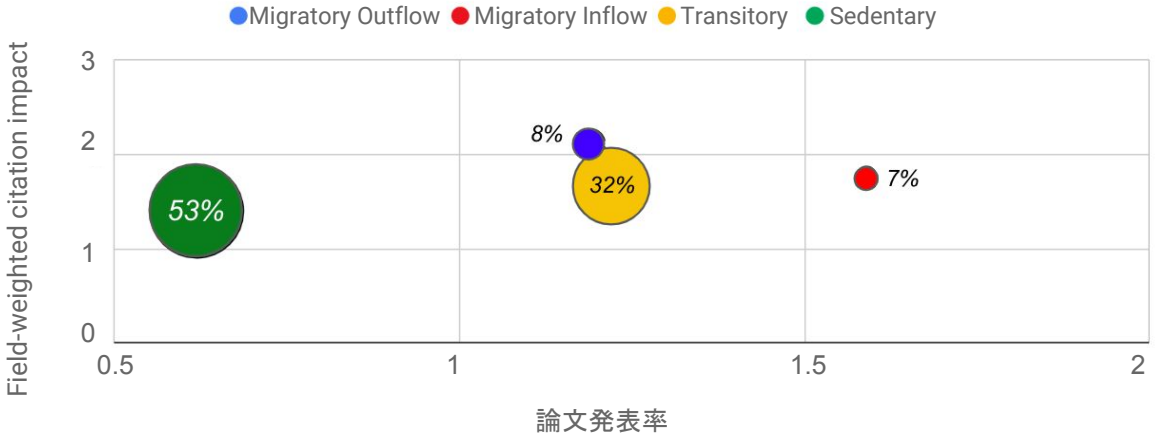
各流動性における論文発表率とCitation Impact — アメリカ (1998–2017)
ソース: Elsevier



各流動性における論文発表率とCitation Impact — 中国 (1998–2017)
ソース: Elsevier



各流動性における論文発表率とCitation Impact — Europe (1998–2017)
ソース: Elsevier



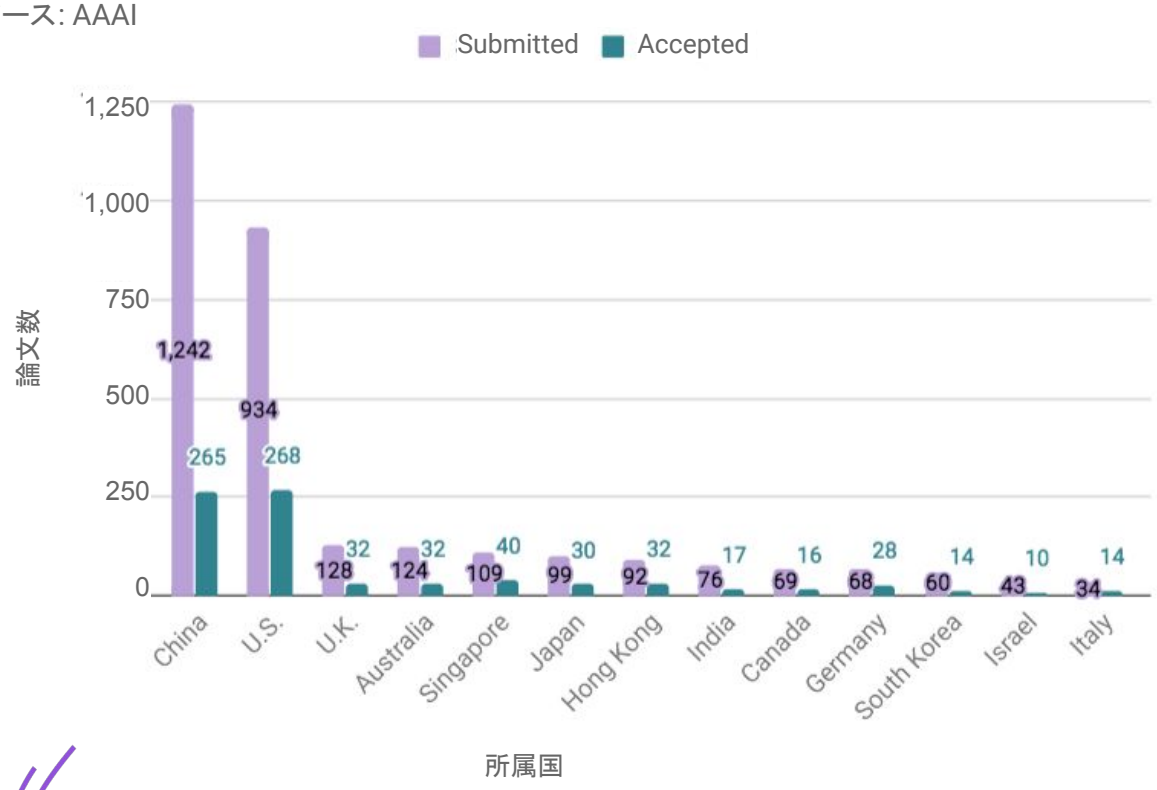
発表論文: AAAI papers by country

2018年2月、ルイジアナ州ニューオーリンズにて開催された AAAIカンファレンスに提出され、採択された各国ごとの論文数をまとめた。

AAAIに提出された論文の約 70%がアメリカ、または中国に連なる論文だった。論文提出数では中国が最多だが、論文採択数に関しては、アメリカから 268本、中国からは265本とほぼ同数の論文が採択されている。2018年、アメリカと中国、それぞれ AAAIの論文採択率は29%（アメリカ）、21%であった。ドイツ、イタリアから提出された論文に関して、提出本数は少ないものの、41%という最高の採択率を誇った。

データと手法の詳細は Appendix4に記載。カンファレンス出席者数の経時的なデータは 26 ページを参照。

提出された論文数と採択された論文数 - 2018 AAAIカンファレンス
ソース: AAAI



70% of AAAI papers are from the U.S. or China

At the 2018 AAAI conference, 70% of submitted papers and 67% of accepted papers were affiliated with the U.S. or China.

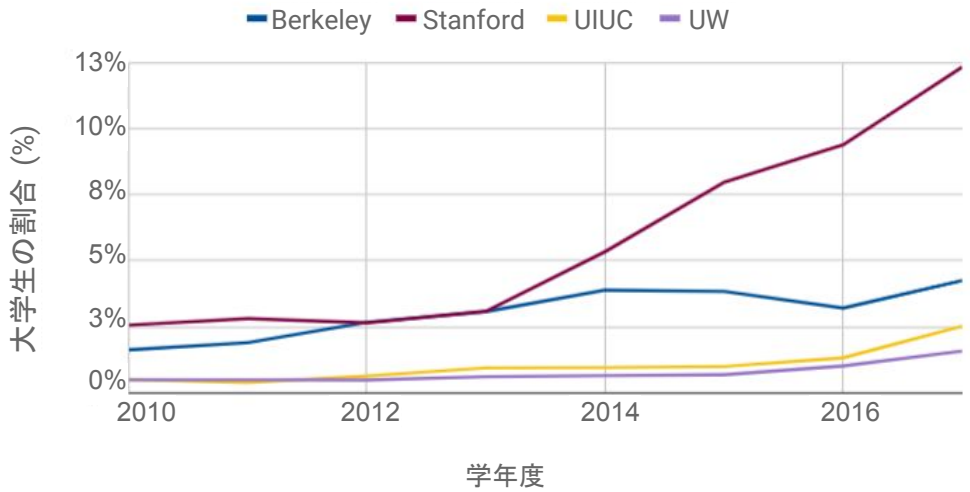
講座受講者：アメリカの大学におけるAI関連講座

AI・機械学習入門講座を受講した大学生の割合。この指標は大学の規模を考慮したもので、本指標における大学の選抜基準、実際の受講者数、および大学の正式名称は Appendixに記載している。

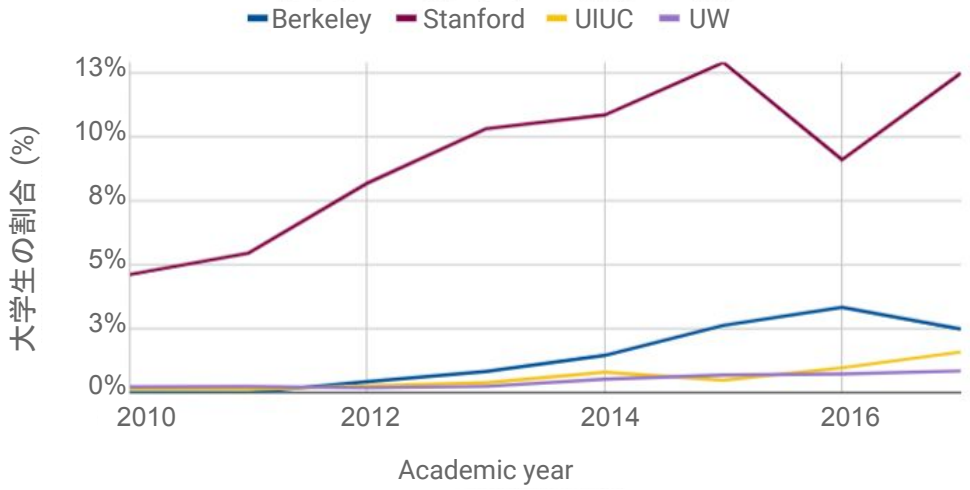
AI入門講座の受講者割合は、機械学習入門の受講者割合よりもわずかに高い傾向（ AI入門講座：平均5.2%vs機械学習入門講座：平均 4.4%）にあるが、機械学習入門講座は受講者数の伸び率がAI入門講座より比較的高くなっている。この結果は、 AIの派生分野として、機械学習の重要性が増していることを示唆している。成長比較は次ページを参照。

データと手法の詳細は Appendix5に記載。

AI入門講座を受講した大学生の割合 (2010–2017)
ソース: University provided data



機械学習入門講座を受講した大学生の割合 (2010–2017)
ソース: University provided data



性別の多様性に関する備考:
17年度 Stanford AI入門講座: 74.45% 男性 (合計869名); 17年度 Berkeley Intro to AI: 73.37% 男性 (合計973名)
17年度 Stanford 機械学習入門: 75.91% 男性 (合計876名); Berkeley '17 Intro to ML: 78.67% 男性 (合計720名)

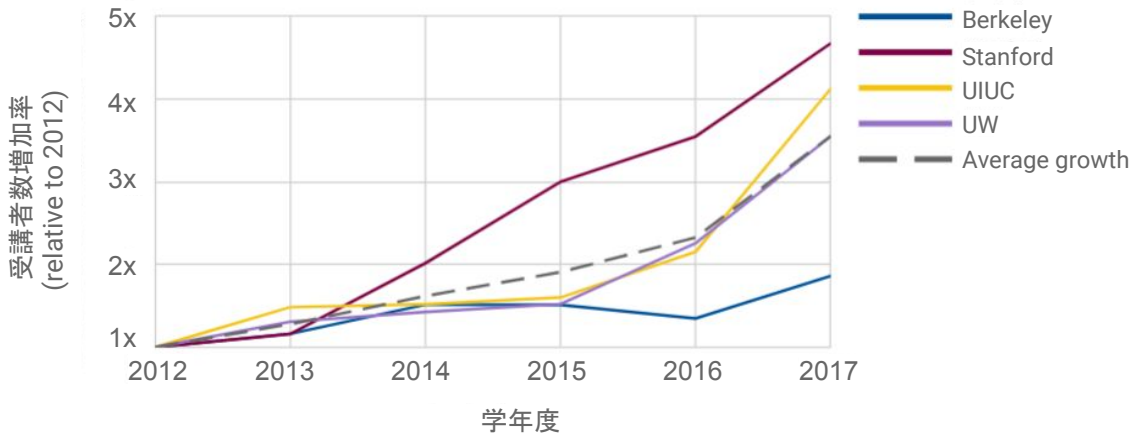
講座受講者：アメリカの大学におけるAI関連講座

コンピュータサイエンス分野でアメリカを牽引する大学における AI、機械学習講座受講登録者数の増加。次の大学からのデータを使用し、下記の要点をまとめた： UC Berkeley, Stanford, UI Urbana-Champaign, University of Washington, Carnegie Mellon University
大学の選抜基準、実際の受講者数、および大学の正式名称は Appendixに記載。

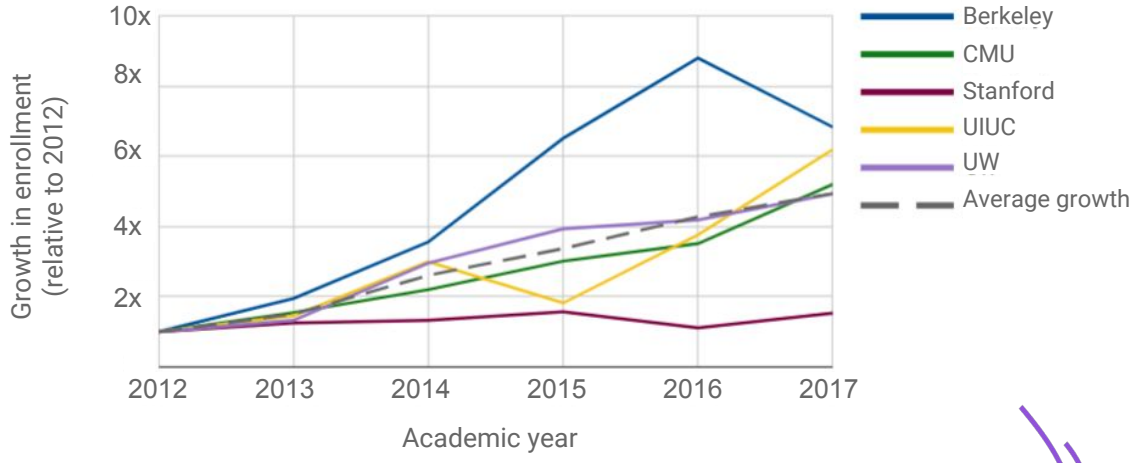
2017年におけるAI入門講座への登録者数は 2012年と比べ、3.4倍に増加している一方、機械学習入門講座への登録者数は、同期間中で 5倍に増加している。UC Berkeleyにおいては、2012年から2017年の間で、機械学習入門講座に登録した学生数は 6.8倍に増加している。この増加は、本レポートで調査した大学の機械学習入門講座の中で最大の増加率である。

データと手法の詳細は Appendix5に記載。

AI入門講座受講者数増加率 (2012–2017)
ソース: University provided data



機械学習入門講座受講者数増加率 (2012–2017)
ソース: University provided data



2017 introductory ML course enrollment was 5x greater than it was in 2012

講座受講者：国際的なデータ

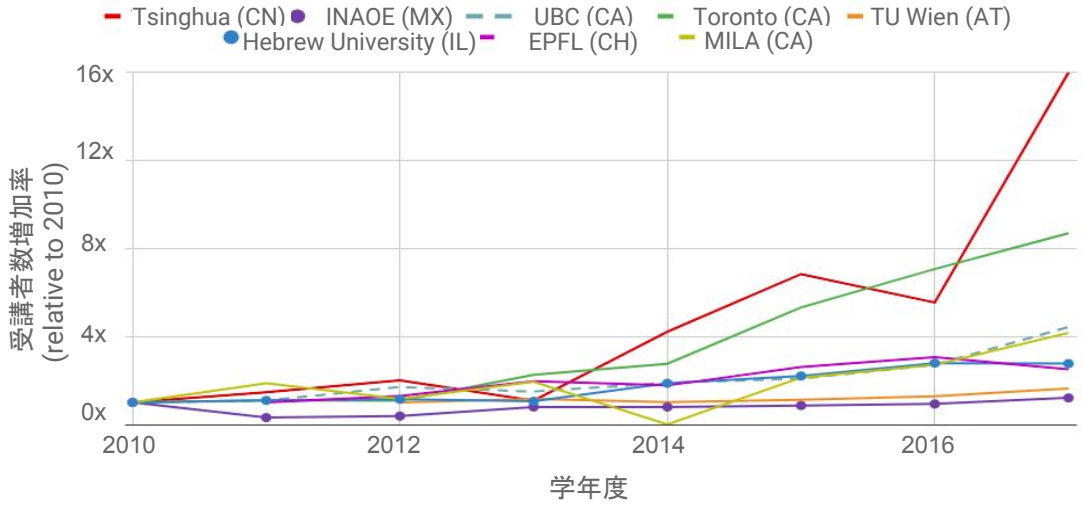
アメリカ国外の、コンピュータサイエンス分野を牽引する大学における AI・機械学習講座登録者数の増加。

グラフ1は、世界の大学における2010年から2017年のAI・機械学習講座登録者数の相対的成長率を表し、グラフ2では、2016年から2018年のAI・機械学習講座登録者数の相対的成長率のみを提供した大学をまとめている。大学の選抜基準、実際の受講者数、および大学の正式名称はAppendixに記載。

清華大学における2017年のAI・機械学習講座登録者数は 2010年と比べ、16倍に膨らんでいる。この大学調査から、AI講座登録者数の増加は大学によりばらつきがあることがわかった。また、地理的要因は、特段の影響を与えていない。今後の調査では、この仮説を立証していきたい。

データと手法の詳細は Appendix6に記載。

AI+機械学習講座受講者数増加率 ー アメリカ国外 (2010ー2017)
ソース: University provided data

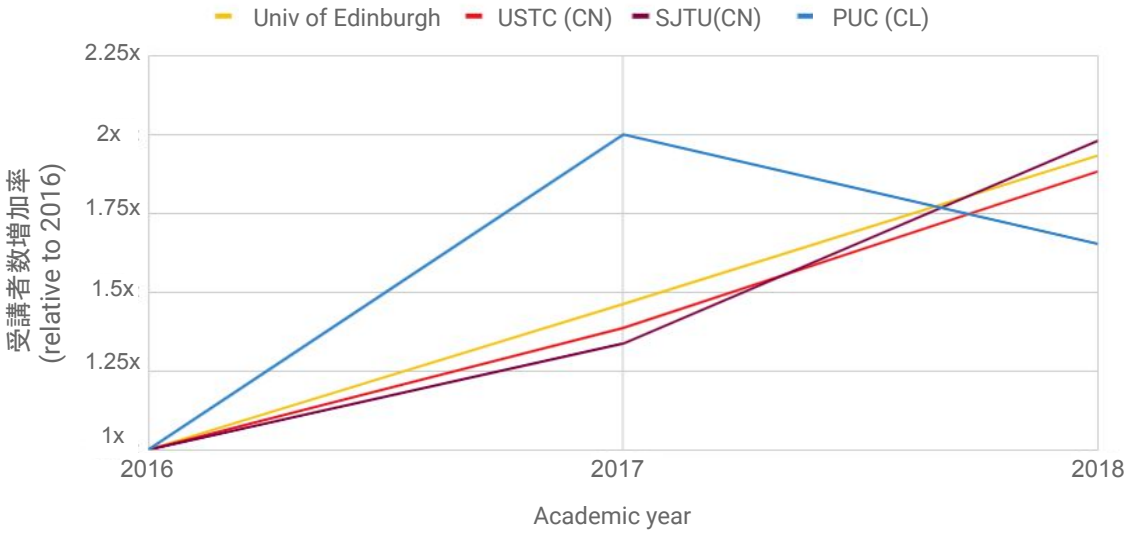


備考: 上記の図は過去のデータにアクセス可能な大学を示している。増加率は2010年時の講座受講者数その相関である。(EPFLは2011年度、TU Wienは2012年度)

講座受講者：国際的なデータ

前ページの解説を参照。

AI講座受講者数増加率 ー アメリカ国外 (2016ー2018)
ソース: University provided data



備考: 上記の図は過去データへのアクセスが限定されている大学を示している。増加率は2016年度の受講者数と相関関係にある。

教職員の多様性

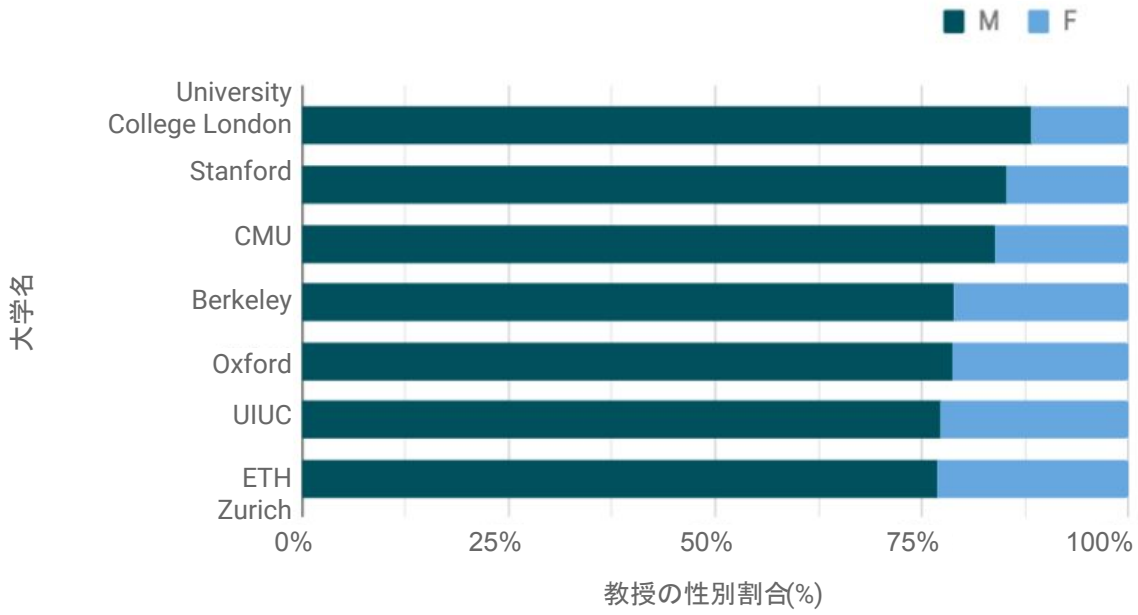
コンピュータサイエンス分野を牽引する大学における教授の性別内訳。データは　2018年9月21日時点の教職員名簿から取得したものである。大学選定は、コンピュータサイエンス分野を牽引する大学であること、教職員名簿へのアクセスが容易であることを基準に進めた。

多様性を改善する上での重大な障壁は、産業界および学界における多様性に関するアクセス可能な統計データの不足が挙げられる。
AI分野における多様性を改善させるために、多様性に関する統計の透明性を確保するよう、各機関に勧めていきたい。

この大学調査から、AI分野における教授の平均 80%が男性であることが判明した。この結果に関して、地理的要因による大きな差異はみられない。しかし、調査を行なった大学数が限られているため、全体像の一部だけを捉えたものだとも言える。

データと手法の詳細はAppendix 7に記載。

AI関連教授の性別内訳 ー 主要大学 (2018年9月)
ソース: University faculty rosters



80% of AI professors are male

On average, 80% of professors from UC Berkeley, Stanford, UIUC, CMU, UC London, Oxford, and ETH Zurich are male

