AI
artificial
intelligence
index
2018 annual report

AI 인덱스 2018 연례보고서

요아브 쇼함, 레이몬드 페로, 에릭 브린욜프슨, 잭 클라크, 제임스 매니카, 후안 카를로스 니블스, 테라 라이언스, 존 에치멘디, 바바라 그로스,
조이 바우어 지음

이상훈, 이동훈 옮김

운영 위원회

요아브 쇼함(Yoav Shoham)*(위원장)*
스탠포드 대학교

레이몬드 페로(Raymond Perrault)
SRI 인터네셔널

에릭 브린욜프슨(Erik Brynjolfsson)
MIT

잭 클라크(Jack Clark)
오픈AI

제임스 매니카(James Manyika)
맥킨지 글로벌 연구소

후안 카를로스 니블스(Juan Carlos Niebles)
스탠포드 대학교

테라 라이언스(Terah Lyons)
인공지능을 위한 파트너십(PAI)

존 에치멘디(John Etchemendy)
스탠포드 대학교

바바라 그로스(Barbara Grosz)
하버드 대학교

프로젝트 매니저

조이 바우어(Zoe Bauer)

AI 인덱스 2018 보고서

AI 인덱스의 목적은 인공지능에 관한 데이터 기반의 담론을 제시하고자 함이다.

AI 인덱스는 인공지능과 관련된 데이터를 수집, 비교, 정제 및 시각화 하고자 하는 노력의 결과이다. 본 보고서는 정책 입안가, 연구자, 경영자, 언론인 및 일반 대중이 인공지능이라는 복잡한 분야에 대한 이해를 높이는데 도움이 될 수 있는 종합적인 데이터와 분석 자료의 역할을 하고자 한다.

보고서 출처 표기법:

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 스탠포드 대학교, 『AI 인덱스 2018 연례 보고서』는 크리에이티브커먼즈 라이선스 중 저작자표시-변경금지 4.0 국제 (CC BY-ND 4.0)에 따라 이용할 수 있습니다. https://creativecommons.org/licenses/by-nd/4.0/deed.ko

(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

목차

AI 인덱스 2018 연례보고서 소개 5
개요 6
활동량 8
- 연구
  - 출간된 논문 9
  - 수강신청 21
  - 참가 26
  - 로봇 소프트웨어 29
- 산업계
  - 스타트업 / 투자 31
  - 일자리 33
  - 특허 35
  - 인공지능 도입 설문조사 36
  - 분기 실적 38
  - 로봇 설치 41
- 오픈 소스 소프트웨어
  - GitHub 프로젝트 통계 42

대중적 관심
- 언론 보도 동향 43
- 정부 언급 44

기술적 성능
- 시각 47
- 자연어 이해 50

기타 지표
- 파생 지표 55
- 정부 정책 인간 57

수준의 성능 마일스톤 59
보고서에서 다루지 못한 내용 63
감사의 말 66
부록 69

AI 인덱스 2018 연례보고서 소개

AI 인덱스 2018 연례보고서를 내놓게 된 것을 기쁘게 생각한다. 올해 보고서를 통해 두 가지 목표를 달성했다. 먼저 2018년도 보고서에서는 작년 보고서의 측정 항목을 개선하였다. 다음으로 본 보고서는 항상 국제적인 관점을 제시하려고 하였다. 인공지능에 관한 담론을 제시하는 것은 지속적으로 양적·기술적 발전 상황을 추적하는 것을 의미하게 때문에 첫번째 목표는 인덱스의 목적을 달성하는데 핵심적이다. 두번째 목표 또한 필수적이다. 국제적인 관점없이 기술된 인공지능 관련 이야기는 존재하지 않는다. 2017년 보고서는 북미 지역의 인공지능 관련 활동에 초점이 맞추어져 있었다. 이는 AI 인덱스 프로젝트의 제한된 국제적 파트너쉽에 의한 한계였지 내재적인 편향 때문은 아니었다. 2018년 AI 인덱스 팀은 이러한 국제적 관점에 있어서의 한계를 좁히기 시작했다. 우리는 AI 인덱스 보고서가 진정으로 포괄적인 보고서가 되기 위해서는 더 많은 협력과 외부 참여와 함께 장기적인 노력이 필요함을 인지하고 있다.

우리는 이미 인공지능이 전세계와 관련되어 있다고 주장할 수 있다. 스코퍼스(Scopus) 논문 인용 분석 데이터베이스에 따르면 2017년 인공지능 관련 논문의 83%가 미국 외 국가에서 게재되었으며 28%가 유럽에서 게재되었으며 유럽이 단일 지역으로 가장 높은 비율을 보였다. 인공지능과 머신러닝 수업 수강신청 수 또한 전 세계적으로 증가 추세를 보이고 있으며, 특히 중국 칭화 대학교에서 가장 두드러진 증가세를 보였다. 칭화 대학교의 2017년 인공지능 및 머신러닝 수업 수강신청 수 증가율이 2010년과 비교하여 16배 증가하였다. 또한 이러한 증가 추세는 미국, 중국 및 유럽 지역에만 국한되어 나타나는 것에 그치지 않는다. 한국 및 일본은 2014년 인공지능 관련 특허 보유건수에서 미국에 뒤 이어 2위 3위를 차지했다. 또한 남아프리카공화국은 전 세계에서 가장 큰 규모의 머신러닝 학습 컨퍼런스인 제2회 딥러닝 인다바 컨퍼런스(Deep Learning Indaba conference)를 개최하였으며 20개국 이상의 아프리카 국가에서 500명 이상의 참가자가 참여하였다.

인공지능의 다양성은 지리적으로만 국한된 것이 아니다. 오늘날 인공지능을 위한 국제 협력단체인 인공지능을 위한 파트너십(PAI)의 회원 중 50% 이상이 미국시민자유연맹(ACLU), 옥스포드 대학 인류 미래 연구소(Future of Humanity Institute), 유엔 개발 계획(UNDP)등의 비영리 단체이다. 또한 인공지능의 발전에 있어 성별·인종 다양성의 중요성에 대한 인식이 높아지고 있다. 예를 들어 AI4ALL 및 머신러닝 분야의 여성 모임(Women in Machine Learning, WiML)와 같이 소수 집단의 참여를 촉진하는 단체의 참여가 늘어났다.

AI 인덱스 보고서 개요

본 보고서는 4개 섹션으로 구성되어 있다:

1. 데이터: 활동량 및 기술적 성능
2. 기타 지표: 최신 정부 정책, 파생 지표 및 인간 수준의 수행능력
3. 보고서에서 다루지 못한 내용
4. 부록

데이터

활동량 지표는 학계, 기업, 사업가 및 대중의 인공지능 관련 활동량을 측정한다. 양적 데이터는 인공지능을 공부하는 학부생 수부터, 인공지능 관련 일자리에 지원하는 여성의 비율, 인공지능 관련 스타트업에 투자하는 벤처캐피탈 투자액의 증가세까지 모두 포함한다.

기술적 성능 지표는 인공지능의 성능이 어떻게 변화해 왔는지를 측정한다. 예를 들어 데이터 문답의 질과 컴퓨터가 객체를 탐지하도록 훈련하는데 걸리는 시간 등을 측정한다.

2018 AI 인덱스 보고서는 2017년 측정한 로봇 설치 및 인공지능 컨퍼런스 참석자 수와 같은 지표에 국가별 분석을 더하였다. 이에 더해 2018년 보고서에는 특허, 로봇 운영체제의 다운로드 수, 일반적인 언어 이해 평가(GLUE; General Language Understanding Evaluation) 지표 및 COCO(Common Objects in Context Challenge) 리더보드와 같은 새로운 지표를 추가하였다. 전반적으로 2018년 보고서에서는 작년 보고서의 주요 시사점을 지속적으로 제시하고 있다. 인공지능 관련 활동은 거의 모든 분야에서 증가하고 있으며 기술적 성능은 전반적으로 향상되고 있다. 그러나 올해 보고서의 시사점 중 특히 주목할만한 항목이 있었다. 자연어 처리의 상당한 진전 및 교내의 제한적인 젠더 다양성 등이 특히 주목할만하였다.

기타 측정 항목

작년처럼 ***파생 지표 섹션***에서는 지표 간 관계를 분석하였다. 또한 AI 활성 지수(AI Vibrancy Index)라는 실험적인 지수를 다루었다. AI 활성 지수는 학계와 산업계의 트렌드를 결합해 인공지능이 하나의 분야로서 얼마나 활성화되어 있는지를 측정한다.

올해 우리는 ***최신 정부 정책***이라는 새로운 질적 지표를 보고서에 추가하였다. 해당 지표는 인공지능과 관련된 최근 정부 투자를 단순화하여 제시한 것이다. 지표는 미국, 중국 및 유럽 국가들의 정책을 포함하고 있다. AI 인덱스는 더 많은 기관과의 협업을 통해 보다 다양한 정부 데이터와 분석을 추후 보고서에 추가하고자 한다.

인간 수준의 수행능력은 인공지능이 인간이나 그 이상의 능력을 보인 예를 시간 순으로 나열한 것을 바탕으로 한다. 올해 보고서에는 2018년의 성과를 추가하였다.

AI 인덱스 보고서 개요 (계속)

마지막으로 인공지능 커뮤니티에서 이러한 논의를 시작하기 위해, ***보고서에서 다루지 못한 내용*** 섹션에서는 AI 인덱스가 보다 종합적이고 정확한 보고서로 거듭 날 수 있도록 인공지능 분야 전문가들이 제안하는 아이디어들이 기술되어 있다.