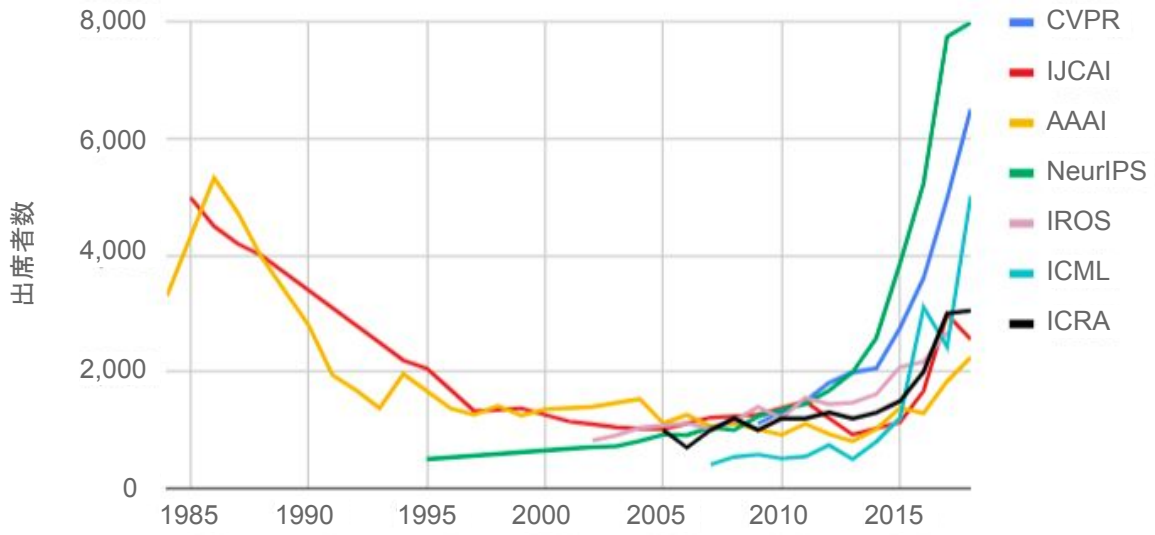
参加者数: 大規模AIカンファレンス

大規模AIカンファレンスの代表サンプルへの出席者数、および 2012年と比較した大規模カンファレンスへの出席者の増加率。大規模カンファレンスの定義として、 2017年に2,000人以上が参加した会議とする。

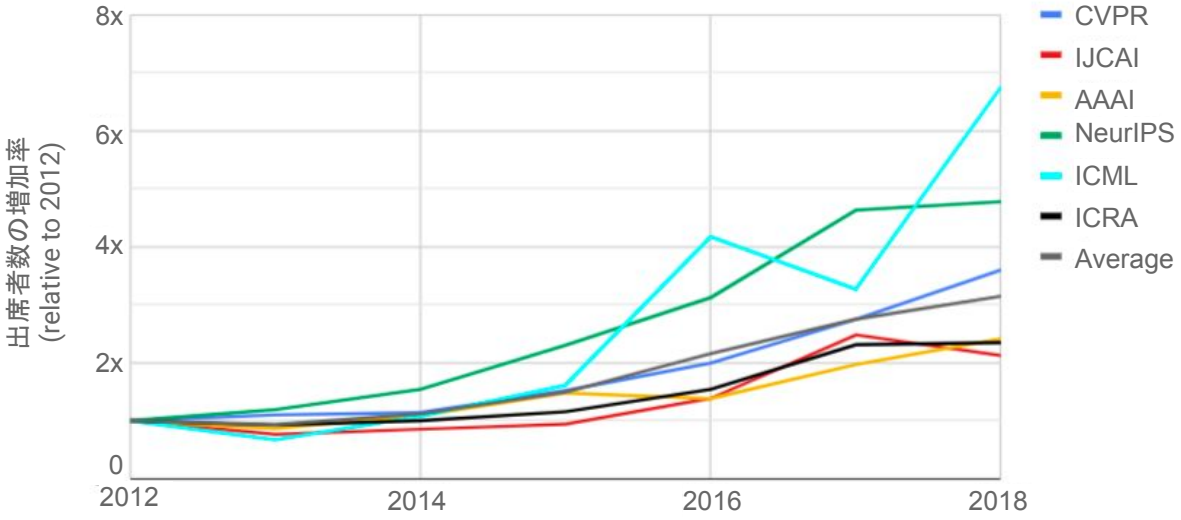
2018年、NeurIPSやCVPR、ICMLといった世界最大規模の AI研究カンファレンスは、2012年以降で最高の出席者数の増加を記録した。 2012年と比べ、NeurIPSへの出席者数は4.8倍、ICMLにおいては出席者数が 6.8倍に増加している。この指標は、AIに関する派生分野として、機械学習への関心が高まり続けていることを示している。一方、象徴的推論に焦点を当てたカンファレンスでは、相対的成長はほとんど見られない。

データと手法の詳細は Appendix 8に記載。

大規模カンファレンスへの出席者数 (1984–2018)
ソース: Conference provided data



大規模カンファレンスの出席者数増加率 (2012–2018)
ソース: Conference provided data



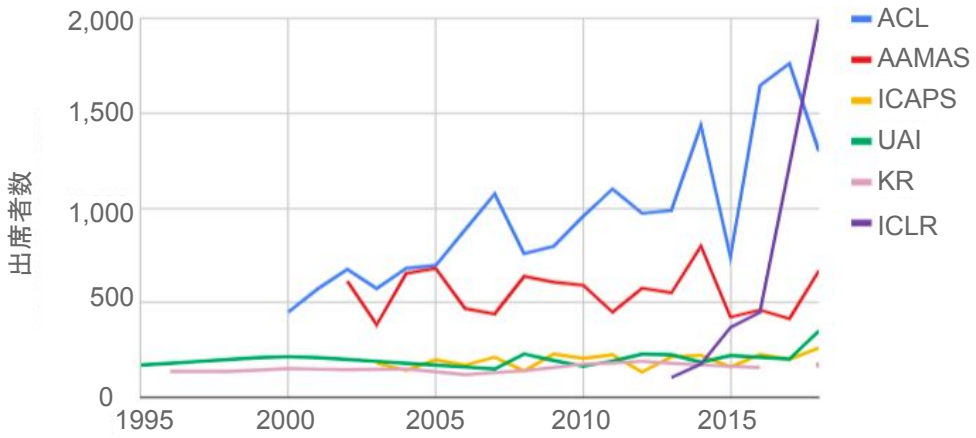
参加者数: 大規模AIカンファレンス

小規模AIカンファレンスの代表サンプルへの出席者数、および 2012年の出席者に対するカンファレンス出席者数の増加率。 小規模AIカンファレンスの定義は、2017年の参加者数が2,000人以下のものとする。

2018年、ICLRの出席率は2013年の20倍に増加している。この増加は、今日の AI分野で、深層学習と強化学習が特に注目されていることを示す結果と考えられる。

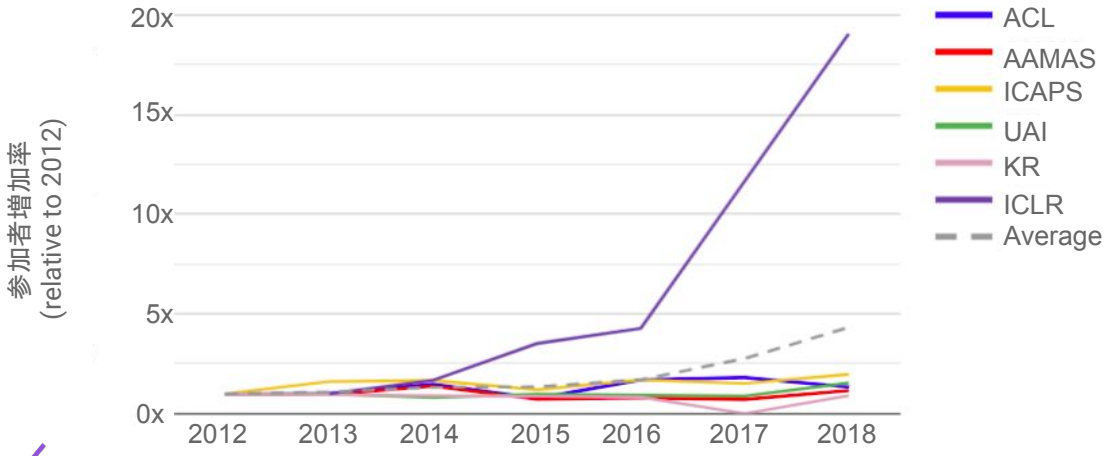
データと手法の詳細は[Appendix 8](#)に記載

小規模カンファレンスへの出席者数 (1995–2018)
ソース: Conference provided data



備考: 2018年、KRカンファレンスが初めてワークショップを開催した。一貫性を保持するため、ワークショップ参加者はKRカンファレンスとしてカウントしていない。

小規模カンファレンスへの出席者数増加率 (2012–2018)
ソース: Conference provided data



ICLR's 2018 conference attendance is 20x larger than it was in 2012

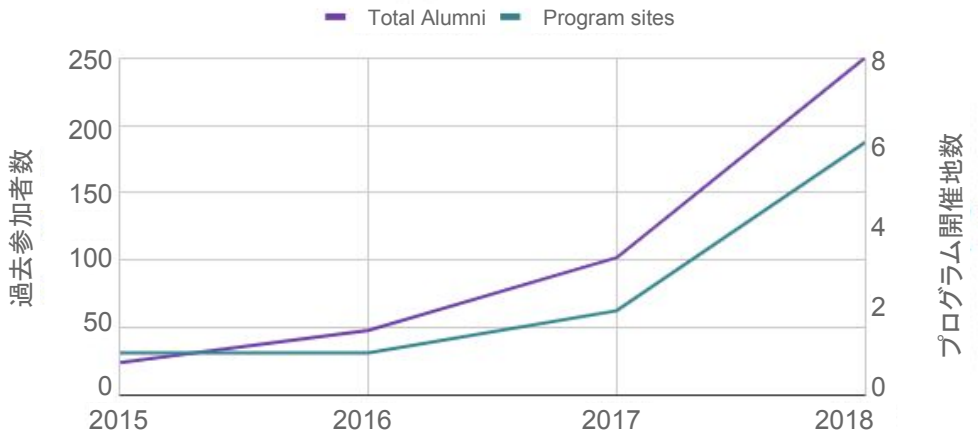
参加者数：多様性に関する団体

人工知能の多様性と包括性を高める AI教育へ取り組む団体、AI4Allへの過去参加者数と、機械学習の女性を支援する組織である Woman in Machine Learning（WiML）への登録者数。

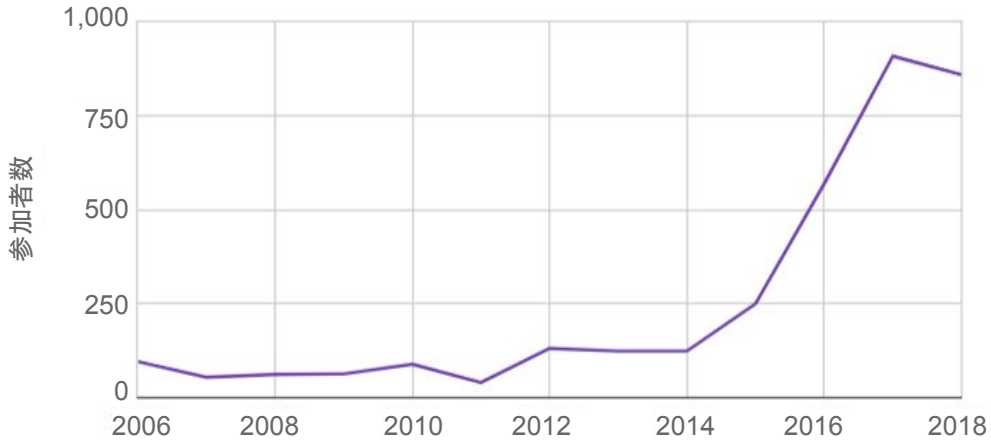
WiMLとAI4ALLの両方でエンゲージメントが大きく増加しており、特に AI4ALLは2015年のプログラム開始以来、同窓生の数が 900%以上に増加している。WiMLへの参加者は、2014年以来600%近く増加している。これらの数値は、AI分野に女性や少数グループを組み入れる取り組みが継続的に行われていることを示している。

データと手法の詳細は Appendix 8に記載。

AI4ALL 過去参加者数とプログラム開催地数 (2015−2018)
ソース: AI4ALL



WiML ワークショップ登録者数 (2006−2018)
ソース: WiML



備考: 2017年のWiML ワークショップへの登録者数は、1-dayから2-dayワークショップへ変更されたため、些か誇張した値となっている。

"...This data reflects the growth we've seen in the WiML community in recent years and, we think, correlates to increased attention to intentionally diversifying the communities within AI. But we have a long way to go, since attendance for WiML events still numbers in the hundreds, while overall attendance at commonly co-located conferences like NeurIPS is in the multiple thousands..."
WiML Board

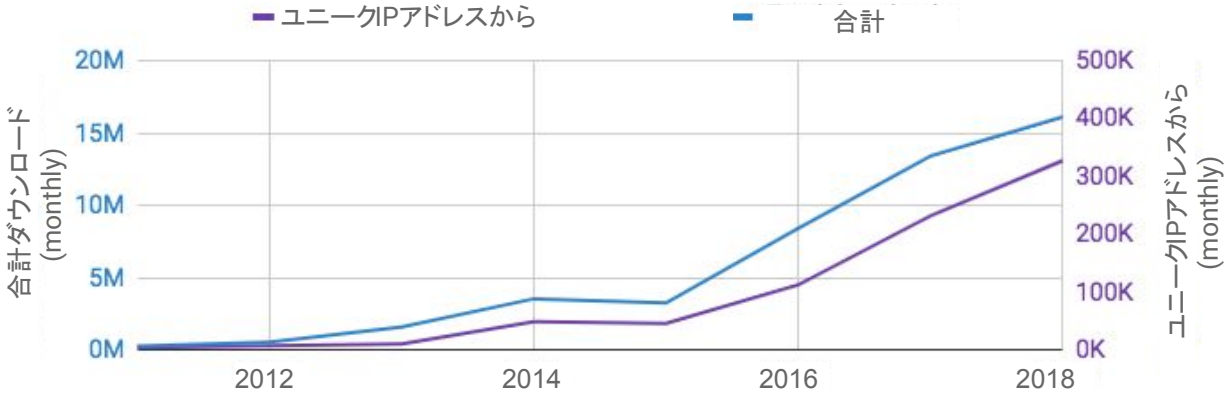
ロボットソフトウェア ダウンロード数

ロボットオペレーティングシステム（ ROS）はロボット工学分野で広く使用されているオープンソースのソフトウェアスタックである。 ROSは製造業者や研究者に広く利用されている。下記のデータは、各期間中に ROS.orgからダウンロードされたバイナリパッケージの数を示している。左軸は月間平均ダウンロード数の合計を示し、右軸はユニーク IPアドレスからの月間平均ダウンロード数を表している。

2014年以降、合計ダウンロード数とユニークダウンロード数は、それぞれ 352%、567%の増加を見せてきた。この数値は、ロボット工学への関心の高まりと、ロボットシステム利用の広がりを表している。ユニークダウンロード数は合計ダウンロード数の増加より早いペースで成長していることから、ROSのダウンロード頻度の増加ではなく、むしろユーザー数が増加していると断言できる。

データと手法の詳細はAppendix 9に記載。

ROS パッケージダウンロード数(2011−2018)
ソース: ROS metrics report



Since 2014, unique ROS package downloads from ROS.org have increased by 567%

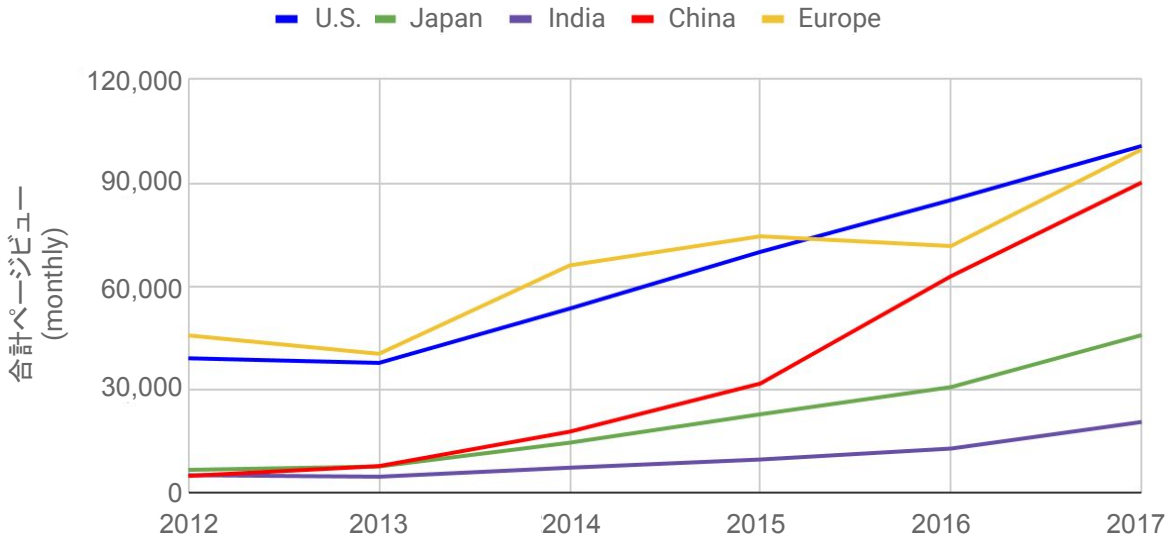
ロボットソフトウェア ダウンロード

2012年以降のROS.orgのPV数。下の図は、ページビュー上位 5地域を示している。

ROSページビューはアメリカとヨーロッパで、最も多くなっている。 2015年から2017年にかけ、中国は、アメリカやヨーロッパのほか、インドと日本を含めた 5地域の中で、過去最高の成長率を記録した。成長後の中国における ROSページビューは、アメリカ、ヨーロッパに引けを取らない規模となった。2017年における、中国からのページビューは、 2012年における同国からのページビューの 18倍まで拡大している。

データと手法の詳細は、Appendix 9に記載。

ROS.org ページビュー (2012–2017)
ソース: ROS metrics report



ROS.org pageviews in China were 18x greater in 2017 than in 2012

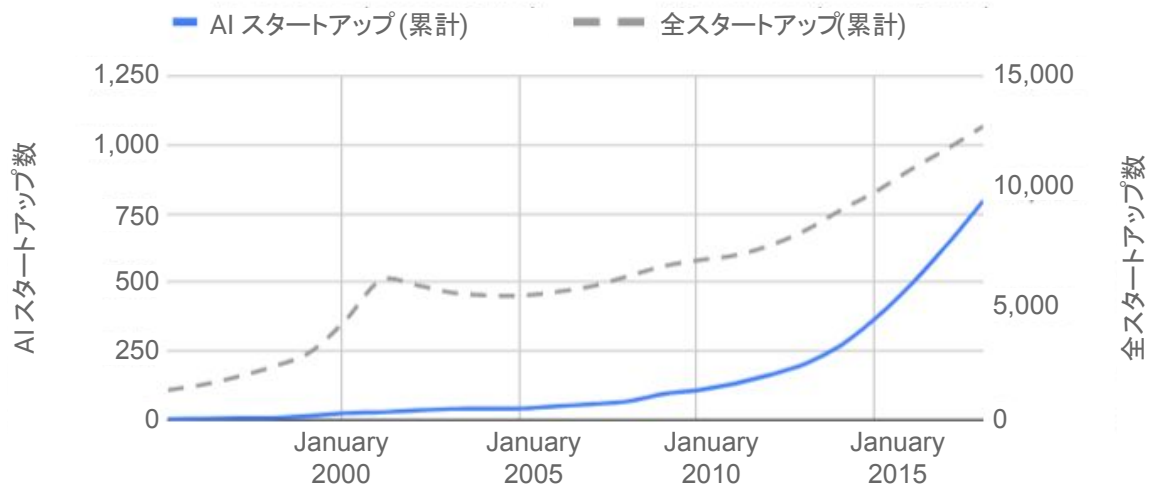
スタートアップ / 投資：AIスタートアップ

特定の年にベンチャーキャピタルからの支援を受けたアメリカの株式未公開スタートアップの数。青線（左軸）はAIスタートアップのみを示し、灰色の線（右軸）は AIを含むすべてのベンチャーキャピタルから支援を受けているスタートアップを示している。 このグラフは、毎年 1月のスタートアップの総数を整理したである。 スタートアップがデータセットから削除された例（詳細については付録を参照）を除いて、スタートアップの数は累計で年々増加している。

2015年1月から2018年1月の間で、スタートアップの総数が 1.3倍に増加した一方、AIスタートアップの数は2.1倍に増加した。ほぼ全ての期間において、スタートアップの総数は比較的安定な成長を記録しているが、AIスタートアップの数は指数関数的に増加している。

データと手法の詳細は Appendix10に記載。

AI スタートアップ (U.S., January 1995 — January 2018)
ソース: Sand Hill Econometrics



備考: 上記のAIスタートアップの大半がAIシステムを開発している。一部のスタートアップでは、AIをビジネスの機関部として活用しているが、システム自体は自社開発していない。Appendixに詳細を記載。

Active AI startups in the U.S. increased 2.1x from 2015 to 2018

Meanwhile, active startups as a whole increased 1.3x.

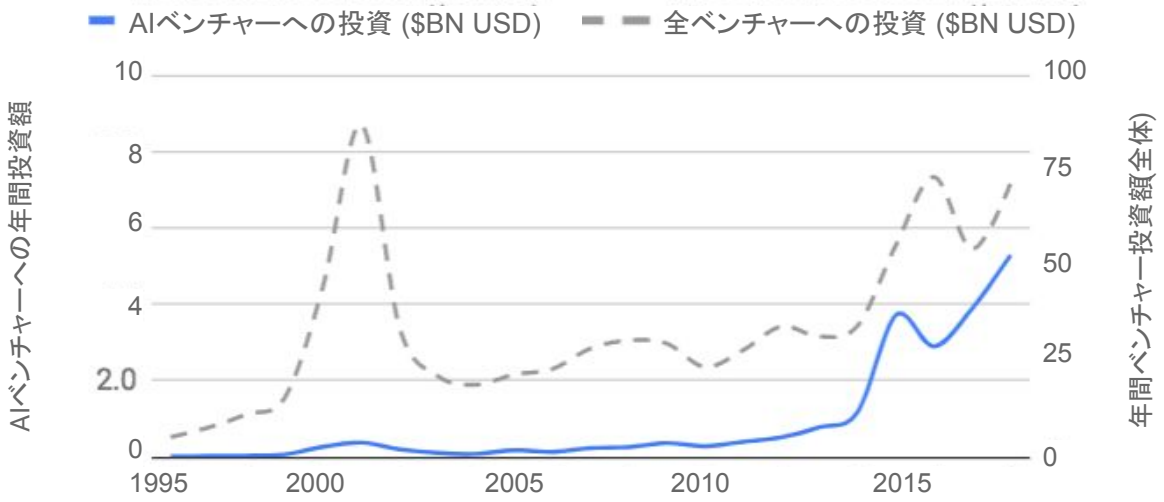
スタートアップ / 投資：ベンチャーキャピタル

ベンチャーキャピタルによるアメリカのスタートアップに対する年間投資総額。投資は、全ての資金調達段階のスタートアップを対象とする。青線（左軸）は AIスタートアップに対する投資のみを示し、灰色の線（右軸）はベンチャーキャピタルから支援を受けている AIスタートアップを含む全スタートアップに対する投資を示している。 以下のデータは、前ページのデータと同様、1年ごとの数値であり、累計の数値ではない。

2013年から2017年にかけて、AIスタートアップに対するベンチャーキャピタルからの投資額は4.5倍に増加しているのに対し、全スタートアップに対するベンチャーキャピタルからの投資額は2.08倍にとどまっている。1997年から2000年にかけてのベンチャーキャピタルから全スタートアップ企業へ注がれた資金額の急激な上昇は、ドットコムバブルの影響によるものである。2014年と2015年の投資額上昇は、2014年の比較的大きな経済成長を反映している。

手法については Appendix10に記載。

ベンチャーキャピタルによるAIスタートアップへの年間投資額 (U.S., 1995 – 2017)
ソース: Sand Hill Econometrics



VC funding for U.S. AI startups increased 4.5x from 2013 to 2017

Meanwhile, VC funding for all active startups increased 2.08x.

AI関連職：AI関連技能ごとの求人数

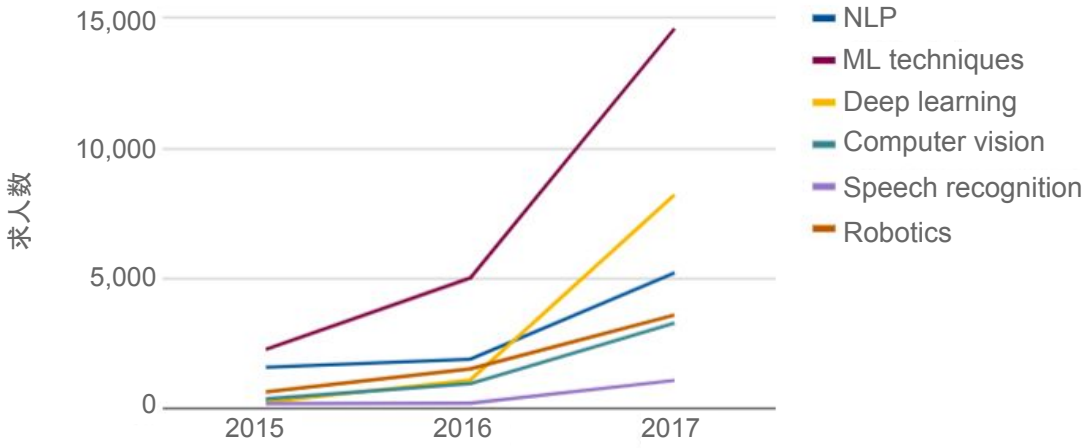
1年あたりの求められる専門性ごとの求人数と、求人数の相関的成長率。多くの求人が、複数の専門性を求めているため、求人数は二重に計測される場合もある。

就労要件として最も多い技能は機械学習だが、成長率で見た場合、深層学習を就労要件に含む求人が最高の成長率となっている。2015年の就労要件と2017年の就労要件を比較した時、深層学習が就労要件に含まれる求人数は 3年間で35倍に増加している。

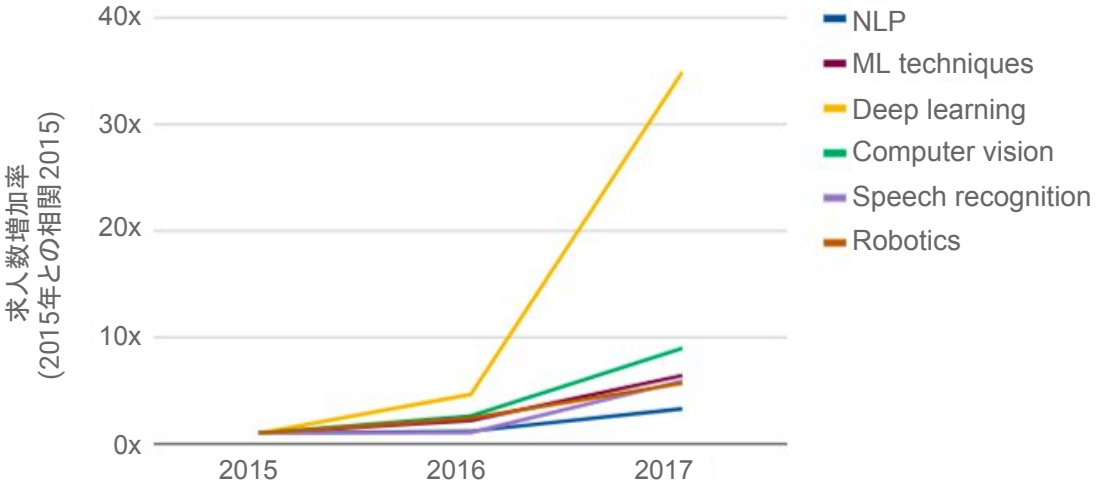
AI関連の求人は全面的に増加しているが、コンピュータエンジニアの求人における割合としては依然として少ない。2018年12月1日の、Monster.comにおける検索結果では、「 Java」のキーワードでは83,230件もの求人が見つかったが、「機械学習」で検索した場合だと、 20,114件しか見つからなかった。

手法の詳細はAppendix11に記載。

AI関連技能ごとに見る求人数 (2015 – 2017)
ソース: Monster.com



各AI関連技能における求人数増加率 (2015 – 2017)
ソース: Monster.com



備考: AI関連の求人数が増加している一方、コンピュータエンジニア全体における求人数としてはいまだに少数のままである。

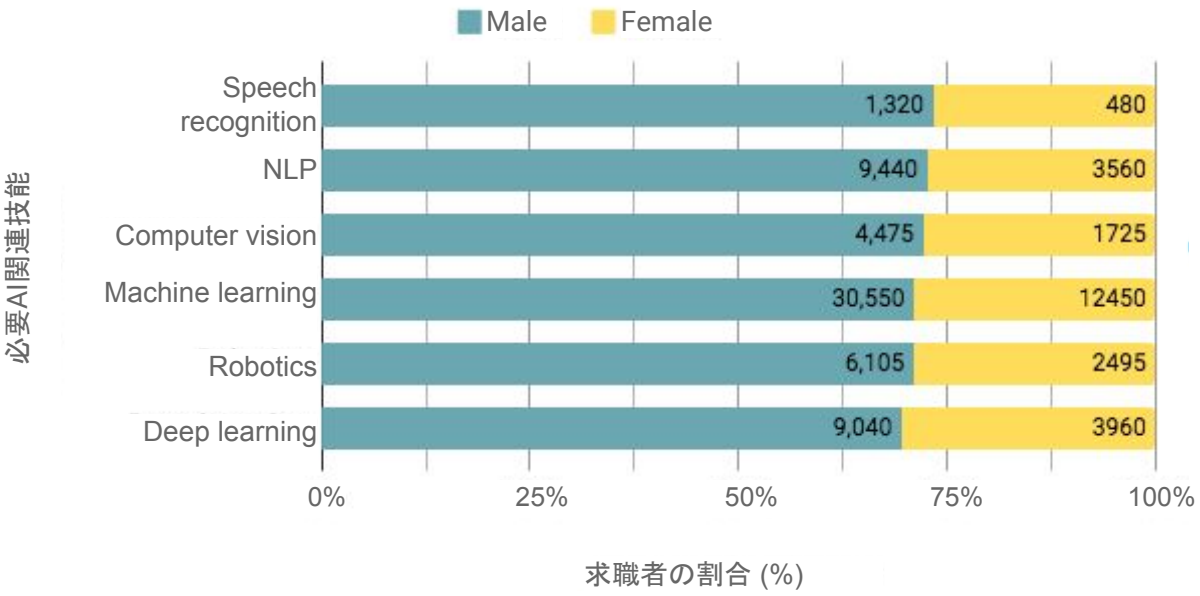
AI関連職：求職者の性別における多様性

2017年、AI関連の求人における応募者の男女比内訳。内訳データは求められている技能ごとに分類されている。2つ以上の技能を必要とする求人は二重にカウントしている。
募集者の数は必ずしも雇用数、または産業全体の状況を表しているわけではない。性別は男女に限定して集計した。

アメリカでのAI関連職に対する求職者において、平均 71%の求職者が男性である。
全体の平均値は、機械学習技能を必要とする職が応募者数最多を誇るため、機械学習を必要とする職に対する応募者の男女比が大きな影響を持つ。機械学習に加え、深層学習やロボティクスは他の技能と比べ、わずかに女性の応募が多い結果となっている。

データと手法の詳細は Appendix 11に記載。

性別ごとに見る求職者数 (2017)
ソース: Gartner TalentNeuron



On average, men make up 71% of the applicant pool for AI jobs in the U.S.

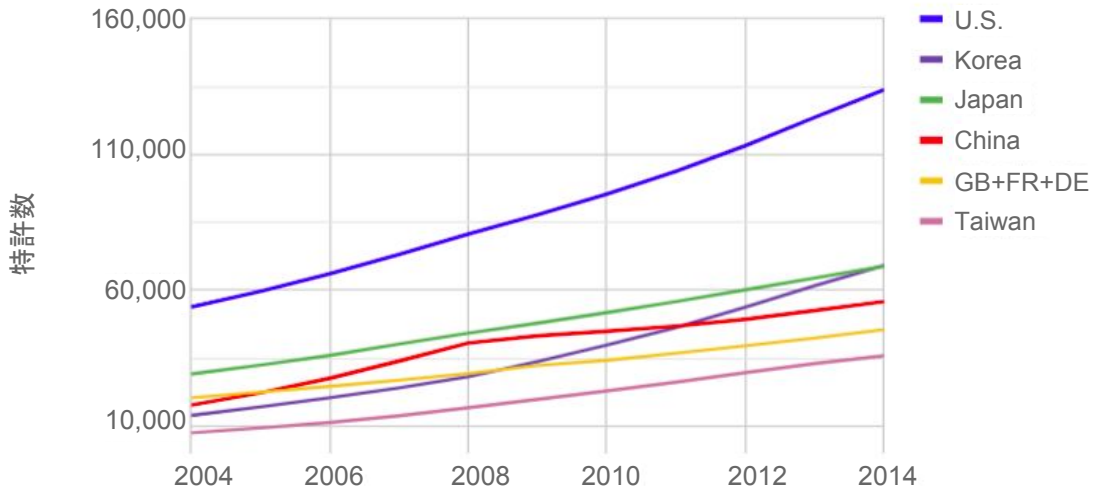
特許

発明者の在住地域・経時的データで見る AI関連特許数とその成長率。
AI関連特許は、*認知と意味の理解*、および*ヒューマンインターフェイステクノロジー* の分野に分類されるものを、IPCコードを使用して集約したものである。特許を長期にわたって追跡することは困難である。
この指標に関する注意事項とニュアンスについて、 Appendixに記載している。

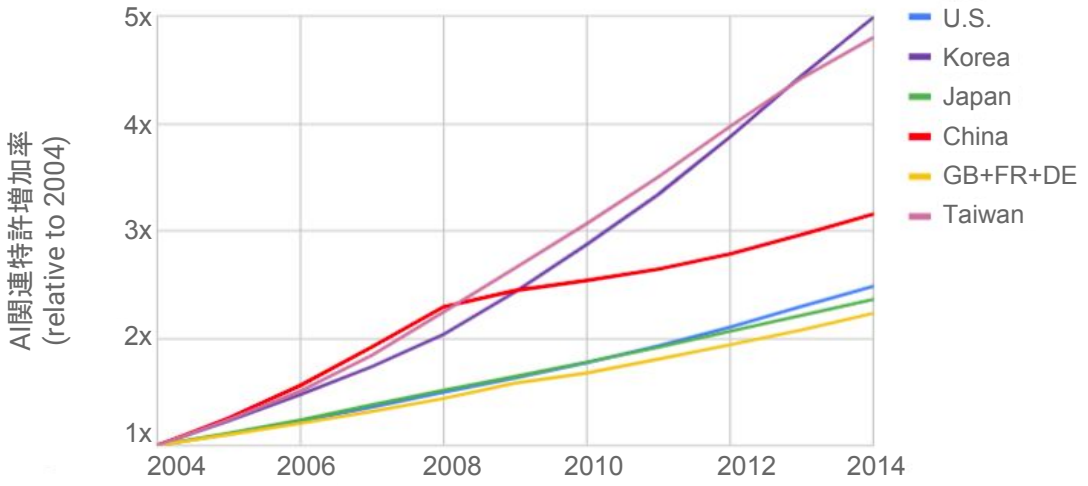
2014年には、AI特許の約30%が米国から、韓国と日本から各 16%の特許が生み出された。発明者の所属する地域の上位では、韓国と台湾が最も成長しており、 2014年のAI関連特許の数は2004年の約5倍に成長している。

データと手法はAppendix 12に記載。

発明者の所属地域から見るAI関連特許 (2004–2014)
ソース: amplified



発明者の所属国から見るAI関連特許増加率(2004–2014)
ソース: amplified



Note: GB + FR + DE refers to a combined number of patents from Great Britain, France, and Germany

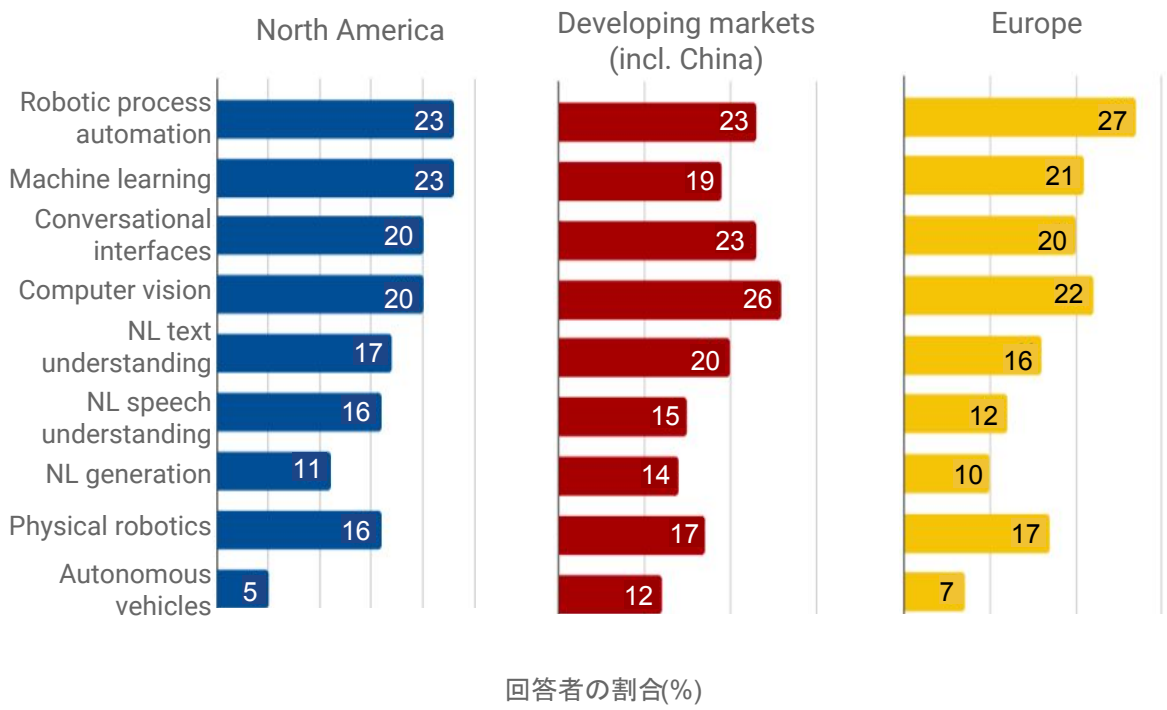
AI導入：地域別

下の図は、McKinsey＆Companyの調査に回答した各組織の代表 2,135人の結果を示している。グラフには、少なくとも 1つの機能または事業単位に AI機能を組み込んでいる回答者の割合が表示されている。本調査において、回答者は複数の AI機能を選択できる。アジア太平洋地域、インド、中東、北アフリカ、およびラテンアメリカのデータは 次のページを参照。

一部の地域では特定の機能が他の地域よりも多く導入されているが、 AIの導入率における各地域ごとの差異は比較的小さい。今後、会社の AI導入が時間とともにどのように変化するか計測することは興味深い結果に繋がる可能性がある。

データと手法の詳細は Appendix 13に記載。

1つ以上の経営機能にAIが導入されている企業の割合 (2018)
ソース: McKinsey & Company



備考: バーのサイズは各地域の導入率と相関関係にある;
North America: N = 479; Developing markets (incl. China): N = 189 (China N = 35); Europe: N = 803

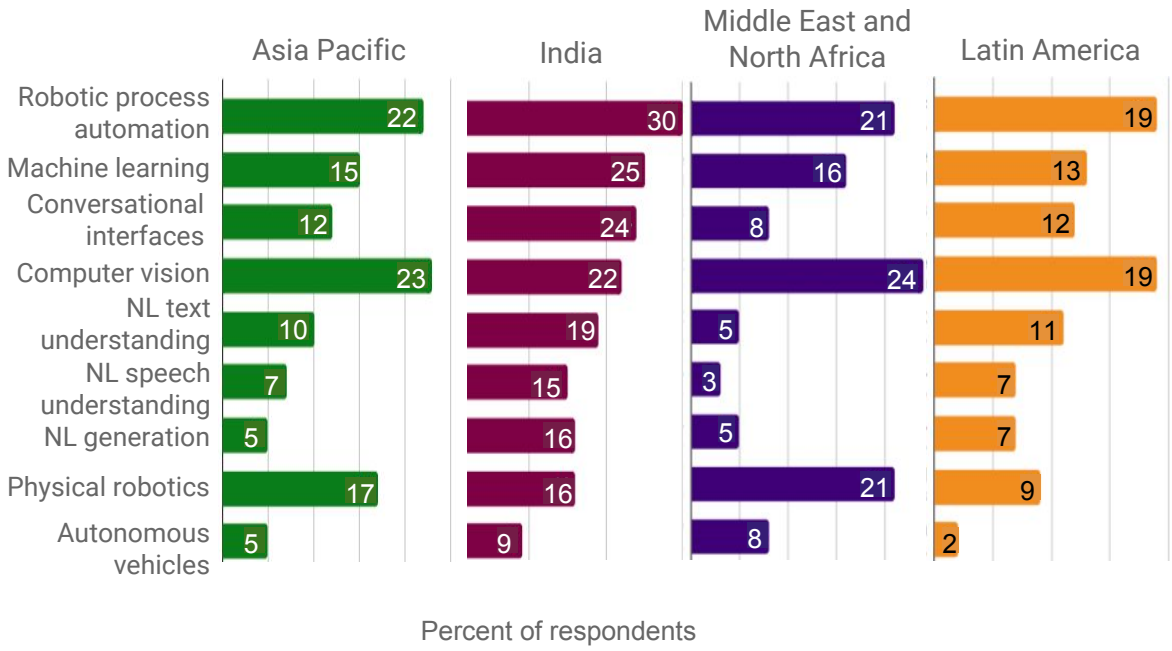
"...We found widespread adoption of different AI technologies across sectors, functions, and geographies around the world; about half of all companies had embedded AI into a corporate business process. However, it's still early; most had not yet adopted the complementary practices necessary to capture value from AI at scale...."

- *Michael Chui, McKinsey & Company and the McKinsey Global Institute*

AI導入：地域別

北アメリカ、発展途上国、ヨーロッパの解説とデータは前ページを参照。

1つ以上の経営機能にAIを導入している企業の割合 (2018)
ソース: McKinsey & Company



Note: バーのサイズは各地域の導入率と相関関係にある;
Asia−Pacific: N = 263; India: N = 197; Middle East and North Africa: N = 77; Latin America: N = 127

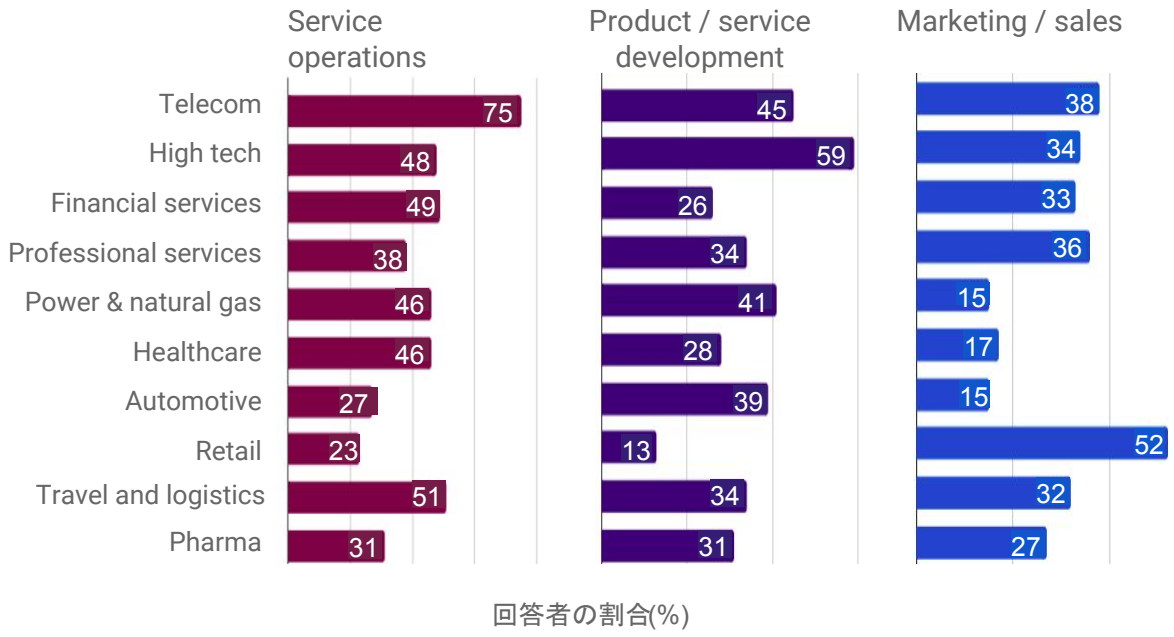
AI導入：産業と機能

下の図は、McKinsey＆Companyの調査に回答した各組織の代表 2,135人の結果を示している。グラフには、少なくとも 1つの機能または事業単位に AI機を導入している回答者の割合が表示されている。本調査において、回答者は複数の AI機能を選択できる。製造、サプライチェーン管理、およびリスクに関するデータは 次のページを参照。

組織は、事業内で最も大きな価値を生む機能に AIを組み込む傾向がある。 たとえば、金融サービスではAIがリスク管理に導入されているが、自動車では製造ラインに組み込まれており、小売ではマーケティング /販売に組み込まれている。この結果は、製造などの特定分野へのAI導入進捗率が、専門性が特に重要である業界での普及率と相関する可能性が高いことを意味している。

データと方法論の詳細は Appendix 13に記載。

各機能と産業ごとのAI導入率 (2018)
ソース: McKinsey & Company

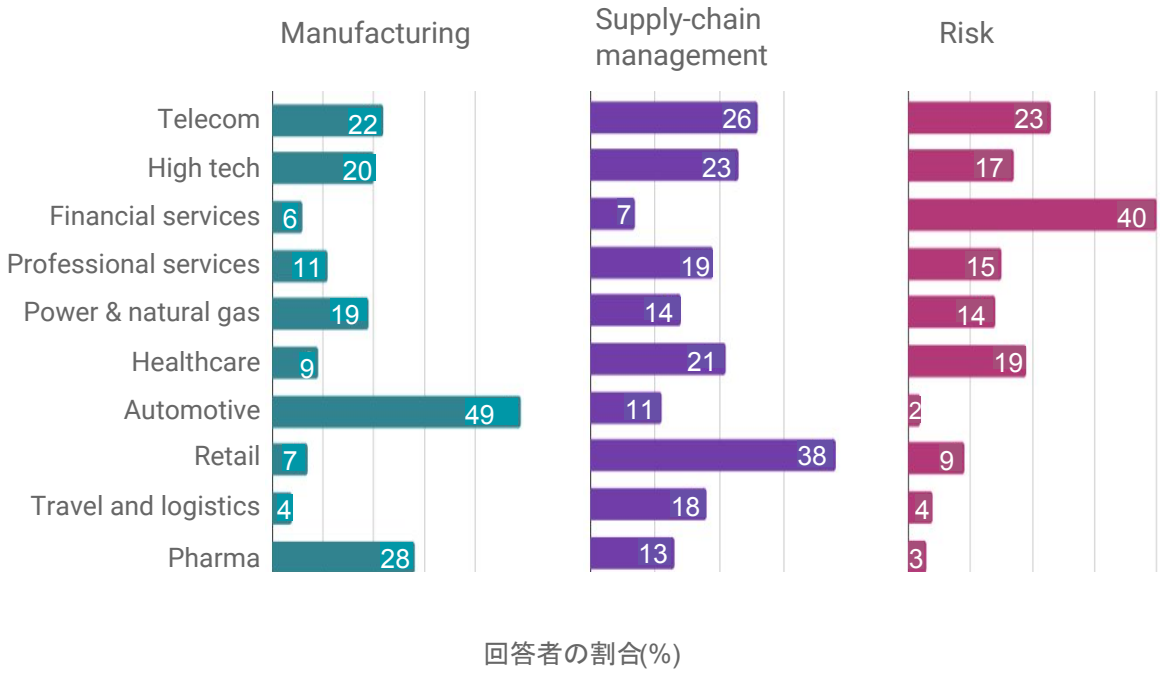


備考: バーのサイズは各機能における産業と相関関係にある:
N = 77; High tech: N = 215; Financial services: N = 306; Professional services: N = 221; Electric power and natural gas: N = 54; Healthcare systems and services: N = 67; Automotive and assembly: N = 120; Retail: N = 46; Travel, transport, and logistics: N = 55; Pharma and medical products: N = 65.