부록

부록에서는 출처, 데이터 수집 방법 및 뉘앙스에 관해 투명하게 서술하고 있다. 본 보고서의 부록은 보고서에서 다룬 거의 모든 그래프에 대한 데이터를 제시한다. 우리는 인공지능 커뮤니티에 소속된 관련자들이 그들의 업무 및 관심사와 가장 밀접하게 관련된 데이터를 활용하게 되기를 바란다.

상징

AI 인덱스 팀은 인공지능의 보편성에 대해 다루는 페이지에는 이하와 같은 지구본 표시를 하였다. 보편성에서는 국가별 비교, 미국 외 지역에 대한 분석 및 인공지능 관련 커뮤니티의 다양성에 관한 데이터 등을 포함하고 있다.

활동량

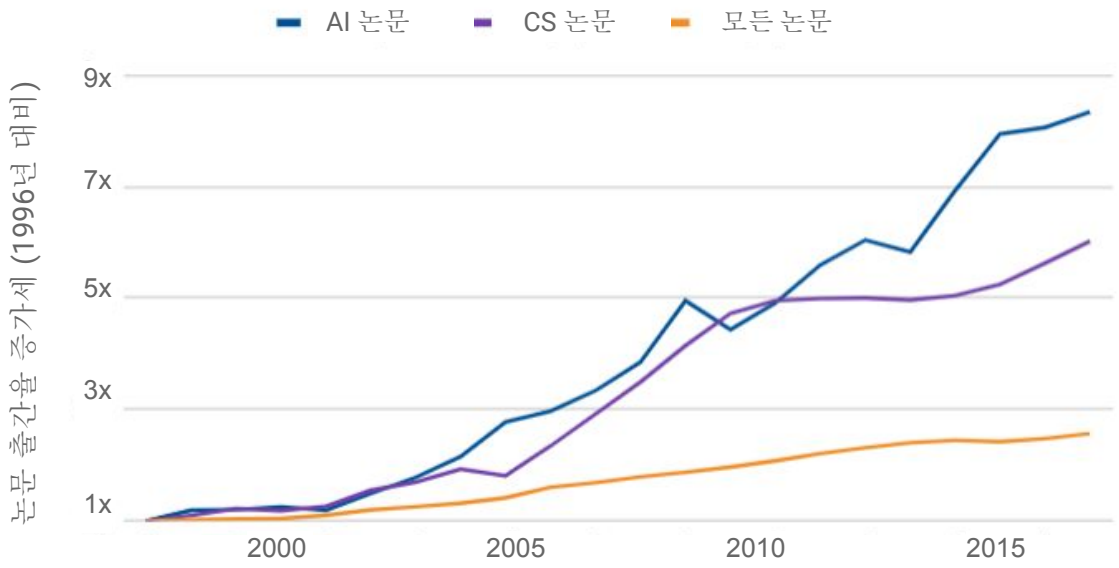
출간된 논문: 주제별 분류

이하 그래프는 1996년 대비 연간 학술 논문 출간율의 증가세를 보여준다. 해당 그래프는 모든 분야, 컴퓨터 과학(CS), 인공지능(AI) 분야에서의 논문 출간율을 비교한다.

인공지능(AI) 분야에서의 연간 논문 출간율이 지속적으로 컴퓨터 과학 분야의 연간 논문 출간율을 상회하는 경향을 보인다. 이는 인공지능 분야의 논문 출간율이 컴퓨터 과학 분야에 대한 높은 관심뿐만 아니라 보다 많은 요인에 기인한 것일 수 있음을 이야기한다.

데이터 및 수집 방법론은 부록 1 참고

주제별 연간 논문 출간율 증가세(1996–2017)
출처: Scopus



노트: 위 그래프는 Elsevier 키워드 검색이 아닌 출간된 학술 논문 데이터베이스 Scopus에 "인공지능"을 질의로 검색한 결과이다. 추가 정보는 부록 참고.

AI 관련 논문 출간율 CS 관련 논문 출간율 상회

Scoups에 등재된 인공지능 관련 논문 출간율은 1996년 대비 7배 증가하였다. 같은 기간 컴퓨터 과학 관련 논문 출간율은 5배 증가하였다.

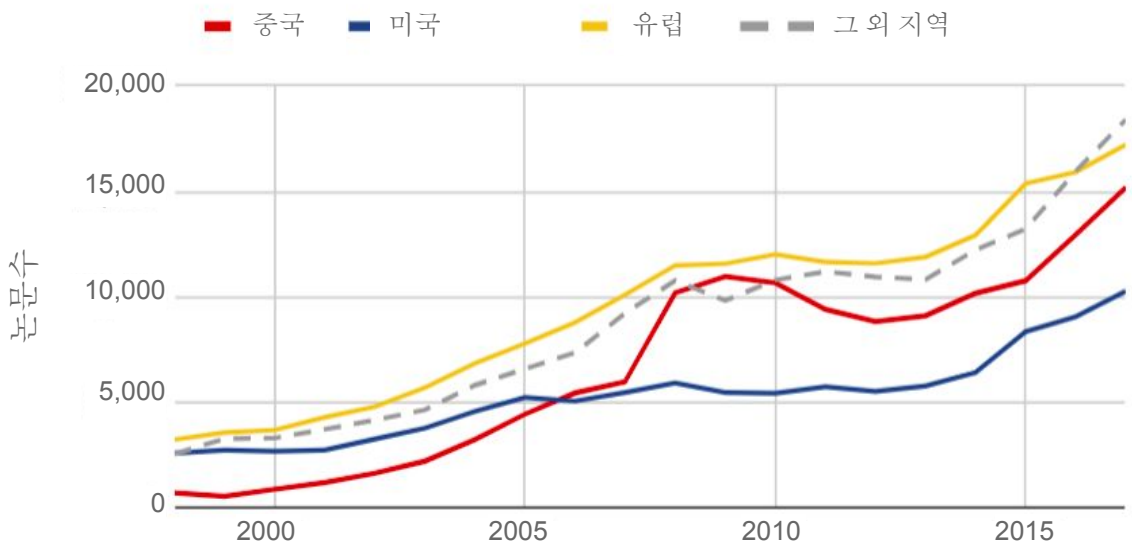
출간된 논문: 인공지능 관련 논문 지역별 분류

이하 그래프는 지역별 연간 인공지능 관련 논문 출간율을 보여준다.

유럽이 지속적으로 인공지능 관련 논문의 최대 출간율을 보인다. 2017년 기준 Scopus에 등재된 논문 중 28%가 유럽에서 출간된 논문이었다. 반면 중국에서는 2008년 전후로 인공지능 관련 논문 출간율이 급등락을 보였음에도 불구하고 인공지능 관련 논문 출간율이 2007년부터 2017년까지 150% 증가하였다.

데이터 및 수집 방법론은 부록 2 참고

Scopus에 등재된 인공지능 관련 논문 출간율 지역별 증가세(1998–2017)
출처: Elsevier



주석: 우리는 2008년 전후 중국에서 인공지능 논문 출간율이 급등한 것이 과학기술 발전에 관한 중장기 계획 (The National Medium- and Long-Term Program for Science and Technology Development (2006– 2020) 및 이와 유사하게 인공지능 연구에 투자 및 인센티브를 제공하는 다양한 정책에 따른 결과라고 추측한다. 이와 유사하게 유럽에서는 2008-2010년 나타난 약간의 증가세가 제7차 프레임워크 프로그램(FP7 (2007–2013)) 및 기타 과학기술 연구 프로그램의 영향일 수 있다.

유럽, 인공지능 관련 논문의 최대 출간율 기록

2017년 유럽 연구진이 출간한 논문이 Scopus에 등재된 논문 중 28%를 차지하였으며 중국(25%), 미국(17%)이 이를 뒤이었다.

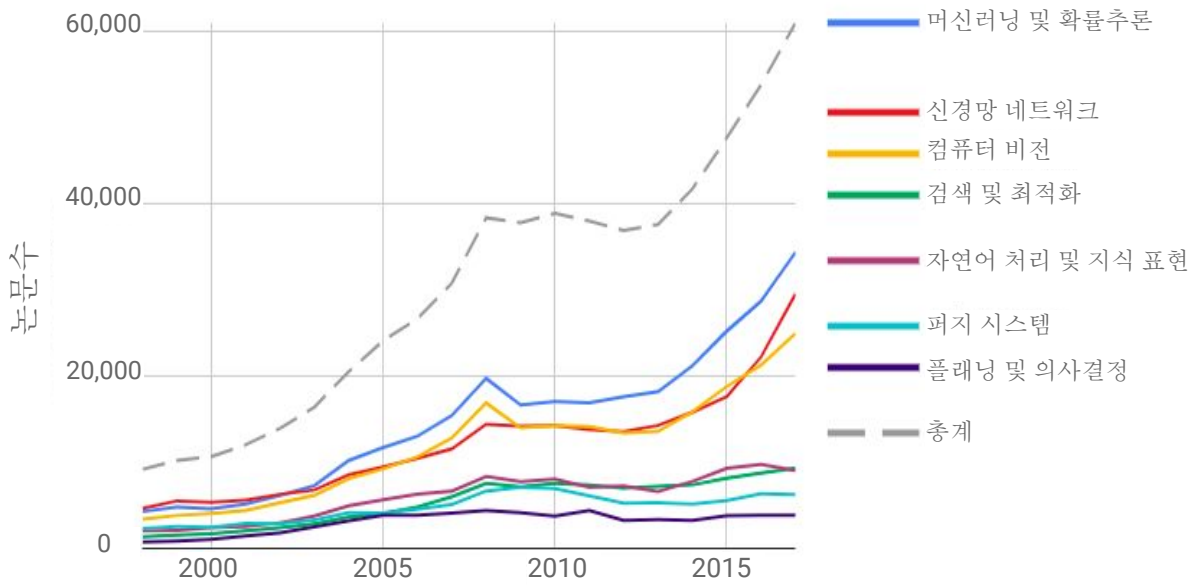
출간된 논문: 세부분야별 분류

이하 그래프는 Scoups에 등재된 인공지능 관련 논문 수를 세부분야별로 분류한 것이다. 범주는 상호 배타적이지 않다.