AI導入：産業と機能

Service operations, *Product / service development*, and *Marketing / sales*に関する解説とデータは前ページを参照.

各機能と産業ごとのAI導入率 (2018)
ソース: McKinsey & Company



備考: バーのサイズは各機能における産業と相関関係にある; Telecom: N = 77; High tech: N = 215; Financial services: N = 306; Professional services: N = 221; Electric power and natural gas: N = 54; Healthcare systems and services: N = 67; Automotive and assembly: N = 120; Retail: N = 46; Travel, transport, and logistics: N = 55; Pharma and medical products: N = 65.

Organizations adopt AI in business functions that provide the most value within their industry

This implies that the rate of AI progress for specific applications will likely correlate to uptake in industries where that specialization is particularly important.

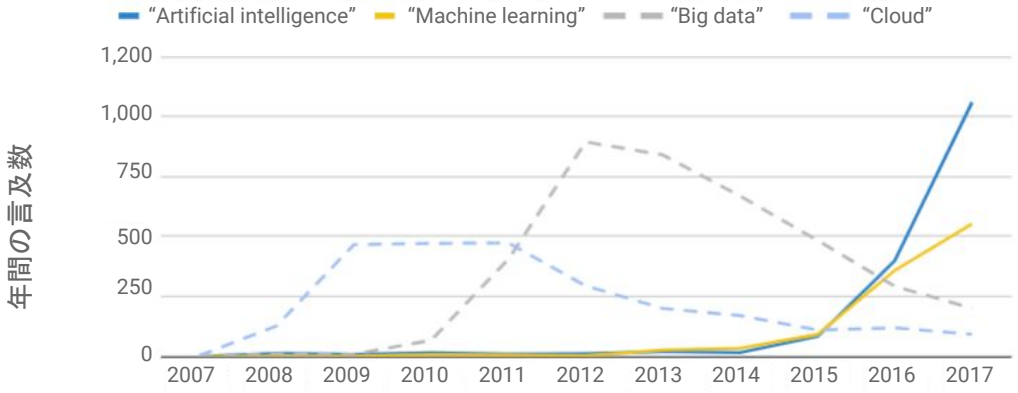
収益報告におけるAIへの言及

各産業で人工知能と機械学習が企業の収益報告で言及された回数。 一つ目の図では、ITセクターの企業に限定したデータをまとめている。
2つ目の図は、IT以外の業界での収支報告における人工知能の言及回数をまとめたものである。ビッグデータとクラウドへの言及も AI / MLへの言及に追加している。分析にはニューヨーク証券取引所で上場されている会社の収支報告を使用している。

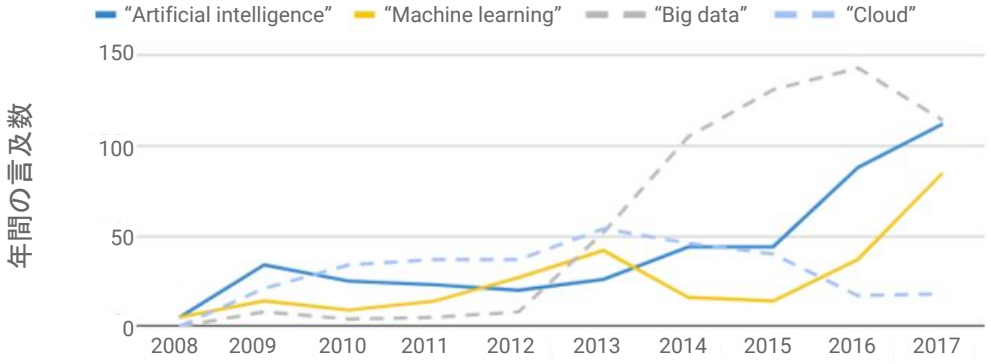
2015年に、ITセクターで人工知能と機械学習の言及が増加し、 2016年には他のほとんどの業界で増加が確認された。 ITセクターに加え、一般消費財、金融、およびヘルスケア業界で、 AIについての言及が多くなっている。

データと手法の詳細は Appendix 14 を記載。

企業の収益報告における言及 — IT 企業 (2007—2017)
ソース: Prattle



企業の収益報告における言及 — 他産業の合計 (2008—2017)
ソース: Prattle



Top 10 companies — AIへの言及 (2017)
NVIDIA Corporation (93), LivePerson (50), Pegasystems (44), Facebook (40), Axon Enterprise (37), salesforce.com (36), Intel Corporation (35), Microsoft Corporation (34), Genpact Limited (34), Applied Materials (33)
Top 10 companies — 機械学習への言及 (2017)
Alphabet(57), Nasdaq (26), ServiceNow (21), Progress Software Corporation (21), Cadence Design Systems, Inc. (20), Splunk Inc. (18), Twitter (17), Overstock.com (14), Synopsys, Inc. (14), Aspen Technology (14)

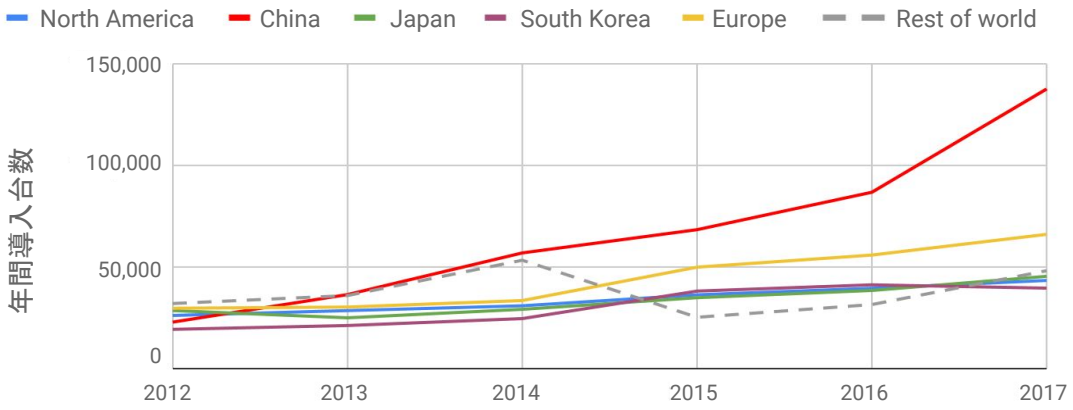
ロボット導入

地域別産業用ロボットユニットの年間導入数。
一つ目の図は、ロボット導入台数における上位 5ヶ国のロボット導入数を示しており、二つ目の図は他地域におけるロボット導入数上位 5ヶ国を示している。

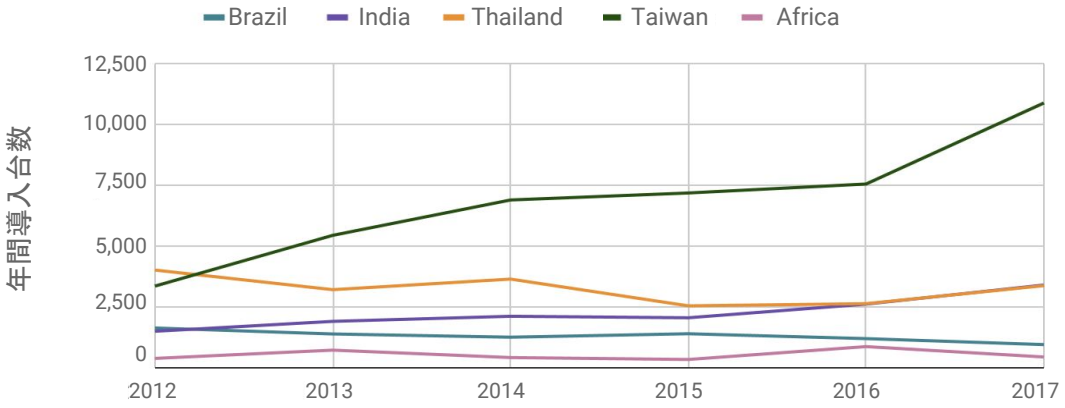
中国は最大のロボット導入国であり、他のロボット導入数上位の国と比較しても、最大の成長率率を誇っている。 2012年以降、中国の年間導入は 5年間で500%増加している一方、韓国やヨーロッパなど、他の急成長地域はそれぞれ 105%と122%の増加を見せている。 小規模生産国の中では、台湾の年間ロボット導入台数が最多で、 2012年から2017年にかけて最も成長している。

データと手法の詳細は Appendix 15 に記載。

ロボット導入 – 上位地域 (2012–2017)
ソース: ifr.org



ロボット導入 – その他の地域 (2012–2017)
ソース: ifr.org



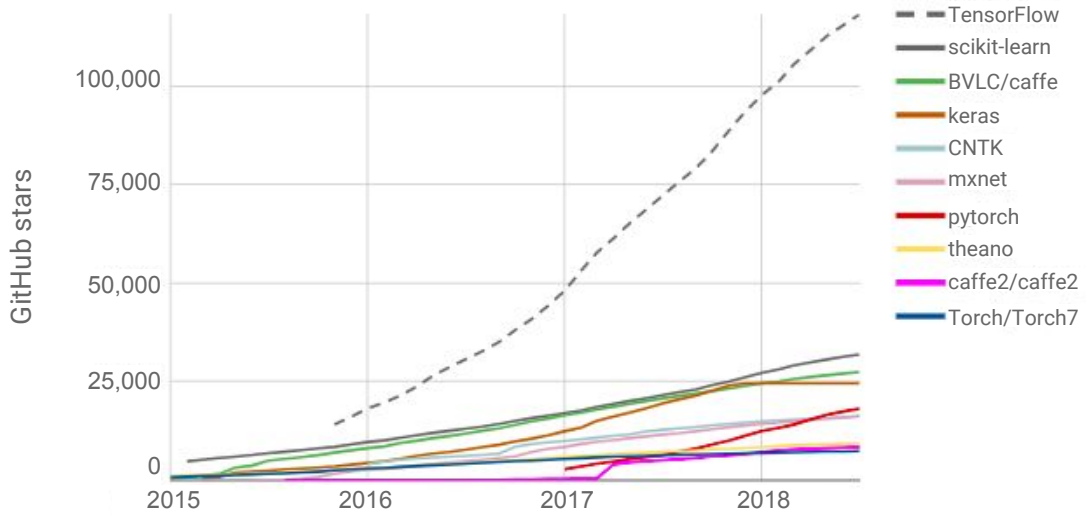
備考: 国内生産台数+輸入-輸出 = 導入台数

GitHub stars

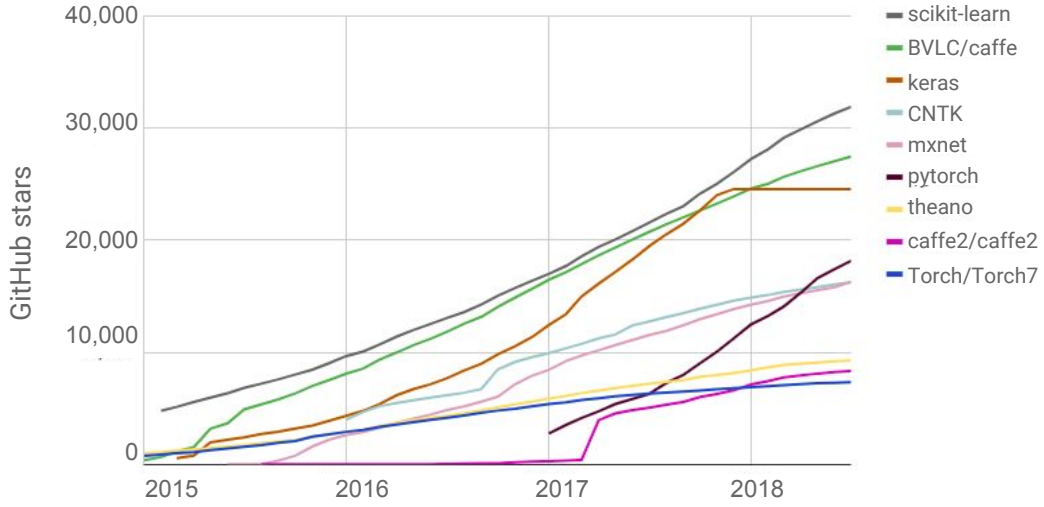
GitHubでさまざまなAIおよび機械学習ソフトウェアパッケージが獲得した Starの数。このデータはさまざまなAIプログラミングフレームワークの大まかな人気度を示している。 大手企業に支えられたフレームワーク {Tensorflow（Google）、Pytorch（Facebook）、mxnet :(Amazon）}の人気の高まり、および TensorFlowに関する言語に対する著しい人気は近年で最も注目すべき潮流である。

データと手法の詳細は Appendix 16に記載。

GitHubにおいて各AIライブラリが獲得した累計Star数 (2015—2018)
ソース: GitHub



Tensorflowを除く各AIライブラリがGitHubで獲得した累計Star数 (2015—2018)
ソース: GitHub



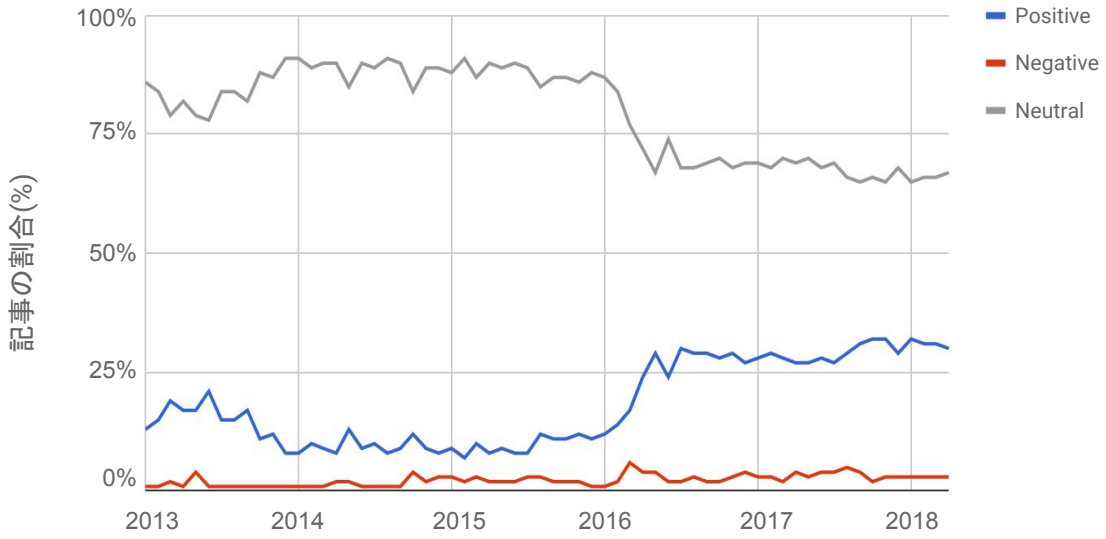
メディア報道における感情

「人工知能」という用語を含むメディア記事を「ポジティブ」、「ネガティブ」、または「ニュートラル」に分類した際の割合。

2016年初めから2017年7月までの6ヶ月間で、ポジティブな記事の割合が 12%から30%上昇したところから、記事はニュートラル主流からポジティブ主流に変化している。それ以降、ポジティブな記事の割合は 30%前後を行き来している。

データと手法の詳細は Appendix 17に記載。

「人工知能」を含む記事のセンチメント (2013−2018)
ソース: TrendKite



Articles on AI became 2.5x more positive from 2016 to 2018

政府によるAIに関する言及

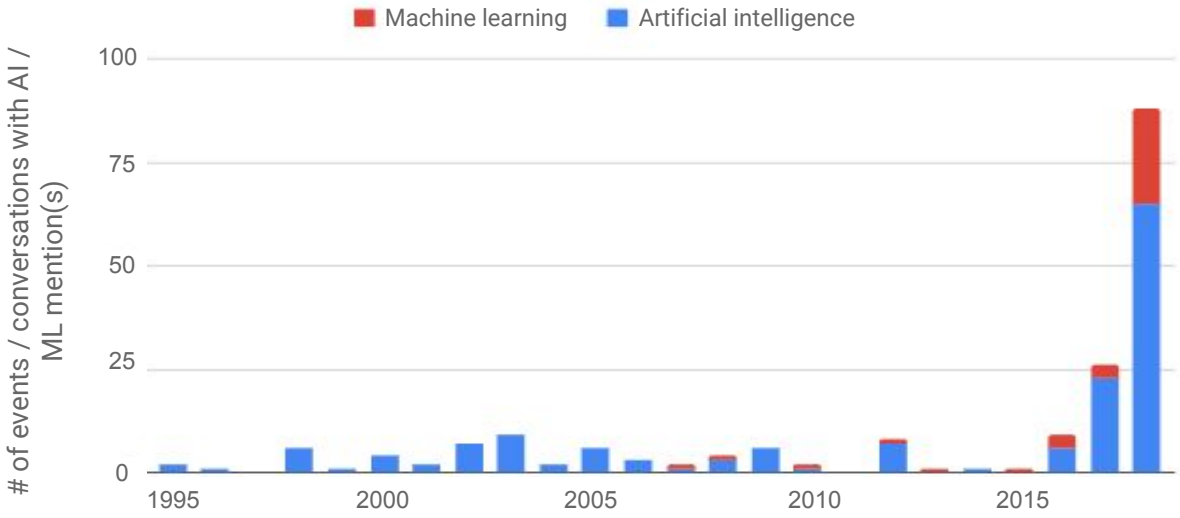
アメリカおよびカナダ、英国の議会の議事録（ Hansards）に人工知能または機械学習という用語が含まれた回数。

2016年、各国におけるAI・機械学習の言及回数が急激に上昇した。 2016年以前、3つの国では、機械学習が言及されるほとんどなく、人工知能と比較して、ごく一部の言及にとどまっていた。

また、手法が異なるため、国ごとの比較が困難という点に注意が必要である。例えば、米国のデータの各カウントは、機械学習または人工知能が特定のイベントまたは会話の中で少なくとも一度は言及された回数を示している。一方で、イギリスのデータの各カウントは、機械学習または人工知能が与えられたコメントの中で少なくとも一度は言及された回数を示している。そのため、国ごとの比較ではなく、国内の経時的傾向を比較している。

下記の図はアメリカに関するデータとなっている。カナダとイギリスは次のページに記載している
データと手法の詳細は Appendix 18に記載。

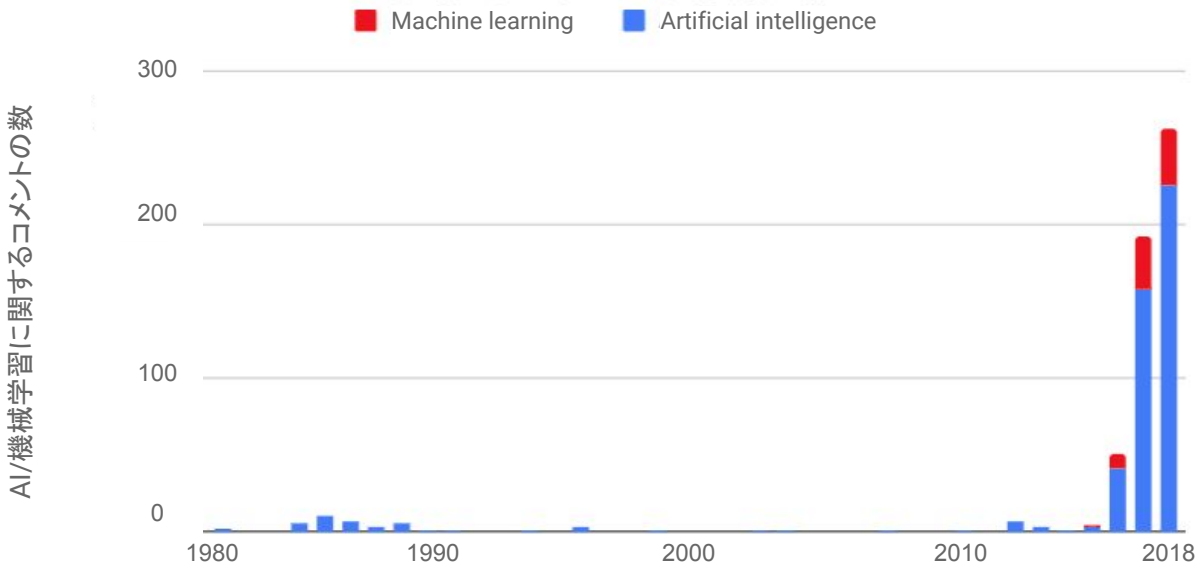
アメリカ議会におけるAIと機械学習に関する言及(1995−2018)
ソース: U.S. Congressional Record website, McKinsey Global Institute analysis



備考: 分析にはアメリカの上院、下院が含まれる。2018年のデータは11/20/18時点まで。

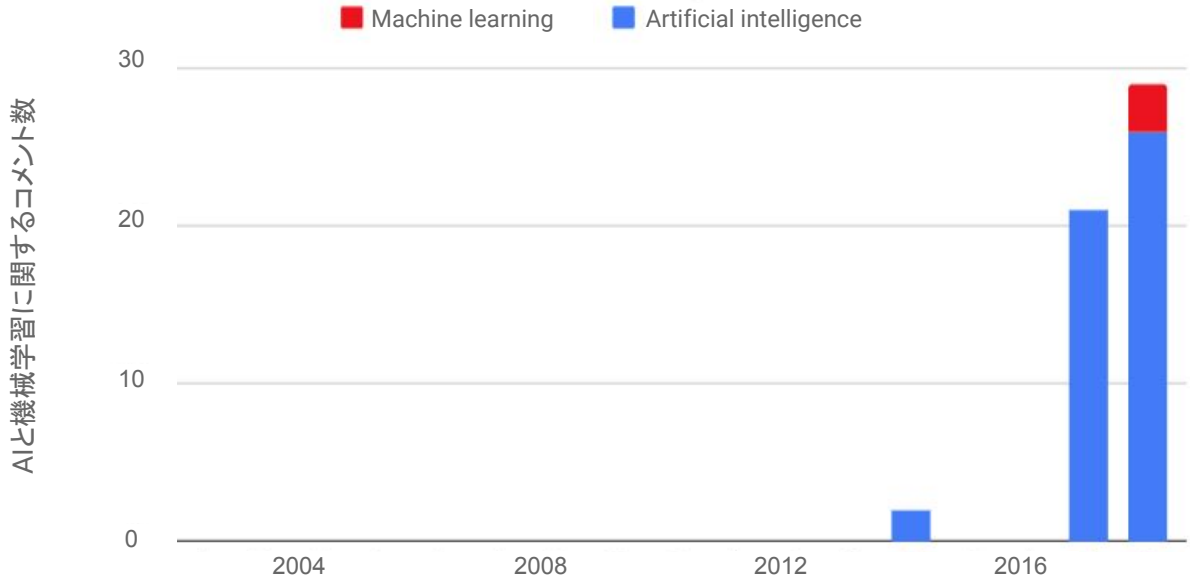
政府によるAIに関する言及

イギリス議会におけるAI と 機械学習に関する言及 (1980–2018)
ソース: Parliament of U.K. website, McKinsey Global Institute analysis



備考: イギリスの下院、上院、ウェストミンスターホール、委員会はすべてこの分析に含まれています。2018年のデータは11/20/18時点まで。

カナダ議会におけるAIと機械学習に関する言及 (2002–2018)
ソース: Parliament of Canada website, McKinsey Global Institute analysis



Note: Only the Canadian House of Commons is included in this analysis. Data for 2018 is through 11/20/18. Please note that this is an unofficial reproduction for non-commercial use. Readers seeking to use this chart should note that the Speaker's Permission does not extend to commercial purpose of financial gains