2017년 Scopus에 등재된 인공지능 관련 논문 중 56%는 머신러닝 및 확률추론(Probabilistic Reasoning) 분야였으며 2010년에는 해당 분야의 논문이 28%를 차지하였다. 이하 분석한 분야 중 대부분은 2010-2014년과 비교하여 2014-2017년에 더 빠른 증가율을 보였다. 가장 주목할만한 분야인 신경망 네트워크(Neural Networks) 같은 경우에는 2010-2014년 3%였던 연평균 논문 출간 증가율이 2014-2017년 37%로 증가하였다.

데이터 및 수집 방법론은 부록 2 참고

Scopus에 등재된 인공지능 관련 세부분야별 논문수(1998—2017)
출처: Elsevier



2014-2017년 Scopus에 등재된 논문 중 신경망 네트워크(Neural Networks) 관련 논문 연평균 증가율 37%기록

출간된 논문: arXiv에 수록된 논문

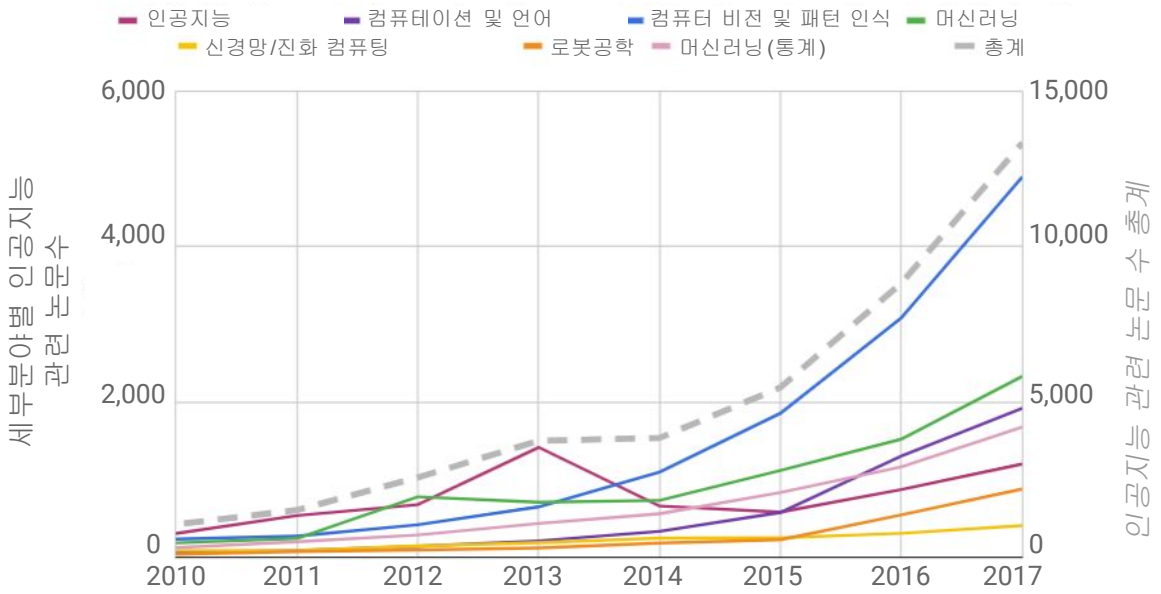
이하 그래프는 arXiv에 수록된 인공지능 관련 논문을 각 논문의 주요 세부분야(primary subcategory)로 분류한 것이다. 오른쪽 축은 arXiv에 수록된 인공지능 관련 논문 수의 합을 나타낸다(회색 점선 그래프).

arXiv에 수록된 인공지능 관련 논문의 수는 여러 세부분야에서 전반적으로 증가하고 있다. 이는 인공지능 관련 논문 저자들이 동료평가 과정이나 인공지능 컨퍼런스에 출간하는 것과 무관하게 논문을 공개하는 경향을 보여준다. 여기서 우리는 인공지능 분야의 경쟁적 환경을 엿볼 수 있다.

컴퓨터 비전(Computer Vision) 및 패턴 인식(Pattern Recognition)은 2014년 이래로 arXiv에서 가장 큰 부분을 차지하는 세부분야였다. 2014년 전에는 해당 범주의 성장률은 인공지능 및 머신러닝과 유사한 성장률을 보였다. 컴퓨터 비전 (및 컴퓨터 비전의 일반적인 응용)에 대한 증가하는 관심을 보여줄 뿐만 아니라 계산, 언어 및 로봇공학과 같은 여타 인공지능 응용 분야의 성장을 보여준다.

데이터 및 수집 방법론은 부록 3 참고.

arXiv에 수록된 세부분야별 인공지능 관련 논문수(2010—2017)
출처: arXiv



"...출간율 증가뿐만 아니라, 인공지능 커뮤니티에서 연구 결과를 출간하는데 arXiv를 사용한다는 점이 눈여겨 볼만하다. 우리는 상당한 논문이 모이면 커뮤티니 내 참여율 증가를 촉진하는 경우를 많이 목도하였다...."

폴 긴스파그(Paul Ginsparg), 코넬 대학교

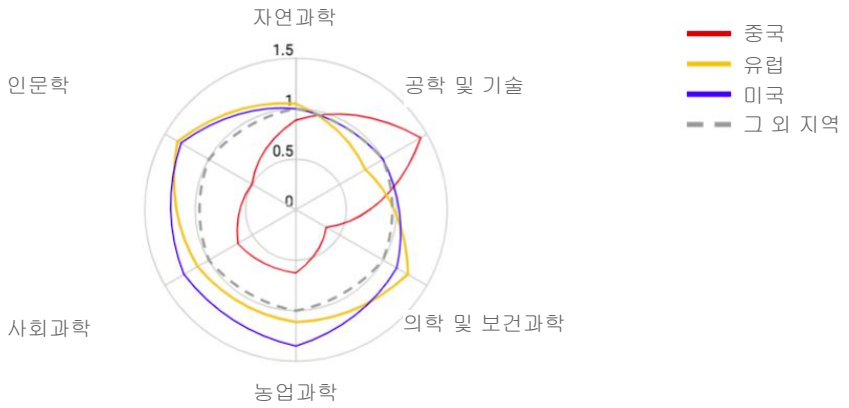
출간된 논문: 지역별 주요 활동분야

이하 그래프는 미국, 유럽 및 중국의 상대적 활동지수(Relative Activity Index, RAI)를 보여준다. RAI는 인공지능 분야에 대한 국제적 연구 활동과 비교하여 각 지역의 분야별 특화 정도를 근사치로 나타낸 수치이다. RAI는 전세계 인공지능 관련 논문 출간율 대비 각 국의 인공지능 관련 논문 출간율의 비중으로 정의한다. 수치 1.0은 한 국가의 인공지능 관련 연구가 인공지능 관련 국제적 연구와 동일한 비중을 차지함을 나타낸다. 1.0보다 높은 수치는 한 국가에서 해당 분야 연구의 중요성이 국제적 추세보다 높다는 의미인 반면 1.0보다 낮은 수치는 한 국가에서 해당 분야 연구의 중요성이 국제적 추세보다 낮음을 의미한다.

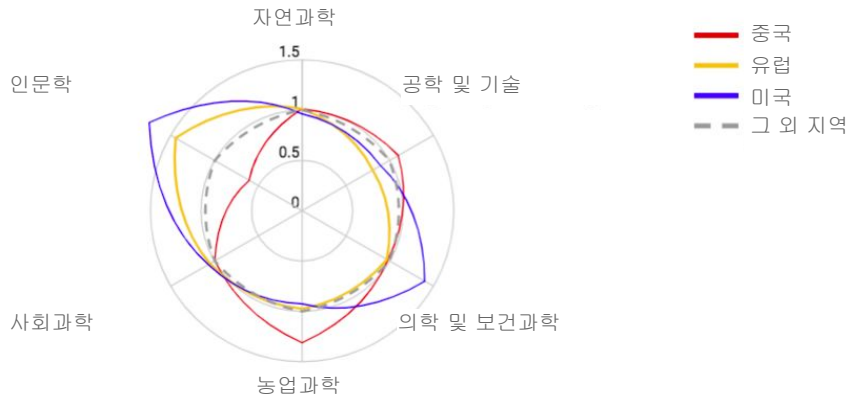
중국에서 인공지능 관련 논문은 공학 및 기술, 농업과학에 더 초점이 맞추어져 있는 반면, 미국 및 유럽의 인공지능 관련 논문은 인문학, 의학 및 보건학과에 더 초점이 맞추어진 경향이 있다. 2000년 데이터 대비 2017년 데이터는 세 지역 모두가 특정 분야에 더 특화되는 모습을 보이며 중국에서 농업과학으로 인공지능 관련 연구의 초점이 옮겨가는 것을 보여준다. 이러한 경향은 중국이 가장 높은 식량 생산량을 가진 국가이며 중국에서는 응용 인공지능에 초점을 맞추어 연구가 이루어진다는 것을 바탕으로 한 우리의 예상과 일맥상통한다.