TECHNICAL PERFORMANCE

物体検出：ImageNet

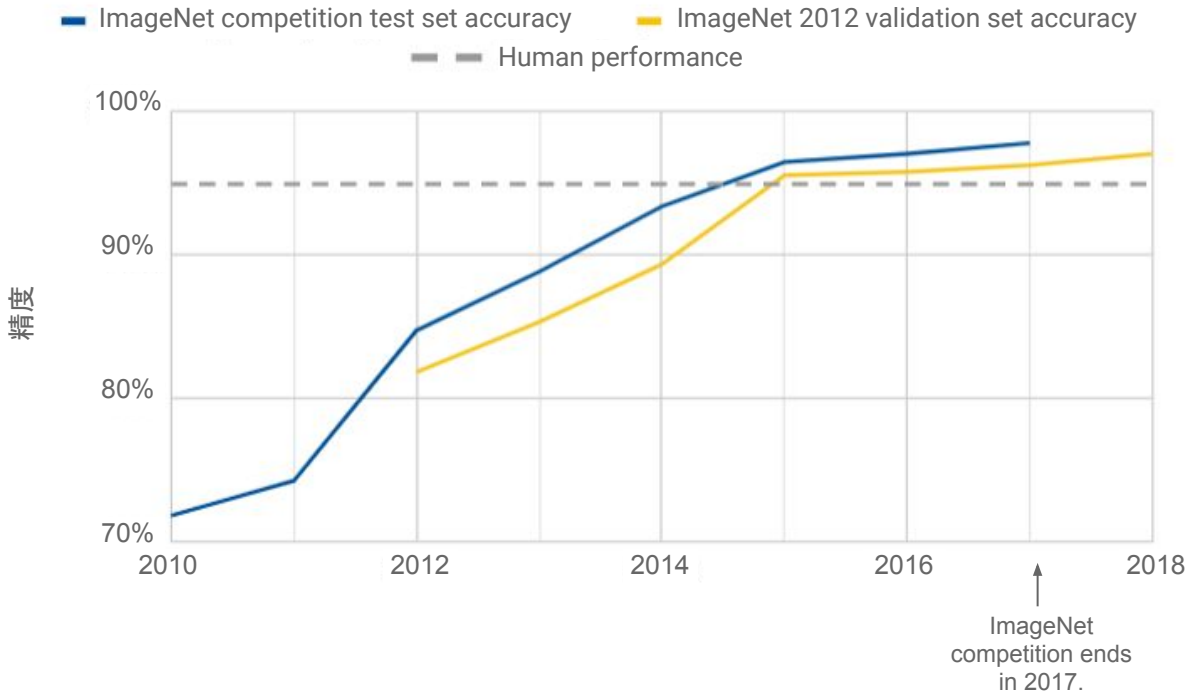
各年ごとのImageNetの精度。ImageNetコンペティションは2017年まで開催され、ホールドアウトされたテストデータセットでモデルをスコア化した。現在、コンペティションは開催されていないため、ImageNet 2012コーパスの検証データセットを使用し、検証している研究論文を通してImageNetの継続的な進歩を追うことを選択した。

検証データセットに対して、コンペティション以外の結果をプロットしているため、モデルを評価するためのホールドアウトテストデータセットは現時点では存在していない。進歩を感覚的に理解するため、2017年と2016年の同じデータセットでは際立った結果を強調している。この図が示すように、ImageNetパフォーマンスは向上し続けている。

この指標はまた、AIの進捗状況を時間の経過とともにモデル化することに内在するいくつかの課題も浮き彫りにしている。例えば、ある研究指標がコンペティションに基づいて構築されている場合、そのコンペティションが終了した場合に、実際の進歩を表すことが難しくなる。しかし、オープンなデータセットが利用可能であれば、いくつかの方法で継続性は確保できる。

データと手法の詳細は Appendix 19に記載。

ImageNet (2010 –2018)
ソース: ImageNet; see appendix

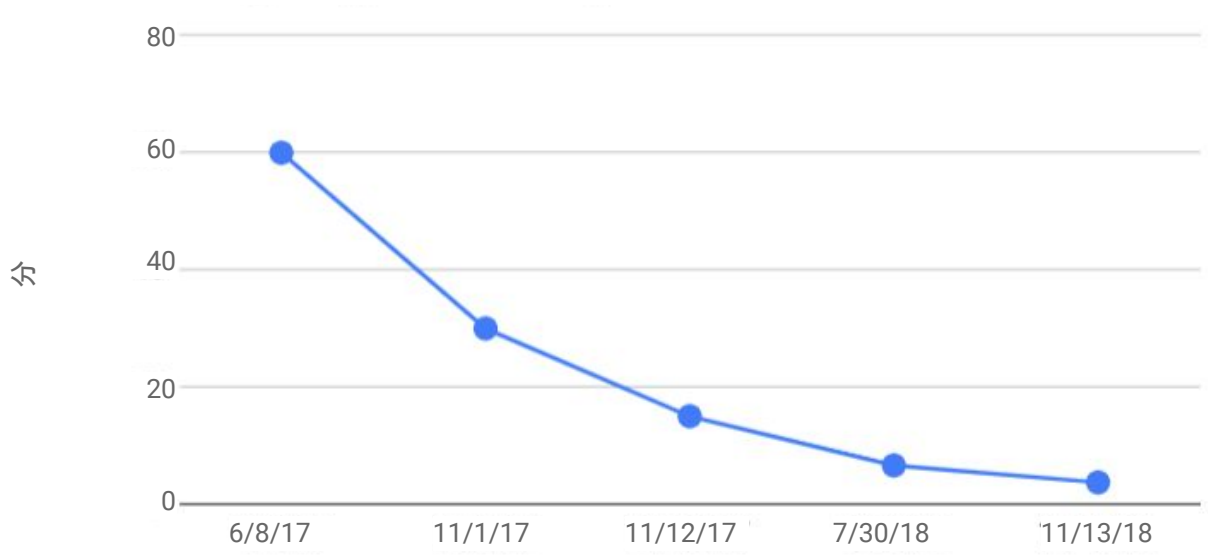


物体検出：ImageNet トレーニングにかかる時間

下の図は、ImageNetコーパス（画像データベース）の画像を最先端の精度で分類する際のネットワークをトレーニングするためにかかる時間を示している。この指標は、大規模ネットワークをトレーニングするために AI分野の資金力のある有力なアクターが画像分類などの AIタスクを実行するのにかかる時間を表すものである。画像分類は（比較的）一般的な教師あり学習タスクであるため、ここでの進歩は他の AIアプリケーションに必要なトレーニング時間とも相関関係にある。トレーニングにかかる所要時間は、ここ 1年半で約1時間から約4分に短縮された。

ImageNetのトレーニング時間を短縮する要因は、アルゴリズムの革新、インフラストラクチャへの投資（ハードウェアなど）、およびトレーニングシステムをサポートするためのソフトウェアの慎重な設計に割り当てられた時間などである。たとえば、トレーニングをさらにスピードアップするために、異なるハードウェアノード間のネットワークを構築することである。
データと手法の詳細は Appendix 19に記載。

ImageNet training time (June 2017 — November 2018)
ソース: arXiv.org; see appendix for authors



ImageNet training time became 16x faster between June 2017 and November 2018

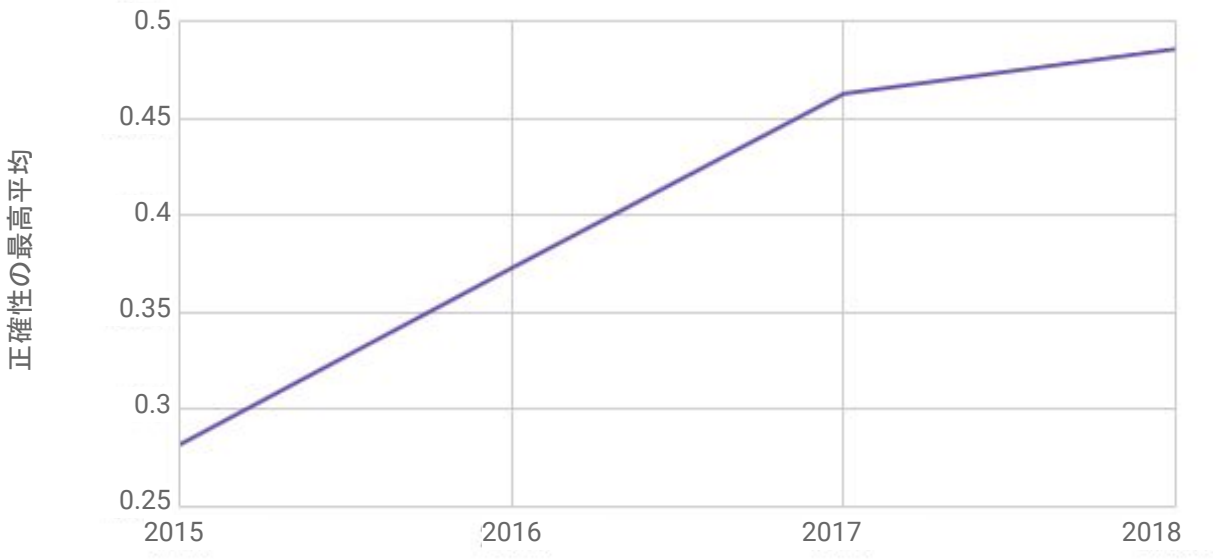
インスタンスオブジェクトセグメンテーション：COCO

2017年、ImageNetチャレンジは、Computer VisionアルゴリズムがImageNetで提供されるオブジェクト検出および画像分類タスクにおいて非常に高い性能を獲得したため終了した。それ以降、研究コミュニティは、より困難なコンピュータビジョンタスクの処理に注力し始めている。つまり、オブジェクトが画像内に存在するかどうかの判断（画像分類）や、対象となるオブジェクトの囲い込んだ矩形の位置特定（オブジェクト検出）の精度がコンピュータアルゴリズムによって向上するにつれて、コミュニティはより詳細な推論を必要とするビジョンタスクに注力し始めたということだ。より詳細な推論を必要とするタスク例としては、ピクセルレベルの精度で境界を囲むオブジェクトをローカライズする（オブジェクトインスタンスセグメンテーション）、またはピクセルレベルの精度でシーンをカテゴリカル領域に分割する（セマンティックセグメンテーション）などが挙げられる。

2018年版AI INDEXでは、Common Objects in Context（COCO）オブジェクトインスタンスセグメンテーションチャレンジから得た成果指標を使用し、より詳細なコンピュータビジョンタスクへと向かう傾向を例証する。アルゴリズムの目標は、入力イメージを取得し、そのイメージ内の各オブジェクトの、定義済みのオブジェクトカテゴリのいずれかに属する正確なピクセルレベルの境界を出力することである。言い換えれば、タスクは関心のある各オブジェクトをローカライズし、その境界を正確に描くことを必要とするものとなる。

2015年以降、COCOチャレンジで達成された最高の平均精度は 0.2ポイント向上し、72%に到達した。
基礎となるデータは こちらを参照。COCOの課題の詳細は COCOチャレンジを参照。

COCO instance segmentation challenge (2015—2018)
ソース: COCO leaderboard



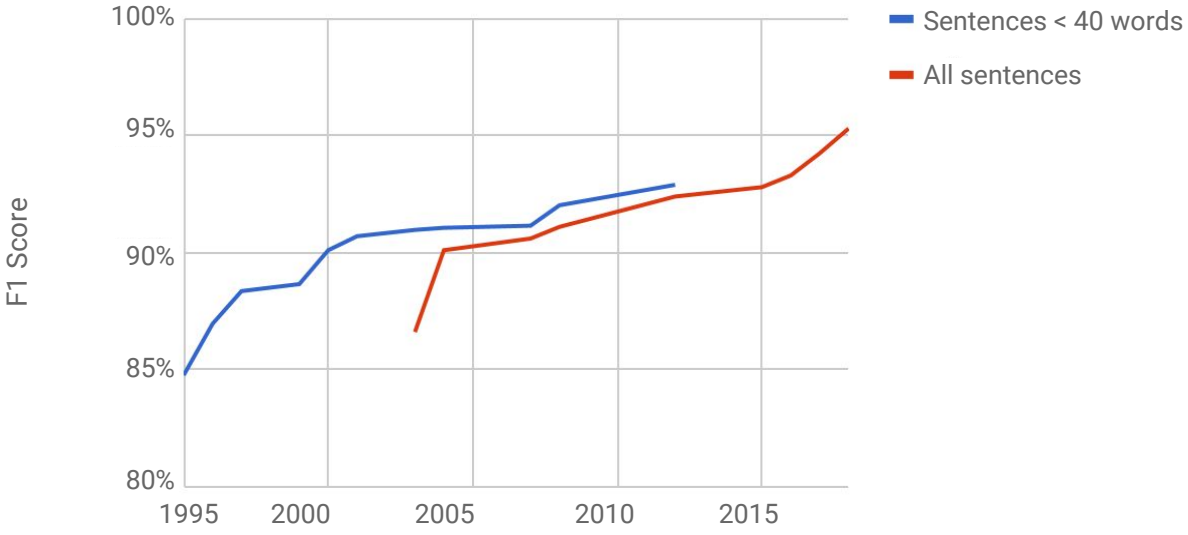
Parsing

文の統語構造を決定するためのタスクに対する AIシステムの性能

これは質問応答などの特定のタスクにおける自然言語を理解するための第一段階である。　もともとプログラミング言語の構文解析に使用されていたものと類似のアルゴリズムを使用して行われていたが、現在では、ほぼ普遍的にディープラーニングを使用して行われている。
2003年以降、全文に対するF1スコアは9ポイント増加した（または 10%の増加）。

データと手法の詳細は Appendix 20に記載。

Constituency parsing — Penn Treebank (1995—2018)
ソース: See appendix



From 2003 to 2018, constituency parsing performance increased by 10%

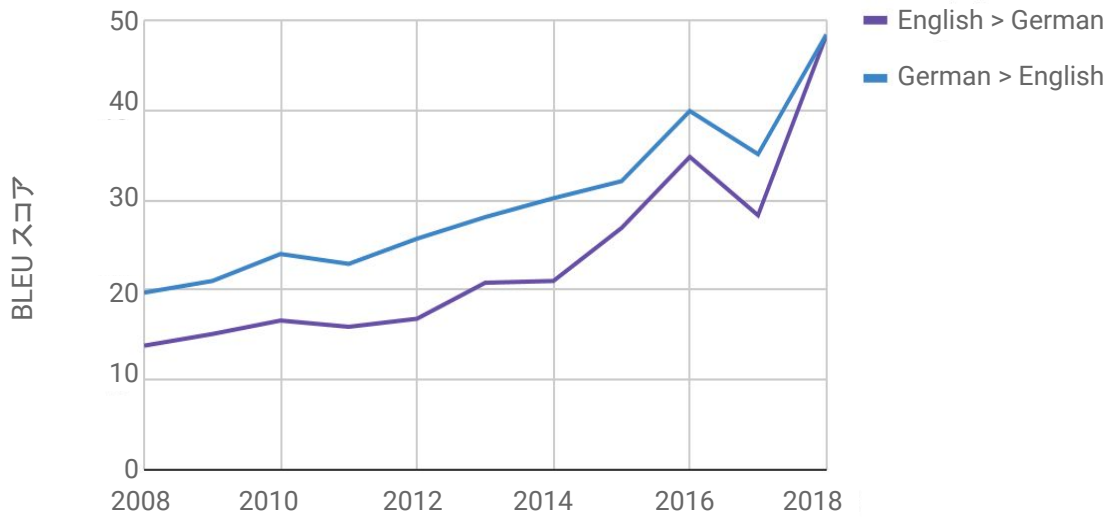
機械翻訳

ニュースを英語からドイツ語に、ドイツ語から英語に翻訳するタスクに関する AIシステムの性能。

2018年、英語からドイツ語へ翻訳した際の BLEUスコアは、2008年の3.5倍となっている。ドイツ語から英語への翻訳は、同期間で 2.5倍に向上した。毎年異なるテストセットを使用しているため、スコアは年間で完全に比較可能なわけではないが（ 2017年のスコア低下の一因 - Appendixを参照）、それを踏まえた上で、BLEUスコアは機械翻訳の進歩を示すために使われている。

データと手法の詳細は Appendix 21に記載。

ニュース翻訳 — WMT competition (2008—2018)
ソース: EuroMatrix



The English to German BLEU score is 3.5x higher today than in 2008

"...There are a couple of things that the top-performing systems do well. They exploit the transformer architecture for Neural MT and they exploit the data more effectively. The transformer architecture can give better results than earlier neural MT models (based on RNNs), but it requires careful choice of hyper-parameters and optimization..."
—Barry Haddow, University of Edinburgh

Commercial machine translation performance
Learn more about the state of commercial machine translation here.

設問に対する解答：ARC

AI2推論チャレンジ（ARC）のデータセットは、7,787問の小学校レベル（アメリカにおけるグレード3～9）の選択式科学問題である。この問題は、高度な質問応答の研究を促進するために集められている。

質問はチャレンジセット（2,590の質問）とイージーセット（5,197の質問）に分けられており、チャレンジセットには、検索ベースのアルゴリズムと単語の共起アルゴリズムの両方が誤回答した質問のみが含まれている。問題はテキストのみの英語試験問題で、ファイルに示されているようにいくつかの学年レベルにわたる。各質問は複数選択構造（通常 4つの回答オプション）を持ち、質問には、ARC関連の知識を含む、14Mの順序付けられていない科学関連の文章の集まりであるARCコーパスが付随している。質問に対する答えがコーパスにあることが保証されているわけではない。

ARCベンチマークは、2018年4月にリリースされ、2018年のパフォーマンスは、Easyセットで63%から69%、Challengeセットで27%から42%に向上した。
データと手法の詳細は Appendix 23に記載。

ARC leaderboard (April 2018—November 2018)
ソース: Allen Institute for Artificial Intelligence

備考: 上記の図は、トレンドラインでつながった主要な投稿を示している。提出日の詳細は、Appendixに記載。

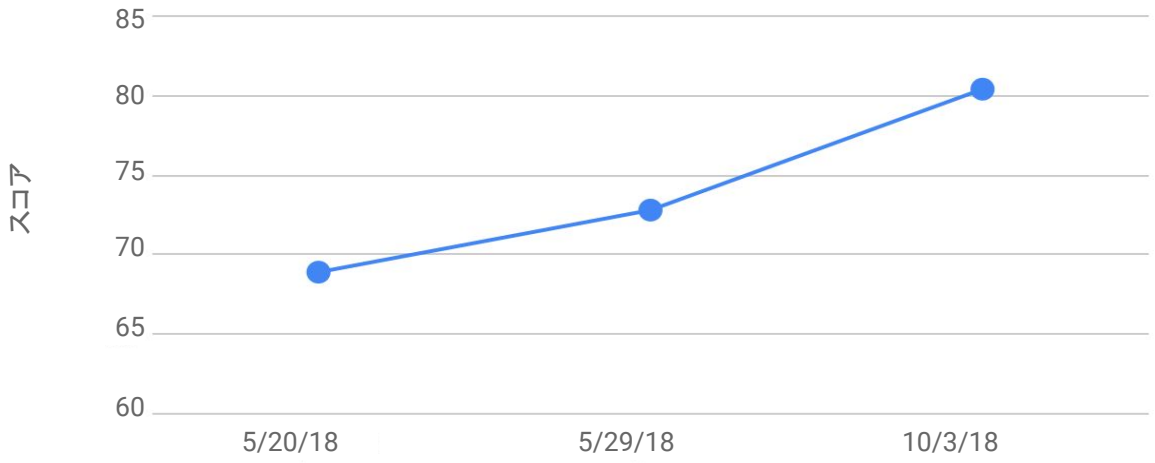
設問に対する解答：GLUE

GLUE（一般言語理解評価）は、さまざまなタスクで自然言語理解（ NLU）システムを検証し、特定のタスクに合わせたシステムの開発を促進することを目的とした新しいベンチマークである。GLUEは9つのサブタスクで構成されている - 2つは単一から言語の受容性と感情を測定するもので、3つは類似性と言い換えに関するもの、そして 4つは自然言語推論に関するものとなっている。企業規模のサイズは従業員数 1,000未満から400K以上までさまざまで、 指標には精度 F1とMatthews Correlation Coefficientが含まれる。

ベンチマークは2018年5月にリリースされたもののみだが、パフォーマンスはすでに公開されており、精度は非専門家レベルの約 90%に達し、最初に公開されたベースラインとの差を約半分に縮めている。

データと手法の詳細は Appendix 24に記載。

GLUE benchmark leaderboard (May 2018—October 2018)
ソース: Gluebenchmark.com



"...We haven't seen a large community emerge around GLUE yet, but I think the two major submissions we've had so far have been recognized as major milestones in representation learning for sentence understanding. Google's BERT—the current state of the art—has been cited eight times despite having been out for just over a month. It was a major focus of discussion in the corridors at the recent NLP conference EMNLP, and it's seeming like it will be the standard baseline to beat in near future work on any language understanding problem, whether or not that work is about pretrained representations..."
—Sam Bowman, NYU

その他のQ&Aデータセット/タスク
Q&Aデータセットは急速に拡大している。AI Index 2018に含まれていないその他いくつかのデータセット: Squad 2.0; OpenBookQA; HotpotQA; MultiRC; NarrativeQA; DrQA; CoQA; QuAC; CWQ (rc setting)

OTHER MEASURES

DERIVATIVE MEASURES

「デリバティブ指標」セクションでは、このページ以前に表示されている潮流間の関係を調査している。 今年も、昨年表示した2つの指標を引き続き見ていく。今後のレポートで新しい指標を検討していきたい。

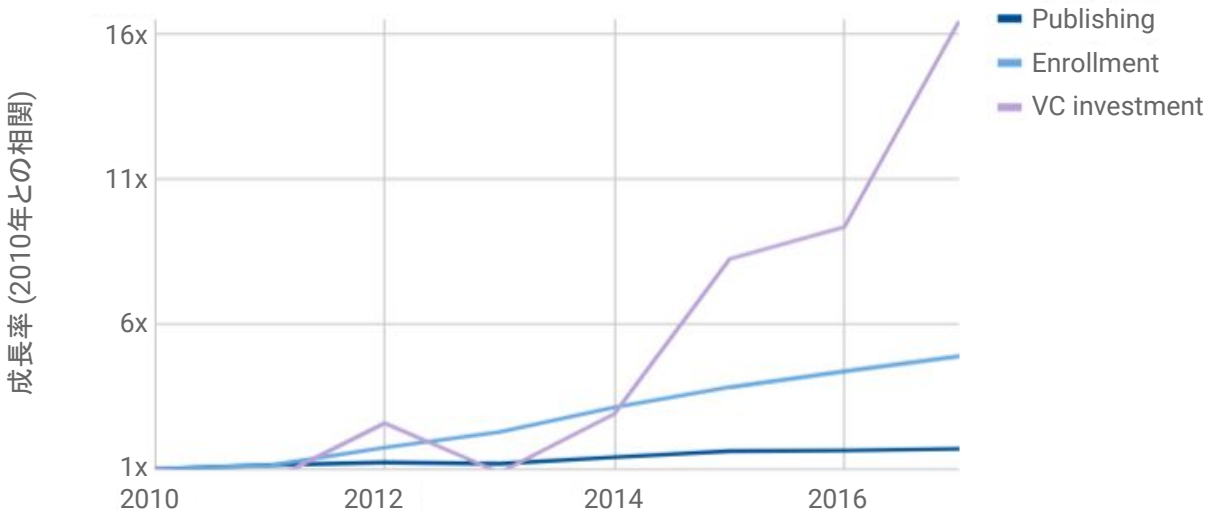
最初の指標、学界 - 産業ダイナミクスは、選択された産業力学の成長とともに、選択された学術指標の成長をプロットする。 第二に、AI Vibrancy Indexは、1つの指数に学術と産業の指標を組み合わせ、表していく。

Academia-Industry Dynamics

学術界と産業界におけるAI関連の活動の関係を調査するため、前セクションからいくつかの代表的な測定値を選択する。特に、ScopusからのAIペーパーの発行、米国のいくつかの大学でのAIおよびMLの入門講座の受講、および AI関連の新興企業に対する VCの投資について見ていく。

これらの測定基準は直接比較できない数値を表しているため、傾向間の関係を分析するにあたり、2010年以降の各測定値を正規化し、絶対数ではなく増加を示す。 2017年のレポートでは、測定値は2000年をベースに標準化されている。現在では、成長率は 2010年をベースに標準化されているため、過去 8年間で出版社の増加と学生の受講者数と比べ、 AI企業のVC支出が大幅に増加していることが確認できる。

学界 - 産業ダイナミクス (2010 – 2017)
ソース: Sand Hill Econometrics, Scopus, university provided data



備考: 2017年のAI Indexでは、受講者数の指標はスタンフォード大学のみの限定的なものだったが、2018年版にはこれらの大学も含んでいる。.

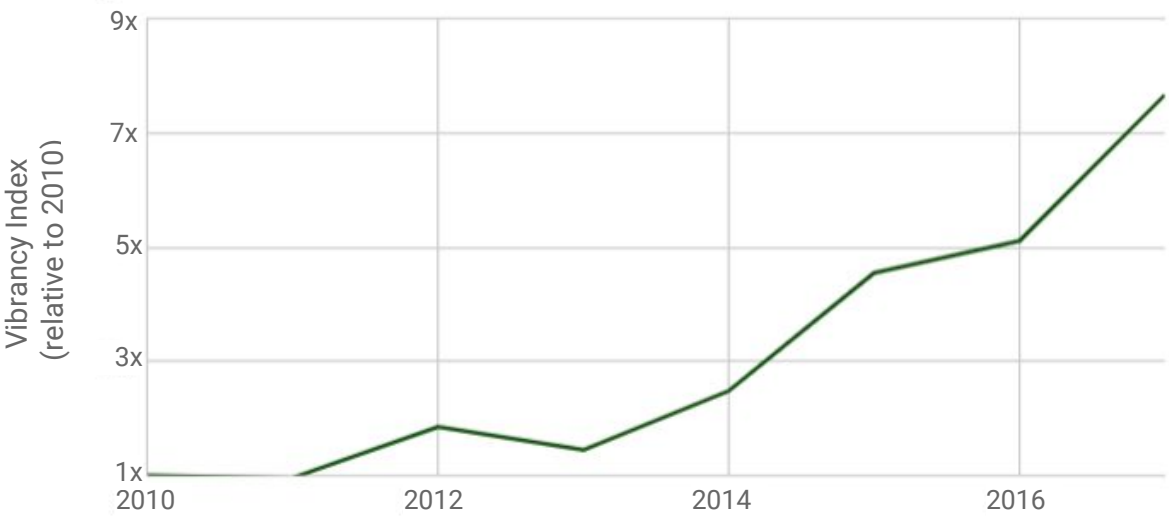
The AI Vibrancy Index

AI Vibrancy Indexに、3つの学術 - 産業指標（出版、登録、および VC投資）を1つの測定値に集約した。AI Vibrancy Indexを使用する目的は AI分野の活気を定量化するためである。 学界 - 産業ダイナミクスと同様に、AI Vibrancy Indexは2010年をベースに標準化されています。

VC投資は他の2つの指標（前のページを参照）と比較して最も成長しており、3つの指標は等しく重み付けされているため、AIのバイブランシー指数は主に VCの投資によって左右される。読者には、AI Indexのデータをダウンロードし、測定基準とその重みを調整して、測定基準の関連性に関する特定の見解と最も一致する Vibrancy Indexを作成してみることを勧める。

Publishing：Publishing Appendixに記載
Enrollment：Enrollment Appendixに記載
VC Investment：VC Appendixに記載

AI Vibrancy Index (2010 — 2017)
ソース: Sand Hill Econometrics, Scopus, university provided data



RECENT GOVERNMENT INITIATIVES

The resources listed below were compiled by the European Commission's Joint Research Center, and by Tim Dutton in his June 2018 Overview of National AI strategies.

USA	EUROPE *(EU, U.K., and France only — does not include several initiatives by other european countries)*	CHINA
2016年10月: オバマ政権が、national AI R&D strategyを提案。国家AI研究開発戦略は、研究への投資、人間と人工知能のコラボレーションのための手法の開発、人工知能の安全性の担保、AIの倫理的、合法的、社会的意義、人工知能訓練のための公的データセットの作成、およびベンチマークと基準を通した人工知能の評価を目的としている。アメリカ政府はまた、AIに関する最初の政策報告書Preparing for the Future of Artificial Intelligenceも発表した。 **2018年5月**:トランプ政権による*Summit on AI*は、(1)AIにおけるアメリカのリーダーシップを維持すること(2)アメリカの労働者を支援すること(3)公的研究開発を促進すること(4)イノベーションに対する障壁を取り除くこと を目指していると発表。*The Select Committee on Artificial Intelligence*が、AI研究開発における優先事項についてホワイトハウスに助言し、学界および産業界とのパートナーシップの創設を検討するために創設された。	***EU*** **2018年4月**:EU加盟国が*Declaration of Cooperation on AI*に署名し、AIに関する連携に同意。ECもまた、(1)EUの研究投資を促進すること(2)社会経済的変化へ備えること(3)倫理的および法的枠組みを構築すること の3点を目的とした*Communication on AI*を発行している。*Communication on AI*はまた、2018年から2020年の間にAI研究を支援するために15億ユーロ(17億USドル)の予算を設定している。これは、Horizon 2020 programから拠出される26億ユーロに追加されたものとなる。ECは、今後10年間でAIの研究に200億ユーロを投資することを目標としている。 **2018年6月**:ECは2021年から2027年までの間に9.2Bユーロ(10.4Bドル)の予算でDigital Europeプログラムを提案。プログラムは、AI技術の進歩と経済と社会全体でのAIを確実に活用することに焦点を当てている。ECはまた、すべての人がアクセスできる「ヨーロピアンアルゴリズムライブラリ」の開発を提案している。	2014年以来、中国政府は2018年までに中国で1470万ドルのAI市場を創出し、2030年までに中国がAIで世界を牽引することを目的とした一連の主要な国内AIイニシアチブを開始した。 **2015年7月**:*Internet+*イニシアチブは、中国におけるインテリジェント製造に焦点を合わせています。AI固有の目的として、スマートホーム、スマートカー、ロボットなどの分野でAIの普及を促進すること、音声、画像、ビデオ、地図、研究を含む大規模なトレーニングデータベースを構築すること、コンピュータビジョン、言語処理、ヒューマンコンピュータインタラクションなどの主要なAI技術の開発と工業化を進めることを含んでいる。 **2016年 - 2020年**:robot industry development planが中国における知的産業用およびサービス用ロボットの発展の道を拓く。

Note: See additional metrics on the following page.

RECENT GOVERNMENT INITIATIVES (continued)

The resources listed below were compiled by the European Commission's Joint Research Center, and by Tim Dutton in his June 2018 Overview of National AI strategies.

USA	EUROPE *(EU, U.K., and France only — does not include several initiatives by other european countries)*	CHINA
2018年9月：DARPAがAIテクノロジーの限界を超えるため、20億ドル以上の投資計画を発表。AI Next programが開始される。The Subcommittee on Information Technology of the U.S. House Committee on Oversight and Government Reformは、AIそのものと、AIが政策に与える影響ついての白書を発行している: Rise of the Machines: Artificial Intelligence and its Growing Impact on U.S. Policy 民間企業は、米国におけるAIの開発/投資において中心的な役割を果たしている。2017年、AmazonやAlphabetのような民間テクノロジー企業は、それぞれ$ 161億と$ 13.9億を研究開発に投資した。これを鑑みた上で、2019年におけるNSFの総予算は、DARPAやDOTの自律システムおよび無人システムへの投資と合わせて、合計53億ドルになる。	**フランス** **2018年**：フランスは以下の3つの報告書を発表している: （1）研究、教育、革新に関する戦略を提案するフランスのAI計画 （2）社会技術上の問題に焦点を当てた議会科学技術部によって作成された報告書 （3）政策と高度知的労働人材の流出を減らすことに焦点を当て、またAI研究における多様性の重要性を取り上げた Mission Villani report **イギリス** **2016年**：イギリス議会上院の人工知能委員会に関する特別委員会が一連のAI policy reportsを発表し始めた。 **2018年4月**：政府は研究/教育を支援し、英国のデータインフラを強化するために9億5000万ポンド（12億ドル）を投資する*AI Sector Deal*を発表した。	**2017年7月**：*New Generation AI Development Plan*は2030年までにAI開発を戦略化し、2020年までに中国のAI技術が世界の最先端レベルと同等になるように計画している。中国は、2025年までにAI理論において大きな進歩を遂げ、AIは産業の高度化と経済再編の原動力となると述べている。そして中国は、2030年までに世界の主要なAIイノベーションセンターになると予測される。

HUMAN-LEVEL PERFORMANCE MILESTONES

HUMAN-LEVEL PERFORMANCE MILESTONES

2017年のAI Indexには、AIが人間レベルのパフォーマンスを達成した、または人間を上回るレベルに達したものを示したタイムラインが含まれていた。リストには、AIによるゲームプレイにおける実績、正確な医学的診断、およびAIが人間レベルまたは人間を上回るレベルで実行したその他の洗練されたヒューマンタスクが記されていた。今年は、そのリストに4つの新しい成果を追加した。しかし、これらの結果を過度に解釈しないことが重要となる。以下のタスクは具体的で、その成果は印象的だが、その他のタスクへの汎用的なシステム性能について示唆するものではない。

1980

オセロ

1980年代、Kai-Fu LeeとSanjoy Mahajanは、オセロをプレイするベイジアン学習ベースのシステム「BILL」を開発した。1989年には、プログラムは米国のコンピュータプレーヤートーナメントで優勝し、アメリカで最高位のオセロプレーヤーである Brian Rose（56-8）を破った。1997年に発表された、Logistelloという名前のプログラムは、オセロ世界チャンピオンに対する 6試合すべてに勝利した。

1995

チェッカーズ

1952年、Arthur Samuelsはチェッカーをプレイしながら自己プレイによって改善していく一連のプログラムを構築した。その後、1995年になって初めて、チェッカー攻略を目指すプログラム「Chinook」が世界チャンピオンを破った。

1997

チェス

1950年代、数人のコンピューターサイエンティストが、 1967年までにコンピューターが人間のチェスチャンピオンを打ち負かすと予測した。しかし、チェスチャンピオンの Gary Kasparovにプログラムが勝利したのは 1997年のことだった。この快挙は、IBM「DeepBlue」によって成し遂げられた。最近では、スマートフォンで実行されているチェスプログラムはグランドマスターレベルでプレイすることが可能になっている。

2011

Jeopardy!

2011年、IBM Watsonが、人気のクイズ番組 Jeopardyに参戦した。 IBM Watsonは元勝者のブ Brad RutterとKen Jenningsに対して勝利し、100万ドルの賞金を手にした。

2015

Atari Games

2015年、Google DeepMindのチームは強化学習により、49のAtari製ゲームのプレイ方法を AIに学習させた。このシステムは、大部分のゲーム（例えばブレイクアウト）で人間並みの性能を発揮することができたが、それでもいくつかはまだかなり手の届かないところにある（モンテスマのリベンジなど）。

HUMAN-LEVEL PERFORMANCE MILESTONES

2016

ImageNetにおける物体検出

ImageNetにおける自動ラベリングのエラー率は、人の平均的なエラー率は約 5%である中、2010年の28%から2016年には3%未満に低下した。

2016

囲碁

2016年3月、Google DeepMindチームによって開発された「 AlphaGo」は、世界最高の囲碁プレーヤーの1人であるLee Sedolを4対1で破った。その後、DeepMindは「AlphaGo Master」をリリースし、2017年3月にトップランクのプレイヤー Ke Jieを打ち負かした。2017年10月、Natureの論文に、オリジナルの「 AlphaGo」を100-0で上回る新しいバージョン「 AlphaGo Zero」の詳細が掲載された。

2017

皮膚ガンの分類

2017年、Esteva氏らが2,032の異なる病気の 129,450の臨床画像のデータセットで訓練された AIシステムについて解説し、21の皮膚科医の診断結果と診断性能を比較した記事を Natureで発表した。著者らは皮膚科医に匹敵するレベルで皮膚癌を分類することができる AIシステムを見つけた。

2017

スイッチボード上での音声認識

2017年、MicrosoftとIBMの両社がスイッチボード上で、人間と同等の音声認識性能を達成した。

2017

ポーカー

2017年1月、Libratusと呼ばれるカーネギー・メロン大学で開発されたプログラムが 4人のトッププレーヤーと1対1で戦い、120,000手に及ぶ激闘の末、全員を破った。この戦いでは、ノーリミットテキサスホールデムルールが適用された。 2017年2月、アルバータ大学で開発された、DeepStackと名付けられたプログラムが 11名のプロポーカープレーヤーと対戦し、 1プレーヤーに対し、3000手以上の戦いを繰り広げた。結果として、 DeepStackはプロポーカープレイヤーに対し、統計的意義を証明するのに十分な勝利をおさめた。

2017

Ms. Pac-Man

Microsoftに買収された深層学習チーム、Maluubaは、Atari 2600でゲームの最大ポイント値である999,900に達する方法を学習する AIシステムを開発した。

HUMAN-LEVEL PERFORMANCE MILESTONES

2018

中国語 - 英語翻訳

マイクロソフトは、ニュース記事を中国語から英語に翻訳する際に人間レベルの品質と正確性を実現可能な機械翻訳システムを開発した。 テストは機械翻訳競技で一般的に使用されるデータセットであるnewstest2017で行われた。

2018

Capture the Flag

DeepMindエージェントは、Quake III Arena Capture the Flag（人気のある3Dマルチプレイヤーの一人称ビデオゲーム）の修正版で人間レベルのパフォーマンスを達成した。 エージェントは、ゲーム内でナビゲート、フォロー、防御などの人間のような行動を示した。訓練を受けたエージェントは、チームメイトと対戦相手の勝率を超え、いくつかの既存の最先端システムを破った。

2018

Dota 2

5つのニューラルネットワークで構成された OpenAI Fiveは、Dota 2のアマチュアチームに勝利した（制限あり）。Open AI Fiveシステムは、180年分に値するゲームの自己プレイを毎日学習したものである。（OpenAI Fiveはプロプレーヤーチームに勝利できていないため、人間を上回る性能とは言い切れない。）

2018

前立腺がんの等級分け

グーグルは、前立腺摘除標本の前立腺がんを等級分けする際に、 70%の精度で等級分けするディープラーニングシステム を開発した。 研究に協力した米国の理事会認定一般病理医の前立腺がん診断の平均精度は 61%であった。さらに、ディープラーニングシステムは、バリデーションセットのすべてのサンプルを採点した一般病理学者の中で特に好成績だった 10人のうち、8人より正確な診断を下した。

WHAT'S MISSING?

WHAT'S MISSING?

AI Indexはまだ発展途上の指標であり、AIが置かれている真の現状を表現するためには、今後、多くの新要素を付け足していく必要がある。新たな要素を付け足していくためには、AIコミュニティの継続的な協力と情報提供が必要不可欠だ。本セクションでは、3人のAI専門家からいただいたAI Indexの今後の拡張に対する提言を紹介する。

FRANCESCA ROSSI
IBM, University of Padova *(on leave)*

常識推論と自然言語理解
AIによる質問/回答システムは非常に便利だが、対話を維持する機能がないことが欠点としてあげられる。自然言語に対する深い理解、さらに、常識推論は依然として大きな課題である。これら 2つの機能を計測できる指標は、ダイアログベースの AIシステム、および AIに関する多くの側面を計測能力の進歩に役立つ可能性がある。 AI Indexは、Winograd Schema Challenge、AI2でのAristoプロジェクトおよび関連するKaggle上のコンペティション、ならびにIBM DebaterなどのDialogベースのチャレンジなど、これらの機能に関連するテストおよびチャレンジの進歩を追跡するための指標を追加してもいいように思う。

人間との連携
AIが人間の知能を置き換えるのではなく強化するのであれば、人間 vs機械の構図から、人間＋機械の構図へ移行することも重要となる。
これにおける進歩を示すための測定基準としては、人間が最終的な意思決定者であるが、機械によってサポートされている多くの非自律システムが向いているように感じる。

AI推論+学習
機械学習は、大量のデータ内に隠れた相関関係を発見できるため、多くのアプリケーションで非常に高い成果を出している。
機械学習により、非常に正確な予測が可能になる。ただし、予測から介入へと移行し、未来を形作るためには、多くの分野で相関以上の因果関係を導き出す必要がある。そのために必要となる指標は AIカンファレンスでの因果関係に関する論文数だと考えている。(MLだけではなく、AAAIまたはIJCAIのようなAI会議も)
この指標を明らかにするには、論文のタイトルをマイニングし、 MLとシンボリックAIを組み合わせた作品を注視することが重要だ。その上で考えられる指標は、 ML +シンボリックAIの組み合わせに関する論文の数となる。この指標を完成させるためには人間による分析が必要となるだろう。

WHAT'S MISSING? (continued)

RODNEY BROOKS
MIT

AIコンポーネント搭載するロボット
ロボットの出荷指標を明らかにしていきたい。ロボット出荷指標の妥当性はソースによって異なり、多くのソースが、産業用ロボットの出荷台数のうち、 AI活用が非常に少ない（またはない）ものを引用しているため、 AIの進歩を示す指標としては不十分である。 Rethink Roboticsは私が知っている唯一の例外的なソースだ。

無人偵察機（SLAMおよび他のAIアルゴリズムを使用する）などの AIコンポーネントを持つロボットを調べることもできるだろう。ルンバなどのホームロボットにも AIコンポーネントが搭載されている。ルンバはこれまでで最も出荷されているロボットであるため、出荷台数を計測することで興味深い結果が期待できるだろう。ソーシャルホームロボットには最近いくつかの失敗が見受けられた、ソーシャルホームロボット分野での成功は　AIが大きく関わってくるだろう。

TOBY WALSH
UNSW Sydney and TU Berlin

政府支出と軍事利用
過去2年間におけるAIの最も重要な進展の 1つは、おそらく政府と業界の両方によって発表された投資の増加だ。例えば、英国政府が 10億ポンド、フランスが 15億ユーロ、そしてドイツが 30億ユーロの投資を決定している。同時に、Alibabaなどの企業が 150億ドルをAIに投資する計画を発表し、 SoftBankのVision Fundが、1000億ドルに上る投資の大部分を AIに集中させている。

増加し続ける投資は、しばしば中国のような国々が AIの使用を通して経済的・軍事的支配を求める野心を宣言する「AIレース」として説明されている。 AI Indexに、この増加する投資を測定する際の定量的測定基準を加えるのは有益だろう。AIの研究開発の焦点がどのように変化しているかを測定することは、国ごとの投資を分析する際に役立つかもしれない。今、懸念が高まっている分野の一つとして、 AIの軍事利用がある。地雷除去のように、これが歓迎されるべき AI活用分野はいくつかあるが、完全自律型兵器の開発のように、深刻に懸念されている分野が他にも多く存在している。したがって、 AI Indexは、世界中の軍隊におけるそのような技術の開発と採用の測定という観点を加えるのも有益かもしれない。

ACKNOWLEDGEMENTS

ACKNOWLEDGEMENTS

AI IndexはOne Hundred Year Study on AI (AI100)の中で考案され、現在スタンフォードの Human-Centered AI Institute（HAI）で開催されているAI100傘下の独立したプロジェクトとして設立された。 私たちはAI100とHAIの双方が繁栄していくことを願っている。

Google、Microsoft Bytedance（Toutiao）、およびAI Journalが提供する追加のスタートアップ資金。

AI Indexの将来的に追加すべき項目に関するコメントを寄せてくださった What's Missing?への貢献者たちに感謝を表したい。

Francesca Rossi, Rodney Brooks, Toby Walsh

また、2018年のAI Indexレポートにおける包括生の指標に対し、専門的な解説を提供してくださった個人にも感謝している。

Paul Ginsparg, the WiML board, Michael Chui, Barry Haddow, Sam Bowman, AAAI, James A. Landay

示唆に富んだデータセットと視覚化をいくつか提供してくださった Elsevierには特に大きな感謝を表したい。以下のElsevierに属する個人からはその過程で大きな支援をいただいた。

Ann Gabriel, Clive Bastin, Sarah Huggett, Mark Siebert, Jorg Hellwig

AI導入に関する貴重な調査結果を提供していただいた上、政府による AIまたは機械学習に対する言及を分析する手助けをしてくださったマッキンゼーアンドカンパニーとマッキンゼーグローバルインスティテュートには感謝の意を伝えたい。特に、McKinseyに属する以下の個人は AI Indexにとって変えがたい重要な存在だった。

Heather Hanselman, Daniella Seiler, Jake Silberg

データおよび/または指標の説明を提供してくださった以下の個人に感謝する。

Susan Woodward, Tully Foote, Sam Davis, Lars Kotthoff, Konstantin Savenkov, Anita Huang, Calvin LeGassick

そして2018年のレポートの日々の制作に欠かせない活躍をしてくださった Ian Danforth（Technical Lead）と Michael Chang（graphic design and cover art）に感謝する。

ACKNOWLEDGEMENTS

また、AI Index2018に不可欠なデータを提供してくださった以下の組織、大学、およびカンファレンスにも恩を受けた。

企業 / 団体

arXiv, ROS.org, Sand Hill Econometrics, Monster.com, amplified, McKinsey & Company, the McKinsey Global Institute, Prattle, IFR.org, TrendKite, The Allen Institute for Artificial Intelligence, The European Commission's Joint Research Center

大学

University of California-Berkeley, Stanford University, University of Illinois at Urbana–Champaign, University of Washington-Seattle, Carnegie Mellon University

Tsinghua University, National Institute of Astrophysics, Optics and Electronics, University of British Columbia, University of Toronto, University of Edinburgh, University of Science and Technology of China, Shanghai Jiao Tong University, Pontificia Universidad Católica de Chile , TU Wien, Hebrew University, EPFL, MILA

カンファレンスと取り組み

AAAI, AAMAS, AI4ALL, ACL, CVPR, ICAPS, ICLR, ICML, ICRA, IJCAI, KR, NeurIPS, UAI WiML

APPENDIX

APPENDIX 1
Scopus上の論文

Return to Scopus papers by topic (pg 9)
See underlying data

ソース
エルゼビアの学術出版物の Scopusデータベースには、約 7000万の文書が索引付けされている。Scopusに関する詳細情報はこちらを参照。

手法
ここでは、Scopusクエリーツールを使用したうえで、トップダウンアプローチを採用し、 Artificial IntelligenceとComputer scienceのキーワードでタグ付けされた論文の数を取得する。この作業は AIインデックスチームがおこなった。(注)この手法と以前のページで説明した手法との違いに注意することが重要となる。

Scopusは、論文にキーワード、発行日、所属国、およびその他のいくつかの指標でタグを付けている。 AI関連の出版物の発行年および地域の取得において、 'opensearch:totalResults'ツール(紙の枚数を返す)を使用して、年/国の組み合わせごとに次のクエリを実行した。

国別AIクエリ

title-abs-key(人工知能) AND PUBYEAR = {} AND AFFILCOUNTRY({})
トピックごとの各年の発表論文については、以下のクエリを使用した:
AIクエリ
title-abs-key(artificial intelligence)AND PUBYEAR AFT {} AND PUBYEAR BEF {}
CSクエリ
SUBJAREA(COMP) AND PUBYEAR AFT {} AND PUBYEAR BEF {}
全Scopusクエリ
PUBYEAR AFT {}とPUBYEAR BEF {}

上記のすべてのクエリは、1996年から2016年までの各年を分析するために作られた。

ニュアンス
- Scopusシステムは遡及的に更新されているため、結果として、与えられた質問に対する論文の数は時間とともに増加する可能性がある。
- Elsevierチームのメンバーは、1995年以降に発表された論文に関するデータが最も信頼できるものであるとコメントしている(したがって、Scopusデータの開始年として 1996年)。
- Scopusの論文数には、すべてのレベル・および AIとの関連性のある論文が含まれている。

地域ごとのAI出版物における固有のニュアンス
- 複数の地域にタグが付けられている場合、それらの論文は二重に数えている。これは、特定の年の一番上の行番号が昨年の年間の論文数と一致しない要因となっている。
- 「その他」には、Scopus上でAI論文を発表したその他すべての国々が含まれている

トピックごとの出版物における固有のニュアンス
- 2017年インデックスは、CSカテゴリ内のAI論文のみ示しているが、2018年版では、CSカテゴリに分類されるかどうかにかかわらず、AIとしてタグ付けされた全論文を含んだものとなっている。
- ScopusのAI論文は一般的にCSの大カテゴリーに分類されるため、CS論文数の増加にはAI論文の増加も関わっている。
- 「全論文」カテゴリには、すべての AIおよびCS論文が含まれる。

Nuance
- The Scopus system is retroactively updated. As a result, the number of papers for a given query may increase over time.
- Members of the Elsevier team commented that data on papers published after 1995 would be most reliable (hence 1996 as a starting year for Scopus data).