데이터 및 방법론은 부록 2 참고.

지역 및 인공지능 연구 분야별 상대적 주요 활동분야(2000)
출처: Elsevier



지역 및 인공지능 연구 분야별 상대적 주요 활동분야(2017)
출처: Elsevier



출간된 논문: 연계부문별

본 페이지 및 다음 페이지의 그래프는 Scopus에 등재된 논문 중 정부, 기업 및 의료기관과 연계된(affiliated) 논문의 수를 나타낸다. 또한 16페이지에서는 세 지역을 비교하였다. 첫번째 그래프는 기업부문과 연계된 논문 출간 증가율을 나타내며 두번째 그래프는 정부부문과 연계된 논문 출간 증가율을 나타낸다.

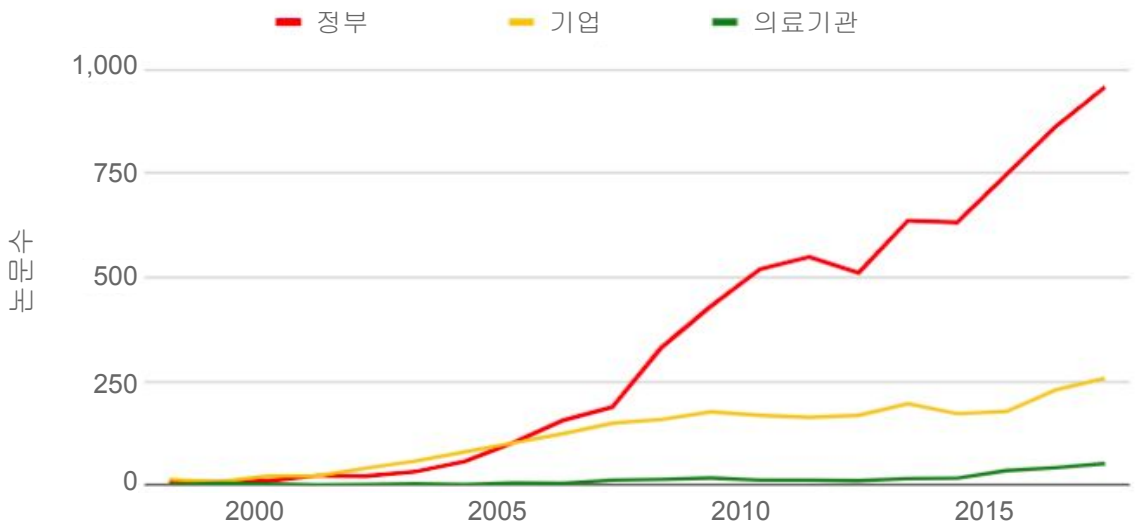
2017년 중국 정부와 연계된 논문의 수가 중국 기업과 연계된 논문 대비 약 4배 이상이었다. 또한 중국은 2007년 이래 정부와 연계된 인공지능 관련 논문의 수가 400% 증가한 반면 기업과 연계된 인공지능 관련 논문은 동 기간 동안 73% 증가하는데 그쳤다.

미국에서는 상대적으로 기업과 연계된 논문이 더 많은 비중을 차지했다. 2017년 미국에서 기업과 연계된 논문의 비중은 중국에서 기업과 연계된 논문과 비교하여 6.6배 높았으며 유럽과 비교하여 4.1배 높았다.

세 지역 모두에서 학계와 연계된 논문의(그래프 없음) 수가 정부, 기업 및 의료기관과 연계된 논문의 수 보다 많았다. 각 그래프 하단의 설명을 참고하길 바란다.

데이터 및 방법론은 부록 2 참고.

출간된 논문: 연계부문별 — 중국(1998—2017)
출처: Elsevier



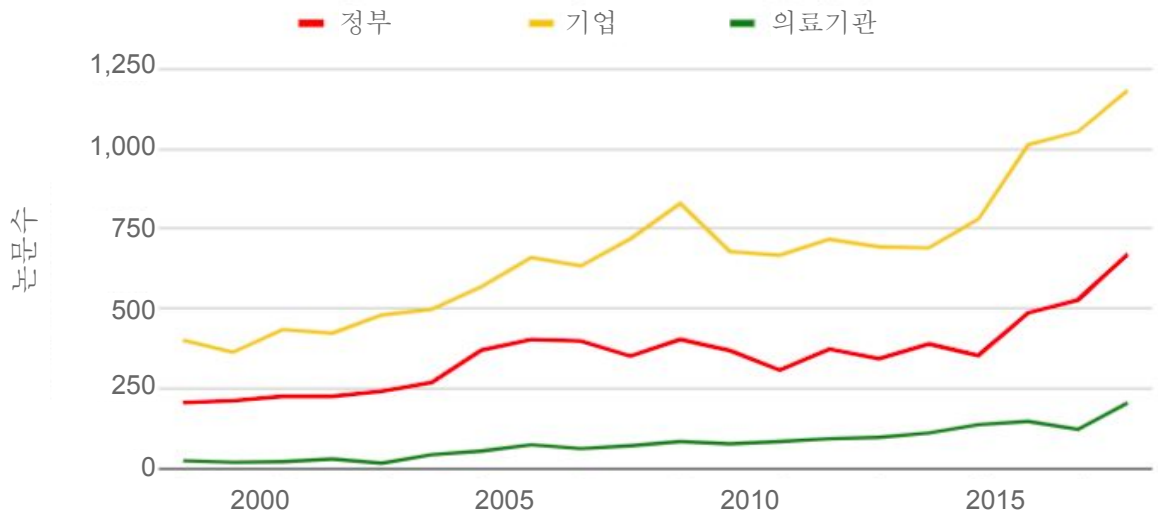
주석: 예를 들어, 2017년 Scopus에 등재된 인공지능 관련 논문 중 중국에서 학계와 연계된 논문의 수는(그래프 없음) 약14,000편으로 중국에서 출간되는 인공지능 관련 논문의 약92%를 차지했다.

중국에서 정부와 연계된 인공지능 관련 논문의 수 400% 증가

2007-2017년 정부와 연계된 인공지능 관련 논문의 수가 400% 증가한 반면 기업과 연계된 인공지능 관련 논문은 동 기간 동안 73% 증가하는데 그쳤다.

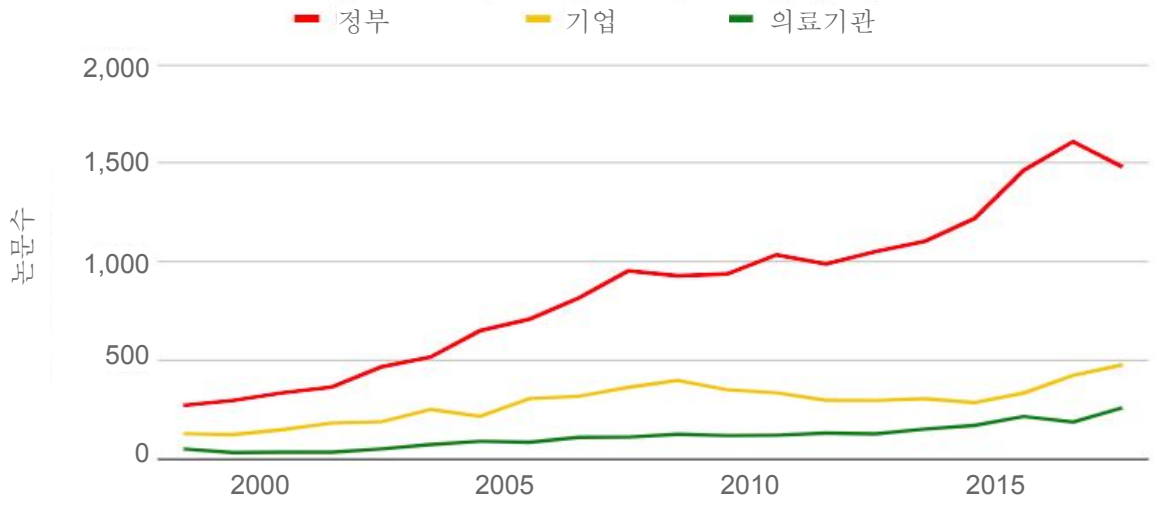
출간된 논문: 연계부문별 (계속)

출간된 논문: 연계부문별 – 미국(1998–2017)
출처: Elsevier



주석: 예를 들어, 2017년 Scopus에 등재된 인공지능 관련 논문 중 미국에서 학계와 연계된 논문의 수(그래프 없음)는 약9,000편으로 미국에서 출간되는 인공지능 관련 논문의 약81%를 차지했다.

출간된 논문: 연계부문별 – 유럽(1998–2017)
출처: Elsevier



주석: 예를 들어, 2017년 Scopus에 등재된 인공지능 관련 논문 중 유럽에서 학계와 연계된 논문의 수 (그래프 없음)는 약16,000편으로 유럽에서 출간되는 인공지능 관련 논문의 약88%를 차지했다.