APPENDIX 2
Scopus上の論文数計測に対するElsevierの手法

Return to Elsevier papers section (pg 10)
See underlying data

ソース
Elsevierの学術出版物に関するScopusデータベースには、約7000万の文書が索引付けされている。これはElsevierによって編集された。Scopusの詳細情報はこちらを参照。Elsevierについてはこちらを参照。

紙の索引付け、所属、地理的対応範囲、およびタイトルに関する詳細な手法は、 Scopus Content Coverage Guideを参照。

手法
ElsevierのAI論文数の取得手法は、約800個のキーワードを含むAI論文のボトムアップアプローチによる集計である。詳細はElsevierの報告書を参照。(注)この手法と次のページで説明する手法との違いに注意することが重要となる。

AIの定義
ElsevierはAIに関連する約800のキーワードを設定した。キーワードは、ElsevierによるAIレポートの基礎となるScopusから、60万のAI論文を特定するために使用された。キーワードは次のように決定された。

まず、Elsevierは、Fingerprint Engineを使用し、AI関連著書、コースのシラバス、特許、およびニュース記事を取り込んだ。Fingerprint Engineは文書を要約し、最も関連性のある一連の重み付けされたキーワードを出力。生成された2万件の関連用語を手動で確認した結果、797個の固有のキーワードが選択された。エルゼビアは、今後数カ月以内にそれらを発表すること考えている。サンプルリストは こちらを参照。

だが、797のキーワードを含むいくつかの論文は、実際には AI関連論文ではなかった。これは誤検知により起こった問題である。例えば、倫理的価値という用語は広義で、 AI以外の論文に含まれる可能性がある。同様に、ニューラルネットワークは、生物学などの非 AI分野に属す用語のため、誤検知を生む可能性がある。

誤検知に対処するため、Elsevierは手動または手動で選択した1,500のAI出版物からなるコアまたは応用 AIについての専門家によるガイダンスを追加した。 Elsevierは教師付き分類子に専門家のガイダンスを使用し、 Scopusからの60万を超える出版物セットの分類作業をし、誤検出出版物を排除した。

全体として、Elsevierのアプローチは、複数の視点とキーワードで AIの分野を広く捉えながら、監視付き分類器の精度も取り入れるように設計されている。

所属国
著者の所属国は、論文の著者の所属組織に基づいて決定される。国際組織においては、著者が所属国を明記していない限り、本部の国籍をデフォルトとして使用している。たとえば、自分の組織として「 Google」と入力した作者は米国に属し、「Google Zurich」と入力した著者はヨーロッパに属す。

複数の地域の著者が共同作業する場合には、論文は二重に数えられている。たとえば、ハーバード大学とオックスフォード大学の著者の論文は、米国で 1回、ヨーロッパで1回と数える。

本レポートでは、EU 44をヨーロッパとして定義している。

これらの規則はセクターの所属にも適用される。

APPENDIX 2
Scopus上の論文数計測に対するElsevierの手法

AIサブカテゴリ（図を参照）
ScopusのAIサブカテゴリを決定するための Elsevierによる手法の概要を以下に記している。 Elsevierの手法の詳細については、ElsevierのAIレポートを参照。

サブカテゴリを幅広く決定するため、 ElsevierはLouvain methodと呼ばれる教師なしクラスタリング手法を採用した。このアプローチは、文書内の共起に基づいて、キーワードをクラスタにマッピングし、それらの関連性を示すものとなっている。共起は、それらのクラスタが独立しているのではなく、互いに強く関連していること（例えばコンピュータビジョン文書におけるニューラルネットワーク）を示している。 Elsevierは、AI技術が次の分野に集中しているようであることを発見した。

確率的推論、ニューラルネットワーク、コンピュータビジョン、検索と最適化、 NLPと知識表現、ファジー論理、計画と意思決定

今後数週間のうちに、Elsevier AIリソースセンターは、クラスターの相互作用型グラフを公開し、ユーザーが地域ごと、あるいは時系列ごと個々の関連性とクラスターを閲覧できるようにするようだ。

相対的な注力点（図を参照）
論文は、その分野と内容に基づき専門の索引作成者によって主観的にカテゴライズされている。主題分野は相互に排他的な訳ではない。

ニュアンス

- 全体的なデータ収集プロセスは、Elsevier内部で行われた - AIインデックスは、キーワード選択の過程、または関連する論文数のカウントには関与していない。
- これらの数値は、Scopus上で異なる手法を用いて集計した AI Index 2017年版レポートの約 2倍となっている。また、Web of Scienceのデータに基づくChina AI Development Reportの約2分の1となっている。
- 音声認識、コンピュータビジョン、ロボット工学、サイバーセキュリティ、バイオインフォマティクス、およびヘルスケアなどの多くの分野で AI活用が急速に進んでいるため、AIの境界は確立するのが困難になっている。しかし、AIはロジック、確率と統計、最適化、写真測量、神経科学、ゲーム理論などの多くの分野に手法的に依存しているため、AIの制限を定義することも困難だ。コミュニティが AIの計量書誌学的分析に関心を持っていることを考えると、これらの研究を作成するグループ、特に基礎となる書誌データベースが異なるもの同士が、結果の再現性を高めるために手法の透明性を追求することは有益だろう。

APPENDIX 3
arXiv上の論文

Return to arXiv papers section (pg 12)
See underlying data

ソース
arXiv.orgは、物理学、数学、コンピュータサイエンス、定量生物学、定量財務、統計学、電気工学およびシステム科学、および経済学の分野における研究論文のオンラインアーカイブである。 arXivはコーネル大学が有し、運営している。arXiv.orgの詳細はこちらを参照。

手法
分析のためのrawデータはarXiv.orgの代表によって提供していただいた。選択したキーワードとそれぞれのカテゴリは以下のとおり。

Artificial intelligence (cs.AI)
Computation and language (cs.CL)
Computer vision and pattern recognition (cs.CV)
Machine learning (cs.LG)
Neural and evolutionary computing (cs.NE)
Robotics (cs.RO)
Machine learning in stats (stats.ML)

ほとんどのカテゴリで、arXivは1999年から2017年までのデータを提供している。今回は、2010年以前のarXiv上には存在しなかったMachine Learning in Statsを分析に含めるため、2010年以降をデータとして使用した。

arXivでの他のカテゴリの投稿率を確認するには、arXiv.orgsの投稿統計を参照。

ニュアンス
- カテゴリは著者自身により分類される - 表示されているものは「主要」カテゴリとして著者に選択されているものとなる。したがって、合理化された分類プロセスではないことに注意が必要だ。さらに、論文著者が他のサブフィールド／キーワードにより、人工知能または機械学習のカテゴリに分類しようとしている可能性がある。
- arXivチームメンバーは、arXivへ参加することはより多くの参加者を呼ぶことにつながると語っている - つまり、arXivの特定のサブカテゴリを主要カテゴリとして選ぶ著者が増加することが、特定のコミュニティによる過剰な参加を促進することにつながる可能性があるということだ。
- 説明で述べたように、arXiv上の掲載論文の増加は、そのトピックに関する論文の実際の増加を反映しているわけではなく、これは、arXiv.orgによる掲載論文数の増加に向けた取り組み、または AIコミュニティにとっての普及の重要性の高まりによる可能性がある。

APPENDIX 4
各国のAAAI論文

Return to AAAI papers by country (pg 20)
See underlying data

ソース

Association for the Advancement of Artificial Intelligence（AAAI）は「AAAI会議」を含む、AI関連会議を毎年開催している。2018年における各国のAAAIへの論文提出/採択に関するrawデータは、AAAIの代表者に提供していただいた。AAAIカンファレンスの 詳細はこちらを参照 。

手法

AAAIチームからAAAIへの提出/採択に関するデータを収集した。 AAAIはこのレベルの詳細情報を得ることができた唯一の会議だった。AI Indexは、将来のレポートに他の会議と同等のデータを含めたいと考えている。

ニュアンス

- ここで分析されている国は、AAAI会議に10以上の論文を提出した国に限定されている。
- このデータは、2018年2月にルイジアナ州ニューオーリンズで開催された会議のもので、提出 /採択された論文の状勢は他の年とは異なる場合がある。
- 論文の採択は空間的制約により大きく制限されている。 AAAI2019会議の空間制限に関するAAAIからのメッセージの引用は以下を参照（ 01/27/19 - 02/01/19 in Honolulu、HI）。

「今年は過去最高の7,700本以上の論文提出があった [2019]。そのうち7,095本を見直したところ、スペースの制限により、合計で1,150枚の論文を受け入れることしかできず、受け入れ率は 16.2%となった。」 - AAAI

APPENDIX 5
アメリカにおけるAIと機械学習講座受講者数

Return to U.S. course enrollment (pg 21)
See underlying data

ソース
コース登録データは各大学から直接収集した。生徒の総人口は学校のアーカイブから収集している（通常はOffice of the Registrarのサイトに格納されている）。次の大学が分析に含まれる。

カリフォルニア大学バークレー校、スタンフォード大学、イリノイ大学アーバナ・シャンペーン校、ワシントン大学シアトル校、カーネギーメロン大学

手法
アメリカの主要なコンピュータサイエンス大学に、数年にわたる入門 AIと入門ML講座の受講登録者数のデータ提供を依頼したところ、いくつかの大学の参加が決まった。大学に十分な履歴データが蓄積されていない場合、またはデータが過度に抽象的なものだった場合、その入学者数は分析に含まなかった。

大学により、学期ごとの受講者数を提供する形や、年ごとの受講者数を提供する形があった。大学同士を相互に比較するために、学期ごとの受講者数データを年単位でまとめ、年度ごとの受講者数を表示している。さらに、いくつかの大学では「入門」と見なされる複数の講座を提供しているが、他の大学では入門講座は 1コースのみの提供となっていた。適切かつ関連性が高い場合、複数のコースを組み合わせて 1つの「入門AI」をトレンドラインとして表示している。

学部生の割合に対する受講者数の割合は、各年度の AI / ML受講者数を、同じ年の学部生人口で割ったものとなっている。これは学校をまたがる平等な競技場への入学動向を示すことを目的とした計算分野です。

ニュアンス
- ほとんどの大学が、近年の受講者数の増加は、学生からの需要ではなく大学の講座供給機能を示したものだと述べた。データには、講座に正式登録された学生数のみが示されているため、応募学生の総数やその他の需要指標は考慮されていない。
- 講座は一般的に学部生に向けて開かれており、通常は専攻学生、非専攻学生関係なく受講できる。一部の講座では名前が変更されているが、今回は 2017年時点の講座名を表記している。また、大学の管理者に提供していただいたその他の詳細情報やニュアンスも記載している。
- 大学の管理者によって言及されていないニュアンスは、以下に記載しない。

<u>AIコース：</u>
バークレーCS 188
スタンフォードCS 221
UIUC CS 440
- UIUCの代表者は、大講義室の利用や開講数の追加は過剰な学生の需要を満たすためだとしている。

UW CSE 415/416（非専攻）およびCSE 473（専攻）
- CSE 416は2017年の新講義のため、2017年のAIコース登録者数の増加に影響を与えている。

<u>MLのコース：</u>
バークレーCS 189
- バークレーの代表は、成長は新規性、主題の興味、そしてイントロ MLが必要条件を満たす講義である専攻科目の成長による組み合わせによるものであると推測している。

スタンフォードCS 229
UIUC CS 446
UW CSE 446
CMU 10-701

APPENDIX 6
国際的なAI講座受講者数

Return to international course enrollment (pg 23)
See underlying data

ソース
コース登録データは各大学から直接収集した。分析に含まれているのは次の大学。

清華大学（中国）、国立宇宙物理光学研究所（メキシコ）、ブリティッシュコロンビア大学（カナダ）、トロント大学（カナダ）、エジンバラ大学（スコットランド）、中国科学技術大学（中国）　、上海交通大学（中国）、カトリック大学チリ大学（チリ）、ウィーン大学（オーストリア）、ヘブライ大学（イスラエル）、 EPFL（スイス）、MILA（カナダ、ケベック）

手法 - Appendix5の手法を参照。

ニュアンス

- ほぼすべての大学が、近年の受講者数の増加は、学生からの需要ではなく大学の講座供給機能を示したものであると述べた。データには、講座に正式登録された学生数のみが示されているため、学生の応募総数やその他の需要指標は考慮されていない。
- アメリカの大学と違い、他国の大学は講座が学部生限定かどうかにより、受講者数が大きく異なる。
- グラフ1はAIとMLコースを合算した成長を表している。グラフ 2はAI講座の受講者数のみを示している。これは、各グラフに同様のデータが表示されるようにするため、この作業をおこなっている。場合によっては、学校の追加データにアクセス可能だったが、学校間で並行した情報を表示するために利用を断念した。追加データは、右上隅の基礎データリンクからアクセス可能。
- 一部の講座は、名称が変更されているが、 2017年時点の講座名で表記している。また、学校の管理者に提供していただいたその他詳細情報やニュアンスも記載している。学校の管理者によって言及されていないニュアンスは、以下には含まない。

INAOE — Courses: C141 (AI) and C142 (computational learning)
Notes: INAOE AI / ML enrollment is greatly affected by the number of students accepted into the INAOE graduate program as a whole. INAOE representatives say that there is a decreasing number of INAOE students, thus affecting AI / ML course enrollment.
USTC — Courses: USTC listed several introductory AI / ML courses across various departments including the Department of Computer Science and Technology, The Department of Automation, the Department of Information Science and Technology and the Department of Data Science.
University of Edinburgh — Courses: Intro applied ML (undergraduate and graduate students) and Informatics 2D — Reasoning and Agents (undergraduate only)
SJTU — Course: CS410 (undergraduate intro to AI)
PUC — Course: Intro to AI
Prior to 2017, the course was only taught once a semester. The large demand in 2017, relative to 2018, is due to the transition from one course to two courses.
Tsinghua — Courses**:** AI (60240052 & 40250182) and ML (00240332 & 70240403 & 80245013 & 80250943)
Open to undergraduates and graduate students
Toronto — Courses: AI (CSC384) and ML:(CSC411)
2016 was the first year that a summer AI course was open. decision to open two semesters of ML in 2015 — due to increased demand
UBC — Courses: AI and ML (CPSC 322, CPSC 340, CPSC 422)
TU Wien — Courses: Intro to AI (undergraduate) and ML (graduate)
Neither course is mandatory for CS students
Hebrew University — Courses: Intro to AI (67842) and Intro to ML (67577)
EPFL — Courses: Artificial Intelligence (undergraduate) and Intelligent Agents (graduate)
EPFL representatives let us know that there is a lot of variation each year due to scheduling and resources.
MILA — Machine Learning (graduate only, non-advanced)
MILA is a program for graduate students only. MILA representatives note that this ML course is not taken by core ML research students, but by students in other specialties who want to learn ML. Core ML students take more advanced courses. There was no course in 2014.

APPENDIX 7
教職員の多様性

Return to faculty diversity (pg 25)
See underlying data

ソース
教員の多様性に関するデータは、2018年9月21日、AI学科のウェブサイトを介して手動で収集された。選択された大学は、容易にアクセスできる教員名簿を持つコンピュータサイエンス分野を代表する大学である。

手法
教員の性別の多様性の内訳を知るため、学校のウェブサイトから教授の名前を集め、名前と写真の両方を確認し、性別を割り当てた。各学校の詳細については、以下を参照。

カリフォルニア大学バークレー校 - 教員リンクを参照
助教授、教員を含む
スタンフォード大学 - 教員リンクを見る
教職員、研究者および関連する教職員を含む
イリノイ大学アーバナシャンペーン校 - 教員リンクを見る
CS学部の教職員および関連教職員を含む
カーネギーメロン大学 - 教員リンクを見る
記載されている全教職員を含む
ユニバーシティ カレッジ ロンドン - 教員リンクを見る
Peopleリンク下にあるすべての教職員を含む
オックスフォード大学 - 教員リンクを見る
教職員セクションのみを含む
ETHチューリッヒ - 教員リンクを見る
タイトルに「Dr.」が含まれている教職員だけを含む。

ニュアンス

- 学校のデータは2018年9月21日のものである。学校の教員は、2018年のAI Indexが発表されるまでに変更される可能性があります。
- データは学校のAI教職員名簿から取得されたもので、客員教授や他の学科に在籍する教授を加味したものではない。同様に、たとえ AI関連教授が別の学部にも属していたとしても、 AI学部の正職員として挙げられている教授として数える。
- 教授の名前と写真を使い、性別を割り当てた - そのことに対し、AI Indexは誤解を与えている可能性があるが、意図としては、性別について誤った仮定をすることではなく、 AI学界における性別表示の現状を示すことである。
- 学界分野での性別表示は、AI業界全体の性別表示を意味する訳ではない（実際、 業界の女性の割合 はかなり低いことが多いが）
- この指標は現状を写したものであり、時系列的な改善を表すものではない。スタンフォード AIの教員であるDr. James A. Landayの発言には以下。

「私たちはより多様な教職員を採用することに注力している。教職員リストに記載されている女性のほとんど過去2 - 3年で採用された職員であることから、多様性は改善してきている。」
- Dr. Landay, Stanford University

APPENDIX 8
参加者数

Return to participation (pg 26)
See underlying data

ソース
会議出席データは会議 /組織の代表者から直接収集したものである。データは以下の会議から収集した。

AAAI — Association for the Advancement of Artificial Intelligence
AAMAS — International Conference on Autonomous Agents and Multiagent Systems
AI4ALL
ACL — Association for Computational Linguistics
CVPR — Conference on Computer Vision and Pattern Recognition
ICAPS — International Conference on Automated Planning and Scheduling
ICLR — International Conference on Learning Representations
ICML — International Conference on Machine Learning
ICRA — International Conference on Robotics and Automation
IJCAI — International Joint Conferences on Artificial Intelligence
KR — International Conference on Principles of Knowledge Representation and Reasoning
NeurIPS — Conference on Neural Information Processing Systems
UAI — Conference on Uncertainty in Artificial Intelligence
WiML — Women in Machine Learning

手法
2018年に2,000人以上が出席した会議を大規模会議として定義し、 2018年に2000人未満の会議を小規模会議と定義している。本レポートで選出された会議は、 AI研究をリードし、年間出席データも提供可能であった会議である。

AIにおける包括性に関する進歩とデータの利用可能が観点から、 AI4ALLとWiMLを選択した。今後の報告にマイノリティグループを代表する組織 /会議を追加することを考えている。

ニュアンス
会議ごとの固有のニュアンス

- 一部の会議開催者は出席の推計のみの提供だった - 今回は推測を正確なもの受け入れた
- 一部の会議は年に一度開催される訳ではなく、何年も開催されなかった会議もある。
- いくつかの会議の代表者は、会議の開催地は 1年以上前に決定されるため、出席の上限が制限されることが多いと述べている。したがって、会議出席者数は必ずしも需要を反映している訳ではない。

AI4ALL / WiMLに固有のニュアンス

- AIにおける包括性を支える公式および非公式のプログラムが他にもいくつか存在することを理解することが重要である。
- AI4ALL / WiMLへの参加は必ずしもその分野における女性や過少代表グループの増加を示すものではない。

APPENDIX 9
ロボットソフトウェアダウンロード

Return to robot software downloads (pg 29)
See underlying data

ソース
ROS.orgのページビューとロボットソフトウェアのダウンロードに関するデータは、 ROS.orgコミュニティ指標レポートから引用した。 詳細は年別のPDFを参照。ROS.orgについてはこちらを参照。

ニュアンス

- ROS.orgのページビューは世界全体からのページビュー合計を表しているのではなく、上位国のページビューを表しているものとなっている。 そのため、90以上の地域情報は含まれていない。
- 2018年のROS.orgに関する指標はROS.orgから入手可能だが、それらはわずかに異なる手法を使用しているためAI Indexには含まない。 詳細はスレッドを参照。
- 月平均は、毎年7月のデータをもとに計算されている（ 2012年と2013年は、それぞれ6月と8月を使用）。
- ROS.orgの代表者は、英語以外の言語を話す国々でのロボットソフトウェアに関する活動は、英語のWebサイトのみを考慮ことから、過小評価されている可能性があると指摘している。
- ROS.orgの代表者は、2015年の数値の下落はサンプリングの問題と該当月のパッチリリースが少なかったことが原因であり、その結果がダウンロード統計に反映されていると推測している。このことに対する説明指標はこちらを参照。

APPENDIX 10
スタートアップ / 投資

Return to startups / investments (pg 31)
Underlying data is not available.