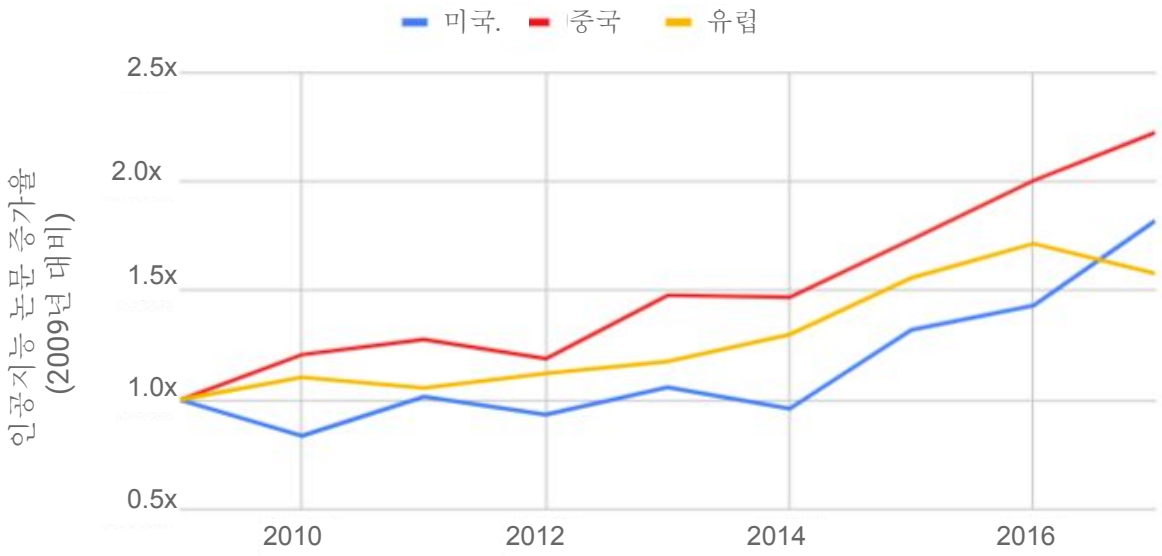
미국에서 기업과 연계된 논문의 비율 중국에서 기업과 연계된 논문 비율의 6.6배

2017년, 미국에서 기업과 연계된 인공지능 논문의 비율은 중국에서 기업과 연계된 인공지능 논문 비율보다 6.6배 높았다.

출간된 논문: 연계부문별 (계속)

기업과 연계된 인공지능 논문 증가율(2009–2017)
출처: Elsevier

정부와 연계된 인공지능 논문 증가율(2009–2017)
출처: Elsevier



중국에서 정부와 연계된 인공지능 관련 논문의 수 2009년 이래로 2배 이상 증가

반면, 미국에서는 기업과 연계된 인공지능 관련 논문이 더 높은 증가율을 보임. 2009년 대비 2017년 기업과 연계된 인공지능 관련 논문 1.7배 증가.

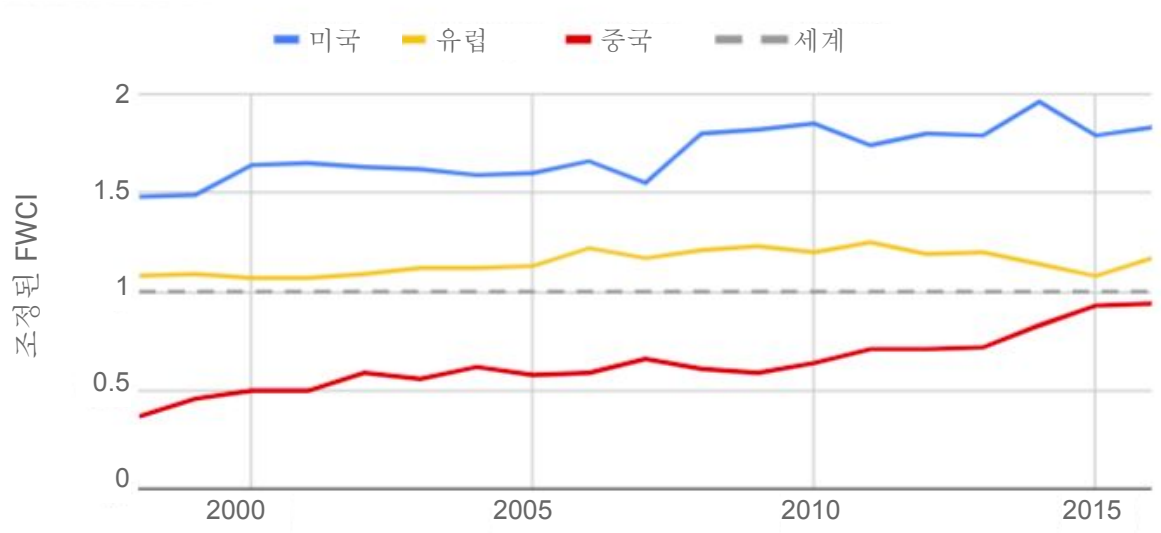
출간된 논문: 지역별 인용 영향력

이하 그래프는 인공지능 관련 논문 저자의 상대적 피인용 지수(field-weighted citation impact, FWCI)의 평균치를 지역별로 나타낸 것이다. 한 지역의 FWCI는 해당 지역에서 인공지능 관련 논문 저자의 피인용 횟수의 평균값을 모든 인공지능 관련 논문 저자의 피인용 횟수의 평균값으로 나눈 것이다. 이하 그래프에선 인용 영향력을 세계 평균과 비례하여 표시한 조정된(re-based) FWCI값을 보여준다. 조정된 FWCI 값 1은 논문 피인용 횟수가 세계 평균과 동일하다는 뜻이다. 조정된 FWCI 값 0.85는 논문 피인용 횟수가 세계 평균에 비해 15% 낮다는 것을 의미한다.

유럽이 가장 높은 연간 논문 출간율을 보이지만, 유럽의 조정된 FWCI 값은 상대적으로 수평 상태를 보였으며 세계 평균치와 같은 수준을 보였다. 반면, 중국의 조정된 FWCI 값이 상당히 증가하는 경향을 보였다. 중국의 인공지능 관련 논문 저자의 피인용 횟수는 2000년 대비 2016년 44% 증가하였다. 미국의 인공지능 관련 논문 저자의 피인용 횟수는 세계 평균에 비해 83% 높으며, 여전히 미국 저자의 피인용 횟수가 다른 지역 저자의 피인용 횟수보다 높다.

데이터 및 방법론은 부록 2 참고.

지역별 인공지능 관련 논문 저자의 상대적 피인용 지수(1998–2016)
출처: Elsevier



미국의 인공지능 관련 논문 저자의 피인용 횟수는 세계 평균에 비해 83% 높다.

출간된 논문: 지역별 저자 이동성

다음 페이지의 그래프는 국제적 이동이 출간율 및 인공지능 관련 저자의 인용 영향력에 미친 영향을 표시한다.

우리는 이동성(mobility)을 정주형(Sedentary), 일시적(Transitory), 이동유입(Migratory Inflow) 및 이동유출(Migratory Outflow) 4가지 범주로 분류했다. *정주형* 저자들은 활동 중인 연구자로서 저자의 거주 지역 밖에서 논문을 출간한 적이 없는 경우이다. *일시적* 저자들은 2년 이하의 기간 동안 타 지역에서 출간한 적이 있는 경우이다. *이동하는* 저자들은 그들의 지역 이외의 지역에 2년 이상 기여를 한 저자들이다. 저자가 *이동유입*으로 분류되느냐 *이동유출*의 경우로 분류되는지는 그래프를 바라보는 관점에 따라 다르다.

X 축은 상대적인 출간율을 나타낸다 - 상대적인 출간율은 각 범주에 속한 저자의 출간율의 평균값을 해당 지역의 전체 출간율의 평균값으로 나눈 값이다. Y 축은 *상대적 피인용 지수(field-weighted citation impact, FWCI)*를 가리킨다. FWCI 값은 각 이동성 범주에 속한 저자의 피인용 횟수의 평균값을 전체 지역의 피인용 횟수 평균값으로 나눈 값이다.

본 분석에서 오직 인공지능 관련 논문 저자만 다루었다. 최소한 30%의 논문에서 인공지능을 다룬 저자를 "인공지능 논문 저자"라고 하였다. 인공지능 저자의 거주 지역(home region)은 해당 저자가 첫 논문을 출간한 지역이다.

미국, 중국 및 유럽에서 정주형 저자들의 출간율이 가장 낮았다. 더불어 세 지역 모두에서 이동하는(이동유입과 이동유출의 조합) 저자들의 FWCI 값이 가장 높았다. 즉, 이동형 저자들이 피인용 횟수 및 출간율에서 모두 높은 경향을 보였다.

세 지역 중에 중국이 가장 높은 정주형 인공지능 논문 저자의 비율을 보였다(76%), 두번째로 높은 지역이 유럽(52%)였으며 다음이 미국(38%)이었다. 더 많은 비율의 중국 저자들이 정주형 이었지만 중국의 비정주형 저자들의 출간율이 유럽 및 미국의 비정주형 저자들의 출간율보다 높은 경향을 보였다. 다시 말해, 지리적으로 이동하는(mobile) 중국 저자들의 비율이 상대적으로 적으나 다른 지역의 이동하는 저자들 보다 평균적으로 더 생산적인 경향을 보였다.

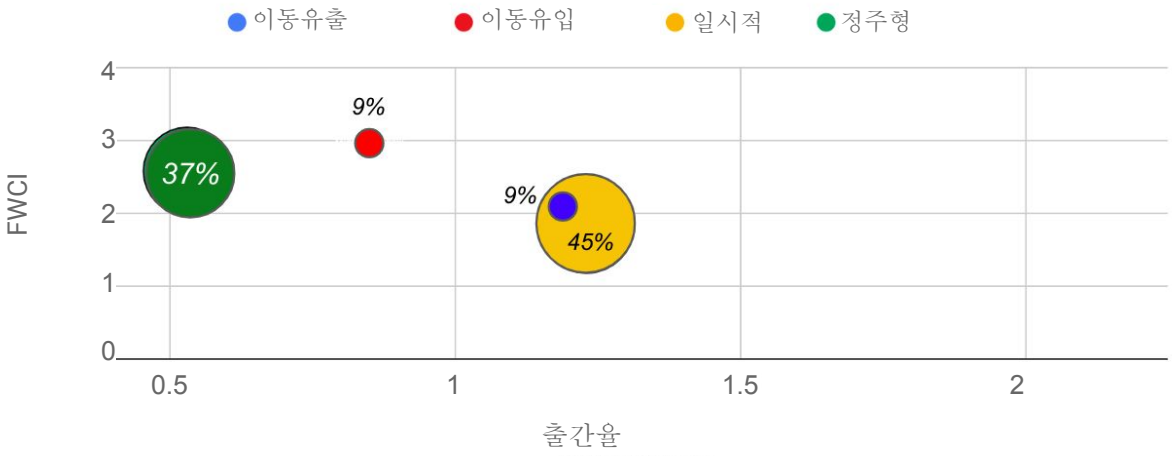
데이터 및 방법론은 부록 2 참고.

이동하는 저자들이 피인용 횟수 및 출간율에서 모두 높은 경향을 보였다.

출간된 논문: 지역별 저자 이동성

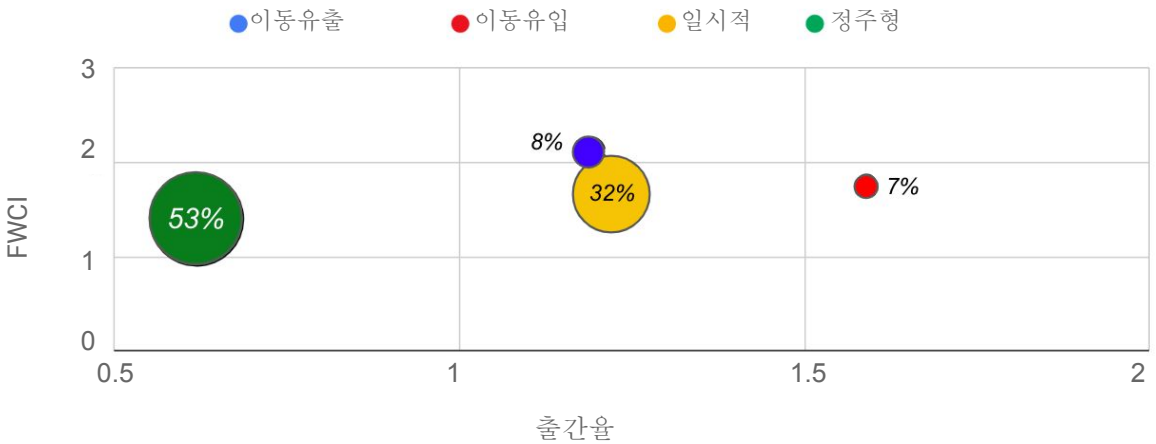
이전 페이지 내용 보기.

이동성 범주별 출간율 및 인용 영향력 — 미국(1998–2017)
출처: Elsevier



이동성 범주별 출간율 및 인용 영향력 — 중국(1998–2017)
출처: Elsevier

이동성 범주별 출간율 및 인용 영향력 — 유럽(1998–2017)
출처: Elsevier



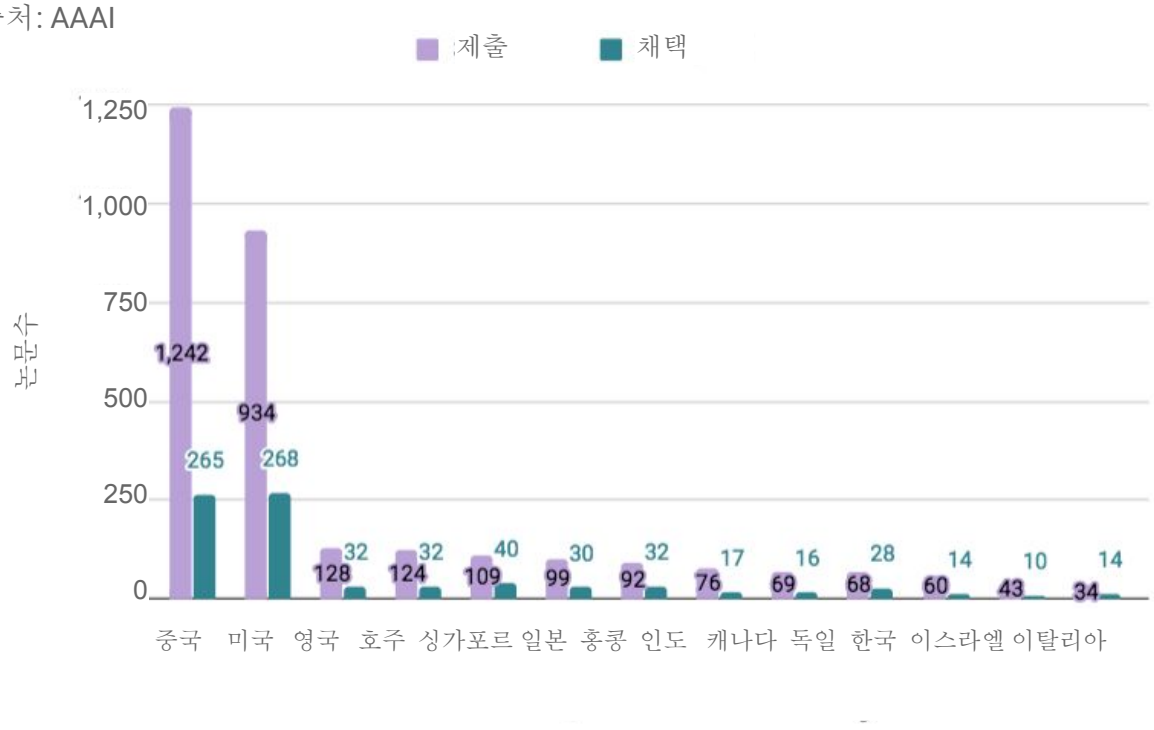
출간된 논문: 국가별 AAAI 논문

이하 그래프는 2018년 전미인공지능학회(AAAI)에 제출되고 채택된 논문의 수를 국가별로 나타낸다. 2018 AAAI 컨퍼런스는 2018년 2월 루이지애나 주 뉴올리언즈에서 열렸다.

2018년 AAAI 컨퍼런스에 제출된 논문의 70%는 미국 또는 중국과 연계된 논문이었다. 중국과 연계된 논문의 수가 가장 많이 제출되었지만, 채택된 논문의 수에 있어서는 미국과 중국이 대동소이했다. 미국과 연계된 논문이 268편 채택되었고, 중국과 연계된 논문이 265편 채택되었다. 결과적으로 미국과 연계된 논문의 채택율(acceptance rate)은 29%인 반면 중국과 연계된 논문의 채택율은 21%였다. 독일 및 이탈리아와 연계된 논문이 상대적으로 적은 수의 논문이 제출되었지만 가장 높은 채택율(41%)를 보였다.

데이터 및 방법론은 부록 4 참고. 컨퍼런스별 참가자 수 변화 26 페이지 참고.

제출 및 채택된 눈문 수 — 2018년 AAAI 컨퍼런스
출처: AAAI



미국 또는 중국과 연계된 논문이 AAAI 컨퍼런스에 제출된 논문의 70% 차지

*2018년 AAAI 컨퍼런스에서 제출된 논문의 **70%**, 채택된 논문의 **67%**가 미국 또는 중국과 연계된 논문이었다.*

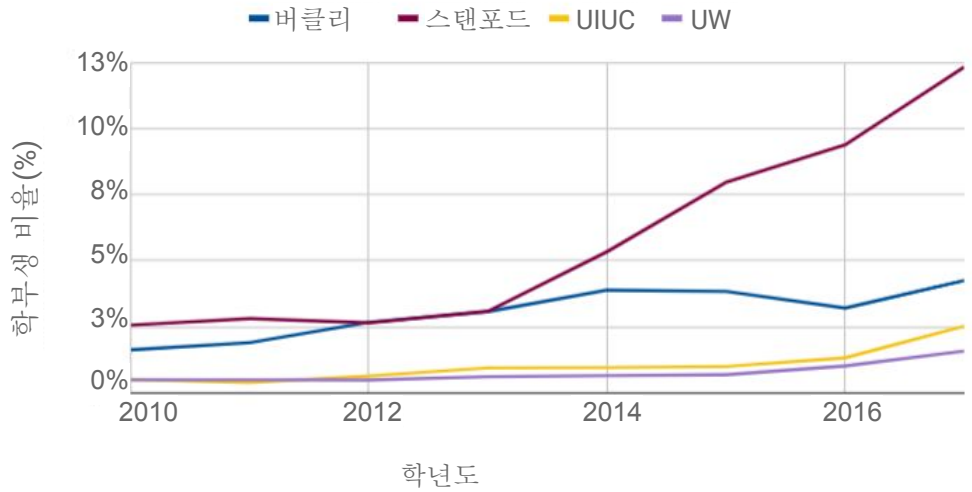
수강신청: 미국 인공지능 관련 수업

이하 그래프는 인공지능 및 머신러닝 개론 수업에 등록한 학부 학생의 비율을 나타낸다. 해당 비율은 학교별 학생수를 반영한 지표이다. 학교 선정 기준, 실제 수강신청자 수 및 학교의 이름은 부록에 제시되어 있다.