ソース
Sand Hill EconometricsはDow Jones VentureSourceのデータベースを使用し、スタートアップと VCの投資データを提供している。 SandHillデータベース内の AI企業のリストを確認するため、Crunchbase APIが使用された。Sand Hill、VentureSource、およびCrunchbaseの詳細は、リンクを参照。

Methodology
以下のキーワードを使用し、Crunchbase上のAI関連企業のリストを取得した。

Category Name	Crunchbase Category UUID
Artificial Intelligence	c4d8caf35fe7359bf9f22d708378e4ee
Machine Learning	5ea0cdb7c9a647fc50f8c9b0fac04863
Predictive Analytics	ca8390d722c65bb5f87022f52f364b1b
Intelligent Systems	186d333a99df4a4a6a0f69bd2c0d0bba
Natural Language Processing	789bbbefc46e1532a68df17da87090ea
Computer Vision	a69e7c2b5a12d999e85ea0da5f05b3d3
Facial Recognition	0b8c790f03bcb2aba02e4be855952d6d
Image Recognition	af9307c9641372aeaac74391df240dd2
Semantic Search	bb1777e525f3b9f2922b0d5defe0d5bb
Semantic Web	ac34c4c6e430f66f44aeef8ca45b52bb
Speech Recognition	24dad27e6a49ccc1ec31854b73d60d16
Text Analytics	e7acbc9932c74c8809ed8fae63eb2559
Virtual Assistant	2eb0b56b6896e0f0545fc747a9ba6751
Visual Search	ad8f33a8c5d0b7d786f389aa2954c119

Crunchbase APIは4,477のAI企業名を返した。ここで得た組織リストは、 VentureSourceデータベース内のすべてのベンチャーキャピタルから支援を受けている企業のリストと照合され、 Crunchbaseの4,477社のうち1,096社が、VentureSourceのAI企業のリストと一致した。 Sand Hillは、一致したリストを集計し、会社の数と経時的に見る調達資金額を表わした。

「All Startup」に関するデータは、VentureSourceデータベースのみを使用して収集された - Crunchbase APIは使用されていない。

ニュアンス
- VentureSourceのデータはVentureSource独自のものであり、一般公開はしていない。
- VentureSourceプロダクトの利用は、すべて Sand Hill Econometrics上で行われた。
- 企業情報は、上場や買収、廃業などが起きた場合、 VentureSourceデータベースから削除されるが、削除前の資金調達ラウンドに関する情報はデータベースに残る。
- CrunchbaseとVentureSourceのどちらを使っても、AIシステムを開発している企業とそれを展開しているAI企業を見分けるのは困難である。内部的には、これらの AIを「メーカー」と「ユーザー」と名付け、VentureSourceで集計した企業のうち 109社（これはVentureSourceの全セットのうちの 10%です）のランダムサンプルをテストした。その結果、そのうちの 68社（または62%）がAI「メーカー」であることがわかった。サンプルの残りの 38%の企業は、AI企業と見なされているが、AIシステム自体を開発していない可能性がある。
- 現在のCrunchbaseのAIカテゴリのリストは、機械学習技術を中心に作られている。
- 活動しているAIスタートアップの数は、毎年 1月のものとなっている。

APPENDIX 11
AI関連職

Return to jobs (pg 33)
See underlying data

ソース

AIスキルによる求人の図（ 33ページ）とAI関連職への応募者の性別の多様性の図（ 34ページ）は、それぞれMonster.comの求人サイトから提供されたものである。 Monster.comは結果を得るためにオンラインタレントマーケットインテリジェンスポータルである Garner TalentNeuronを使った。 詳細はGarner TalentNeuronまたはMonster.comのウェブサイトを参照。

手法

Monster.com

Monster.comはGarner TalentNeuronデータベースを使用して以下の用語を検索した。用語が求人掲載に必須スキルとして記載されている場合、その仕事を集計に含まれる。各用語をキーワード「 Artificial Intelligence」と組み合わせ検索した。たとえば、 Roboticsは、就業要件にArtificial IntelligenceとRoboticsの両方の用語が含まれている場合にのみカウントされる。 Monster.comで検索した求人情報はアメリカ国内のものに限定されている。

Artificial intelligence +
Natural Language Processing, Machine Learning Techniques, Deep Learning, Computer Vision, Speech Recognition, Robotics

ニュアンス

- 仕事に複数のAI関連の専門性が含まれている場合、求人は二重にカウントされる。一部の専門性は、実際の説明には記載されていない他の専門性を必要要件とすることを暗示している可能性がある。

AI関連職の求職者における性別の多様性に係るニュアンス

- 応募者数は、必ずしも採用数を意味するわけではなく、業界全体を代表するものでもない。
- 性別は、男女に限定して計測しているが、将来的にはその他の性別も計測に加えていきたいと考えている。

APPENDIX 12
特許

Return to patents (pg 35)
See underlying data

ソース
特許データは、特許検索サービス amplifiedによって提供されたものである。

手法
手法と基礎となるデータに関しては amplifiedが詳述している。

ニュアンス
2015年から2017年までの特許数はデータが完全に出揃っていないため、図に示されていない（出願から公開までの期間が長いため）。

AI特許を長期間にわたって、また地域を超えて追っていこうとする試みはいくつかあるが、確立された手法は存在しない。そのため、AI Indexにおける特許データは他の組織のものとは異なる場合がある。

その顕著な例が、2018年夏に発表された 清華大学による中国 AI開発報告書 で、そのレポートでは、中国が最も多くの公開特許を保有しており、次に米国と日本が続く形になっている。だが、この結果を得るためにどの特許コードと定義が使用されたかは不明であり、各地域別の実際の特許数は図には容易には反映されない。

特許分野の手法は定義付けが非常に難しいため、 AIインデックスはこの分析で使用されている特許コードとキーワードについて完全な透明性のもと、公開している。 AI Indexの読者には、同じ特許数計測手法を活用することを勧める。

APPENDIX 13
AI導入に関する調査

Return to survey section (pg 36)
See underlying data

ソース
この調査は、McKinsey＆Company（McKinsey）によって分析されたものである。マッキンゼーの AI導入調査の結果はこちらを参照。

手法
McKinseyによる手法に関する発表：
調査はオンライン上で、2018年2月6日から2018年2月16日の間に行われた。2,135人の参加者（20,466人中2,135人参加）から回答が集まった。 参加者はあらゆる地域、業界、企業規模、および職業上の専門性を代表するMcKinsey Online Executive Panelの在籍者である。そのうち 1,657人は、回答の選択肢として挙げられた 9つの会社機能、または事業部門の 1つ以上で、少なくとも 1つの人工知能機能を試験運用していると述べた。全参加者は、McKinsey.comの登録ユーザーで、McKinseyの独自リサーチに参加することを選択したあらゆるレベルの役職、管理職、および従業員を代表するオンラインパネルのメンバーである。サンプルには 108カ国が参加している。各国ごとの回答率の違いを調整するため、データは各回答者の所属国がグローバル GDPに占める割合に合わせて重み付けされている。

McKinseyは、AI Indexに対し、以下の質問への選択肢ごとの回答者の割合を提供した。

Q1. 次のAI機能（ロボットプロセス自動化、機械学習、仮想エージェントまたは会話型インターフェイス、コンピュータビジョン、自然言語テキスト理解、自然言語音声理解、自然言語生成、物理的ロボット工学、自律走行車）を現在の組織でどう使用しているか説明してください。

1. 一つ選択してください：
2. まったく使われていない
3. 少なくとも1つの機能または事業単位で試験運用
4. 少なくとも1つの機能または事業単位の標準ビジネスプロセスに組み込まれている
5. 複数の機能または事業単位にわたる標準的なビジネスプロセスに組み込まれている
6. わからない

Q2。あなたの所属する組織では、次のうちどの事業部門で AIを展開していますか？ （該当するものをすべて選択してください）

1. マーケティングと販売
2. 戦略とコーポレートファイナンス（資本配分の決定など）
3. リスク（リスク管理、詐欺、債務）
4. 人事
5. 製品やサービスの開発
6. サプライチェーンマネジメント
7. 製造業
8. サービス業務（フィールドサービス、カスタマーケア、バックオフィスなど）
9. 上記のどれでもない
10. わからない

Q1では、McKinseyに回答を地理的に分割するよう依頼した。その後、 AI Indexは、少なくとも1つのAIを会社機能または事業単位に埋め込んでいる回答者の割合を取得するため、回答 3と4を合計した。マッキンゼーは Q2の回答を地域別に分割し、そして業界別に分割した。 Q1とは異なり、Q2は、自社にAI機能を組み込んだ、または試験運用した回答者に限定されていた。関連する回答者数は各図の下のメモにあります。

ニュアンス
調査回答は、組織の AI導入に対する回答者の認識によって制限されている。

-

APPENDIX 14
企業の取り組み：企業収益報告

Return to earnings calls (pg 40)
See underlying data

ソース
記載されている収支報告に関するデータは、テキスト分析とセンチメント分析を専門とする投資調査会社のPrattleにより収集された。 Prattleは、FactSetからの収支報告記録を使用している。

手法
Prattleには、人工知能と機械学習という言葉についての収支報告で言及した回数を経時的なデータと、ベンチマークのためのクラウドとビッグデータという用語の同じデータについて提供していただいた。結果は各企業の業界と規模が含まれるものとなった。

通話記録から単語を引き出すために、 Prattleは正規表現の変種（正規表現とも呼ばれる）を使用し、各テキストからAIに関するディスカッションで使用される文字列と言語パターンを検索した。人工知能および機械学習という用語には、それらの略語 AIおよびMLも含まれている。 Prattleはそれから誤検知を取り除いた（例えば、 MLはミリリットルの略語でもあるので、Prattleは用語の前に数字がある MLのカウントを取り除いた）。

セクターの分類は、S＆Pが使用する標準的なグループである GICSを使用して行われた。サイズの分類（これらはレポートでは使用されていませんが、基礎となるデータで使用可能）は、時価総額の分類に基づいている。Prattleが提供するしきい値は以下のとおり。

```
mkt_cap <= '500000000' THEN 'micro' :: text
mkt_cap <= '4000000000' THEn 'small' :: text
mkt_cap <= '20000000000' THEN 'mid' :: text
mkt_cap IS NOT NULL THEN 'large'::text
else Null
```

データの分析はAI Indexによって行われた。

ニュアンス
- We observe that the mentions of both Big data and Cloud have steep initial curves, then trail off. This can be interpreted as the normalization of the technology.
- 言及は必ずしも企業の AIへの投資やAIの導入を表している訳ではない。
- この分析には上場企業のみが含まれる。

Nuance
- We observe that the mentions of both Big data and Cloud have steep initial curves, then trail off. This can be interpreted as the normalization of the technology.
- Mentions are not necessarily a proxy for company investment in or adoption of AI.
- Only public companies were included in this analysis.

APPENDIX 15
ロボット導入

Return to robot installations (pg 41)
See underlying data

ソース
データは、国際ロボット連盟（ IFR）2014、2015、および2017年のInternational Federation of Robotics'(IFR) から直接引き出されたものである。 詳細は下記のレポートへのリンクを参照。 IFRの詳細はこちら。

2012 robot import data
Source: 2014 World Robotics Report — Executive Summary (Table 1)
2013—14 robot import data
Source: 2015 World Robotics Report — Executive Summary (Table 1)
2015—16 robot import data
Source: 2017 World Robotics Report — Executive Summary (Table 4.1)

手法
データは、国別に輸入された産業用ロボットの台数である。 産業用ロボットはISO 8373：2012規格で定義されている。

IFRの手法に関する詳細情報は こちらを参照。

ニュアンス
- どの程度の割合のロボットが「 AI」として分類されるソフトウェアを実行しているかを特定する方法は定まっておらず、AIの発展が産業用ロボットの活用にどの程度寄与しているかは不明だ。
- この指標は、2017年版ではロボット輸入量として表されていた。

APPENDIX 16
GitHub Stars

Return to GitHub (pg 42)
See underlying data

ソース
Google BigQueryに保存されているGitHubアーカイブを使用。

手法
レポートの図は、さまざまな GitHubリポジトリの Star獲得数を経時的に示している。リポジトリは次のとおり。

apache /incubator-mxnet、BVLC/cafe、cafe2/cafe2、dmlc/mxnet、fchollet/keras、Microsoft/CNTK、pytorch/pytorch、scikit-learn/scikit-learn、tensorflow/tensorflow、Theano/Theano、Torch/Torch 7

GitHubのアーカイブデータは Google BigQueryに保存されている。関心のある各リポジトリの「 WatchEvents」の数を数えるために Google BigQueryとインターフェースで連結した。 2016年のデータを収集するためのサンプルコードは以下。

```
 SELECT
        project,
        YEAR(star_date) as yearly,
        MONTH(star_date) as monthly,
        SUM(daily_stars) as monthly_stars
FROM (
SELECT
        repo.name as project,
        DATE(created_at) as star_date,
        COUNT(*) as daily_stars
FROM
        TABLE_DATE_RANGE(
                [githubarchive:day.],
                TIMESTAMP("20160101"),
                TIMESTAMP("20161231"))
WHERE
        repo.name IN (
                "tensorflow/tensorflow",
                 "fchollet/keras",
                "apache/incubator-mxnet",
                 "scikit-learn/scikit-learn",
                "cafe2/cafe2", "pytorch/pytorch",
                 "Microsoft/CNTK", "Theano/Theano",
                "dmlc/mxnet", "BVLC/cafe")
        AND type = 'WatchEvent'
GROUP BY project, star_date
)
GROUP BY project, yearly, monthly
ORDER BY project, yearly, monthly
```

ニュアンスは次のページを参照

APPENDIX 16
GitHub Stars

ニュアンス
GitHub Archiveは、現在のところ、ユーザーがリポジトリから Starを削除したときにstarの数を数える方法を提供していない。したがって、報告されたデータはわずかに星の数を多く見積もっている。 GitHub上のリポジトリの実際のStars数と比較すると、その数はかなり近く、傾向は変わっていないことがわかる。

GitHubのスターデータを取得する方法は他にもある。 結果をスポットチェックするために、star-history toolが使用された。

GitHubプロジェクトのフォークを調査するのも興味深いが、リポジトリ StarsとForksの傾向がほぼ同じであることを発見した。

APPENDIX 17
社会の関心

Return to public interest (pg 43)
See underlying data

ソース
ここでの全データは TrendKiteから取得した。 TrendKiteサービスは一般的なメディア記事を索引付けし、記事を「ポジティブ」、「ネガティブ」、または「ニュートラル」に分類する感情分析分類子を採用している。 TrendKiteの詳細はこちら。

手法
否定的な感情を持つ不適切な量の無関係な記事を紹介したソースを削除するように調整した。 昨年、中立的な記事の一部が暗示されていたが、図には含まれていなかった。 今年は、ニュートラルを含めた。AI関連記事の識別には以下のクエリを使用。

クエリ
”Artificial Intelligence”
AND NOT “MarketIntelligenceCenter.com’s"

NOT site_urls_ll:(
“individual.com"
OR "MarketIntelligenceCenter.com")

フィルター
• Only included English-language articles
• Removed press releases
• Removed financial news
• Removed obituaries

APPENDIX 18
政府によるAIに関する言及

Return to government mentions (pg 44)
See underlying data

国ごとのソース、手法、およびニュアンス:

データ収集と分析は、the McKinsey Global Institute (MGI)によって行われた。

カナダ(下院):
データは、カナダ国会議事堂のウェブサイトの Hansard search 機能を使用して収集された。「人工知能」と「機械学習」(引用符を含む)という用語を検索し、結果を CSVとしてダウンロードした。日付範囲は「すべての討論」に設定されている。データは 11/20/2018時点のものである。データは 2002年8月31日時点のものからオンラインで入手可能。

それぞれの数は、下院の議事録で人工知能または機械学習が特定のコメントまたは発言で言及されたことを示している。つまり、議論の中で、メンバーが発言の中で AIまたはMLを複数回言及した場合、それは 1としてカウントされる。しかし、同じ議論の間に、話者が別のコメントで AIまたはMLに言及した場合(間に他の話者がいる場合)、それは複数回カウントされる。人工知能と機械学習のカウントは別々の検索で行われたため別物としてカウントされている。略語「AI」または「ML」の言及は含まれていない。

イギリス(下院、下院、ウェストミンスター会館、委員会)
英国議会のHansard websiteのFind References機能を使用してデータを収集した。「人工知能」および「機械学習」(引用符を含む)という用語を検索し、結果をカタログ化している。データは、 2010年10月31日時点のもので、データは1/1/1800以降からオンラインで入手可能。 Open Parliament License v3.0の下でライセンス化されている議会情報が含まれている。

カナダのように、それぞれの数は、「人工知能」または「機械学習」が、特定のコメントまたは発言で言及された回数を示す。したがって、あるメンバーが発言内で AIまたはMLを複数回言及した場合、それは 1としてカウントされる。しかし、同じ議論の間に、同じ話者が別々のコメントで AIまたはMLに言及した場合(間に他の話者がいる場合)、それは複数回としてカウントされる。人工知能と機械学習のカウントは別々の検索で行われたため別物としてカウントされている。また、略語「AI」または「ML」の言及は含まれていない。

アメリカ(上院および下院)
データは、US Congressional Record websiteの advanced search機能を使用して収集された。「人工知能」と「機械学習」(引用符付き)という用語を検索し、結果を CSVとしてダウンロードしている。「単語のバリエーション」オプションは選択されておらず、議事録には上院議員、下院、および発言延長が含まれていたが、デイリーダイジェストは含まれていなかった。データは、 2017年11月20日時点のものである。第 104回大会以降(1995年)のデータはオンラインで入手可能。

それぞれのカウントは、議会の議事録に記録された特定の討論の中で、人工知能や機械学習が言及されたことを示している。特定の討論には、請求書の閲覧を含む。ある発言者が発言中に AIまたはMLについて複数回言及した場合、または同じ討論の中で複数の発言者が AIまたはMLについて言及した場合、結果は 1としてカウントされる。人工知能と機械学習のカウントは別々の検索で行われたため別物としてカウントされている。また、略語「AI」または「ML」の言及は含まれていない。

APPENDIX 19
ImageNet

Return to ImageNet (pg 47)

See underlying data: (1) accuracy scores (2) training time

ImageNet精度スコア

ソース

2010～2017年のImageNetの精度に関するデータは、LSRVCリーダーボードを通じて取得された。2016～2018年の検証データセットのパフォーマンスは、arXivの文献レビューを通じて取得。

ImageNetの進歩を視覚的に表すため、以下の論文からスコアを取得した。
ImageNet Classification with Deep Convolutional Neural Networks (2012)
Visualizing and Understanding Convolutional Networks (Nov 2013)
Going Deeper with Convolutions (Sept 2014)
Deep Residual Learning for Image Recognition (Dec 2015)
PolyNet: A Pursuit of Structural Diversity in Very Deep Networks (Nov 2016)
Squeeze-and-Excitation Networks (Sep 2017)
GPipe: Efficient Training of Giant Neural Networks using Pipeline Parallelism (Nov 2018)

人間レベルの性能に関する推定は、Russakovsky et al, 2015.によるものである。LSVRC ImageNetコンペティションおよびImageNetデータセットの詳細はリンクを参照。

手法

ImageNetコンペティションデータ（2010–2017）については、ImageNet Webサイトでホストされている各 LSVRCコンペティションのリーダーボードからスコアを収集した。検証データセットの性能は文献レビューを通して得られた。各年度の最高成績のみが記録されている。

ImageNetのトレーニング時間

ソース

ImageNetのトレーニング時間は、arXivの文献レビューを通して収集された。以下の論文が使用された。

June 8th, 2017: 60 minutes
November 1st, 2017: 30 minutes
November 12th, 2017: 15 minutes
July 30th, 2018: 6.6 minutes
November 13th, 2018: 3.7 minutes

APPENDIX 20

Parsing

Return to parsing (pg 50)
See underlying data

ソース

Wall Street JournalにおけるPenn Treebankの部分は、各センテンスの構成要素ベースの解析ツリーで注釈が付けられたセンテンスのデータセットである。このデータセットのセクション 23は、自動パーサーの研究のための主要なテストセットとなっている。 Penn Treebankの詳細はこちら。

手法

自動パーサーは、自動生成された解析の構成要素とテストセットのゴールド解析の構成要素を比較することによって評価される。生成された構成要素の精度と再現率は結合され、 F1スコアとして報告されます。 報告されるのはPenn TreebankのWSJ部分のセクション23の文に対するパーサーの F1スコアである。長さ40ワード以下の文章と、それぞれが利用可能な文章のセット全体についてこれらのスコアを報告している。選挙区ベースのパースツリーの詳細は ウィキペディアのパースツリーに関するページを参照 。

nlpprogress.comで管理されているリーダーボードに相談し、英語部会の解析を行った。リストされたスコアを確認し、より最先端の結果を見つけるべく
、追加の文献レビューをおこなった。新しいデータポイントは、 2017年のAIインデックスレポートのデータと結合された。特定したパーサーの報告された F1スコアには、単一モデルだけでなくアンサンブルも含まれている。

ニュアンス

自動構文解析研究における初期は、パーサーは通常、計算上および方法論上の理由から、長さ <40語、長さ <100語の文章でのみ評価されていた。利用可能であれば、 40語未満の長さのセンテンスとコーパス内のすべてのセンテンスについてシステムの F1スコアを記録した。

Nuance

- In the early days of automatic parsing research, parsers were typically evaluated only on sentences of length <40 words and length <100 words for computational and methodological reason. We have recorded the F1 scores of systems on sentences of length <40 words and on all sentences in the corpus when available.

APPENDIX 21
機械翻訳 (BLEU)

Return to machine translation (pg 51)
See underlying data

ソース
データはWorkshop on Machine Translation（WMT）コンペティションのための BLEUスコアの指標を提供する EuroMatrixから引き出された。

手法
WMTは毎年ニュースの翻訳作業をホストし、新しいトレーニングとテストデータセットを提供する。参加者のチームは、翻訳作業に参加するために構築した翻訳システムを提出する。

WMTが使用する主な測定基準は、競合するエントリのランク付けを目的としており、前年比の比較はできない。

また、前年比比較は非常に労働集約的である。システムの翻訳と人間が作成した多数の翻訳を大まかに比較する自動方式であるBLEUに戻した。これは、0から1の間の精度検証の修正版である（数値が高いほど高得点）。翻訳ペアのコーパスに対する機械翻訳システムの平均 BLEUスコアを計算することも可能だ。毎年、その年のニュース翻訳タスクへ提出された英語からドイツ語、およびドイツ語から英語へのシステムによって達成された最高平均BLEUスコアを記録しました。

EuroMatrixは2006年以来ニュース翻訳タスクへの投稿の BLEUスコアを英語とドイツ語そしてドイツ語と英語のペアで記録してきた。我々は、特に文章をトークン化するためのプロトコルを定義する BLEU（11b）を使用して、毎年最高のパフォーマンスを発揮するシステムの BLEUスコアを選出した。可能な時は、両方のペアで上位 BLEUスコアを持つシステムのスコアをその年の代表として選出した。

ニュアンス

- BLEUは自動的に計算することができ、それは翻訳品質についての人間の判断と相関することが示されている。ただし、このメトリックをコーパス間で使用することはできず、システム間で BLEUスコアを比較することは誤解を招く可能性がある。
- グラフは全般的に上昇傾向にありますが、 2017年のBLEUスコアが2016年のスコアと比較して大幅に低下したことがわかる（ただし、2017年のスコアは2015年のスコアよりも高い）。MTシステムのパフォーマンスが2016年と比較して低下することはまずないが、ここで提示されている評価スキームは完全ではない。ただし、長期間にわたる傾向を見ると、 BLEUスコアは機械翻訳の分野における進歩の指標となる可能性がある。

APPENDIX 22
ARC

Return to ARC (pg 52)
See underlying data

ソース
AI2推論チャレンジ（ARC）は、Allen Institute for Artificial Intelligenceによって開催されています。 ARCのパフォーマンスデータは ARCのリーダーボードから取得された。easy set とchallenge setのリーダーボードは対応するリンクを参照。

手法
参加者はARCデータセットをダウンロードし、Allen InstituteのWebサイトを通じてリーダーボードに送信する。

Easy developmentコーパスからの質問の例：

> 最近開発された技術はどれか。（A）携帯電話（B）テレビ（C）冷蔵庫（D）飛行機 [Grade 4]
>
> ある学生は、藻は生産者であると仮定している。これが正しいかどうかを判断するのに最も役立つ質問はどれか。（A）藻類は他の生物を消費するか？ （B）どの生物が藻を消費するか？ （C）藻類は日光を使って食物を作るか？ （D）生態系は藻類がなくても生き残ることができるか？ [Grade 8]

チャレンジ開発コーパスからの例：

> JuanとLaKeishaはいくつかのオブジェクトをランプの下に向けて転がす。生徒たちはどのオブジェクトが最も遠くに転がるのかを見たがっている。調査を繰り返すことができるようにするために彼らは何をすべきか？（A）オブジェクトをグループにまとめる。（B）ランプの高さを変える。（C）転がすオブジェクトを選ぶ。（D）調査の詳細を記録する。 [Grade 4]
>
> 高気圧は水が凝縮する可能性がある大気のより寒い地域に空気が上昇するのを防ぐ。高気圧が長期間その地域に留まると、どうなるか。（A）霧（B）雨（C）干ばつ（D）竜巻 [Grade 8]

各質問の正解は 1点とする。モデルは複数の解答を与えることが可能となっている。その場合、 N個の解答を与えるモデルは、その N個の解答の 1つが正しければ 1/Nポイントを、そうでなければ 0点となる。総合スコアは、個々の質問での得点の平均とする。

AI Indexは、以前の投稿から得点を上回る質問のみを収集している。投稿が以前の投稿のどれか一つよりも低い場合は、図に含まれない。

ルールと提出ガイドラインについては こちらを参照。

APPENDIX 23
GLUE

Return to GLUE (pg 53)
See underlying data

ソース
GLUEベンチマークデータは、GLUEリーダーボードから取得した。 GLUEベンチマークの詳細は こちらを参照。

手法
参加者はGLUEタスクをダウンロードし、GLUE Webサイトを通じてリーダーボードに提出する。タスクの個々の測定基準に基づいて、各タスクのスコアが計算されます。全ての測定基準は100倍（すなわち百分率として）で拡大縮小される。これらのスコアは平均されて最終スコアが得られる。複数のメトリクス（ MNLIを含む）を持つタスクの場合、メトリクスは平均化される。

リーダーボードでは、ユーザーのトップスコアの投稿のみが表示またはランク付けされています。他の投稿は、各ユーザーの拡大表示の下で表示可能。競合他社は、結果が表示されないように個人的に送信することが可能。AI Indexの図には、個人の投稿は含まれていない。一致および不一致の MNLIは、スコアリング目的の 1タスクと見なされる。

AIインデックスは、以前の投稿の得点を上回るもののみを収集している。投稿が以前の投稿のどれよりも低い場合、それは図には含まれない。

ルールと提出ガイドラインについては こちらを参照。