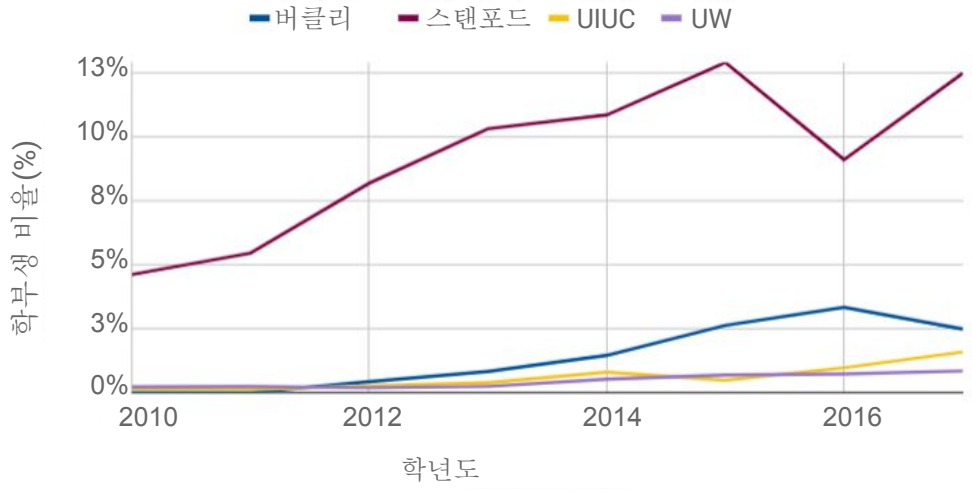
전체 학부생 중 인공지능 개론 수업을 수강하는 학부생 비율이 머신러닝 개론 수업을 수강하는 학부생 비율보다 높은 경향이 있지만(평균적으로 인공지능 개론 수업을 수강하는 학부생은 전체 학부생의 5.2%였으며 머신러닝 개론 수업을 수강하는 학부생은 전체 학부생의 4.4%였다), 머신러닝 개론 수업을 수강하는 학부생 수의 증가율이 더 높다. 다음 페이지에 증가율 비교를 제시하였다. 이러한 경향은 인공지능의 하위 분야에서 머신러닝의 커지는 중요성을 보여준다.

데이터 및 방법론은 부록 5 참고.

인공지능 개론 수업을 수강하는 학부생 비율(2010–2017)
출처: 대학교 제공 데이터



머신러닝 개론 수업을 수강하는 학부생 비율(2010–2017)
출처: 대학교 제공 데이터



성별 다양성에 대한 주석:
스탠포드 '17 AI 개론(Intro to AI): 74.45% 남성(총 학생 869 명); 버클리 '17 AI 개론(Intro to AI): 73.37% 남성(총 학생 973명)
스탠포드 '17 ML 개론(Intro to ML): 75.91% 남성(총 학생 876 명); 버클리 '17 ML개론(Intro to ML): 78.67% 남성(총 학생 720명)

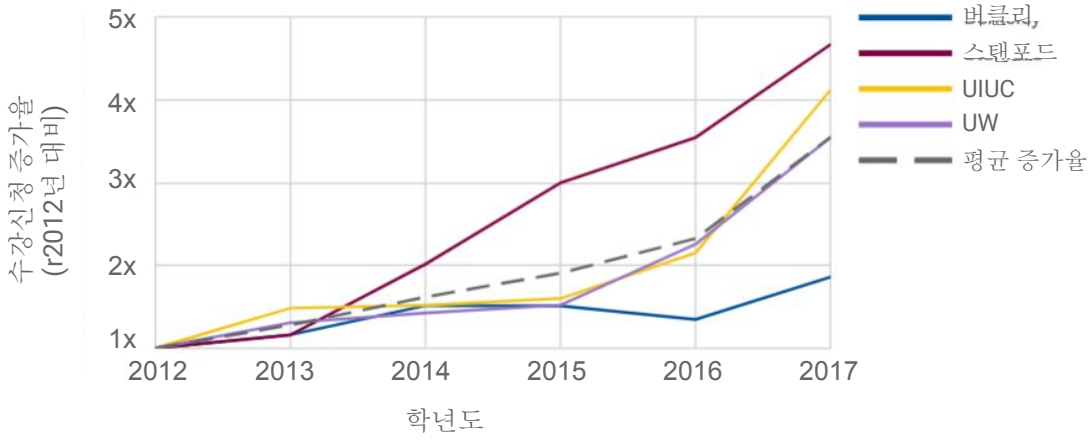
수강신청: 미국 인공지능 관련 수업 (계속)

이하 그래프는 미국의 상위 컴퓨터 과학 학교의 인공지능 및 머신렁닝 수업 수강신청 증가율을 보여준다. 학교 선정 기준, 실제 수강신청자 수 및 학교의 이름은 부록에 제시되어 있다.

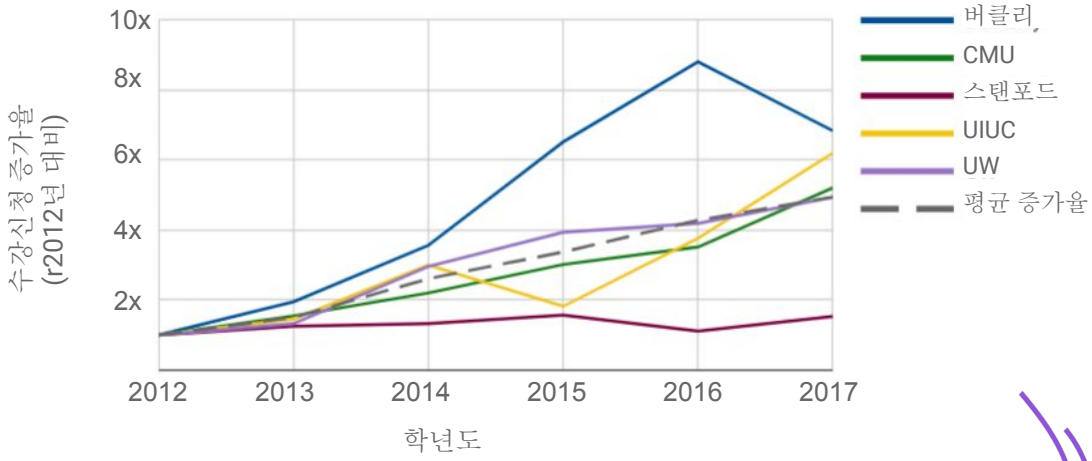
2017년 인공지능 개론 수업의 수강신청자 수가 2012년 대비 3.4배 증가한 반면 2017년 머신러닝 개론 수업의 경우 수강신청자 수가 2012년 대비 5배 증가하였다. UC 버클리 대학교의 2017년 머신러닝 개론 수업의 수강신청자 수는 2012년 대비 6.8배 증가하였다. UC 버클리의 머신러닝 개론 수업 수강신청자 수가 우리가 조사한 머신러닝 개론 수업 중 가장 높았다.

데이터 및 방법론은 부록 5 참고.

인공지능 개론 수업 수강신청 증가율(2012–2017)
출처: 대학교 제공 데이터



머신러닝 개론 수업 수강신청 증가율(2012–2017)
출처: 대학교 제공 데이터



2017년 머신러닝 개론 수업 수강신청자 수 2012년 대비 5배 증가

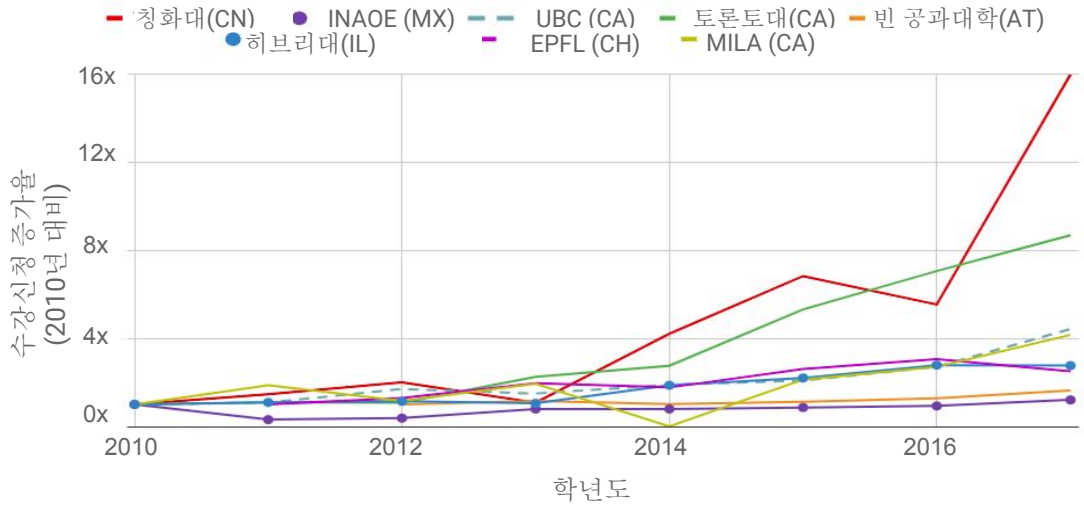
수강신청: 미국 외 지역 소재 대학교 수업

본 페이지 및 다음 페이지의 그래프는 미국 외 지역 소재 상위 컴퓨터 과학 대학교의 인공지능 및 머신러닝 수업 수강 정보를 보여준다. 첫번째 그래프는 2010-2017학년도 데이터를 제공한 미국 외 지역 소재 대학교의 상대적 증가율을 보여준다. 두번째 그래프는 2016-2018학년도 데이터를 제공한 미국 외 지역 소재 대학교의 상대적 증가율을 보여준다. 학교 선정 기준, 실제 수강신청자 수 및 학교의 이름은 부록에 제시되어 있다.

칭화 대학교(Tsinghua University)의 인공지능 수업과 머신러닝 수업 수강신청 수의 합이 2010년 대비 2017년 16배 증가하였다. 분석한 대학교 중 인공지능 수업 수강신청 증가율은 학교에 따라 달랐으며 특별히 지리적 위치의 영향을 받은 것은 아니었다. AI 인덱스는 추후 발간되는 보고서에서 본 가설을 다듬을 계획이다.

데이터 및 방법론은 부록 6 참고.

인공지능 + 머신러닝 수업 수강신청 증가율 — 미국 외 지역(2010—2017)
출처: 대학교 제공 데이터

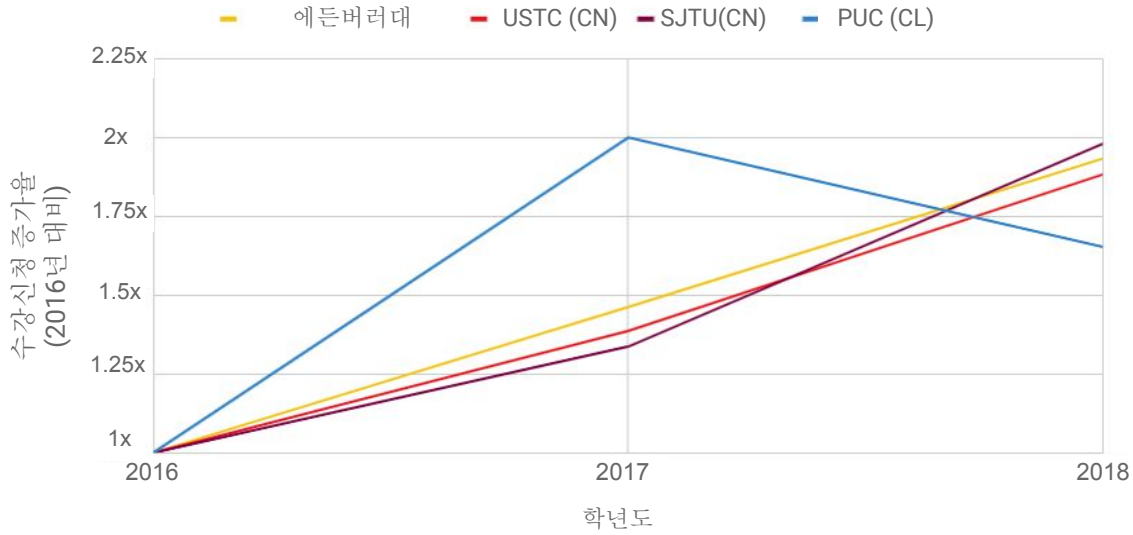


주석: 본 그래프는 2010-2017학년도 데이터가 가용한 대학교의 수치를 보여준다. 증가율은 2010년 수강신청 수를 기준으로 하며 2011년을 기준으로 하는 로잔 연방 공과대학(EPFL), 2012년을 기준으로 하는 오스트리아 빈 공과대학(TU Wien)의 경우는 예외이다.

수강신청: 미국 외 지역 소재 대학교 수업 (계속)

이전 페이지 내용 보기.

인공지능 수업 수강신청 증가율 – 미국 외 지역(2016–2018)
출처: 대학교 제공 데이터



주석: 본 그래프는 과거 자료가 제한적인 대학교의 수치를 보여준다. 증가율은 2016년 수강신청을 기준으로 한다.

교수 다양성

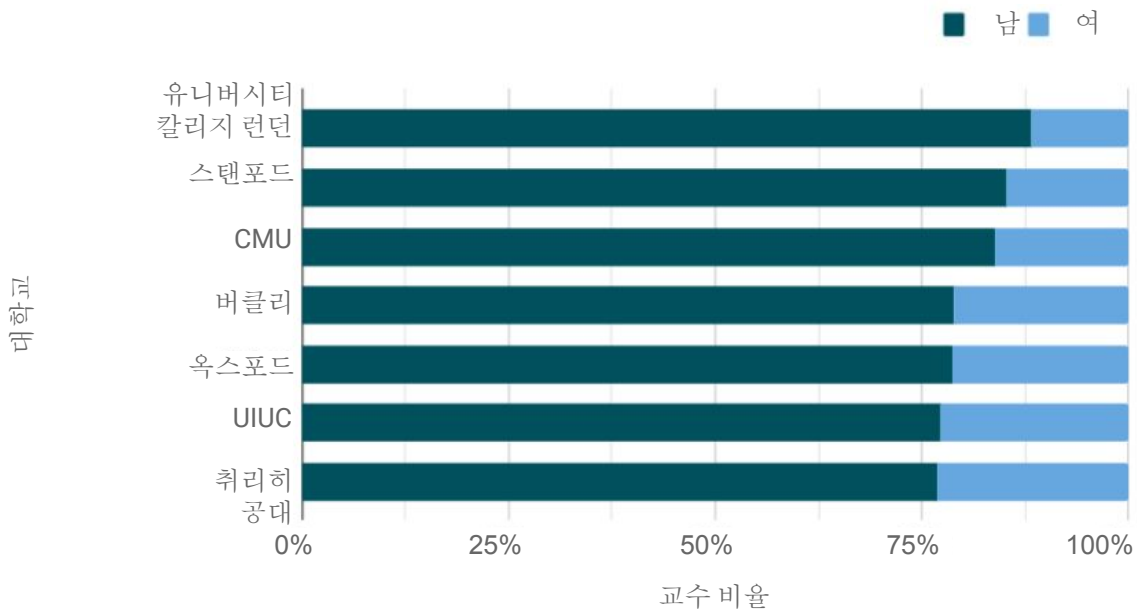
이하 그래프는 컴퓨터 과학 분야 상위 대학교의 인공지능 분야 교수의 성별 분류를 보여준다. 데이터는 2018년 9월 21일 기준 교수 명부를 사용하여 수집하였다. 인공지능 분야 교수 명부에 쉽게 접근할 수 있는 대학교를 분석하였다.

산업 및 학계의 다양성 관련 통계 데이터에 대한 접근성이 낮은 것이 다양성 증진에 큰 장애가 된다. 우리는 교육 기관이 인공지능 분야의 다양성 증진을 위하여 다양성 관련 통계를 더 투명하게 할 것을 촉구한다.

분석한 학교들에서 80%의 인공지능 분야 교수가 남성이었다. 지리적 위치와 무관하게 학교별 인공지능 분야 남성 교수 비율은 큰 차이를 보이지 않았다. 그러나 제한된 숫자의 학교만 분석하였기 때문에 본 분석 결과는 더 큰 그림의 일부에 불과하다.

데이터 및 방법론은 부록 7 참고.

인공지능 분야 교수 성비 — 선택된 학교(2018년 9월)
출처: 대학교 교수 명부



인공지능 분야 교수의 80%가 남성

평균적으로 UC 버클리, 스탠포드, 일리노이 주립대학교, 유니버시티 칼리지 런던, 옥스포드 및 취리히 공대 인공지능 분야 교수의 80%가 남성이다.

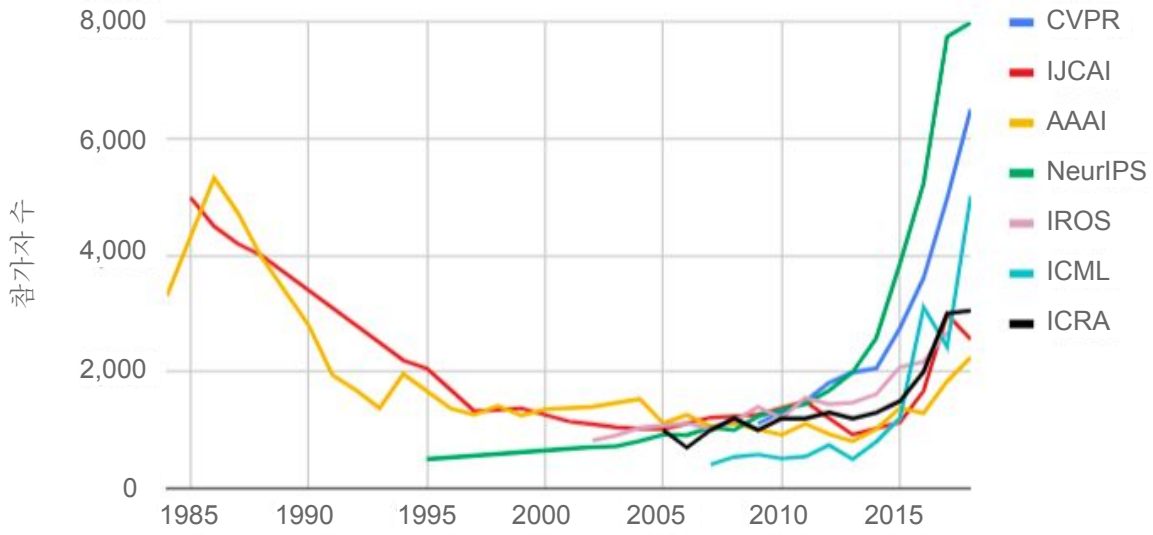
참가: 대규모 인공지능 컨퍼런스

이하 그래프는 대규모 인공지능 컨퍼런스 참가자 수와 2012년 참가자 대비 참가자 수의 증가율을 보여준다. 대규모 인공지능 컨퍼런스는 2017년 2천명 이상이 참가한 컨퍼런스를 의미한다.

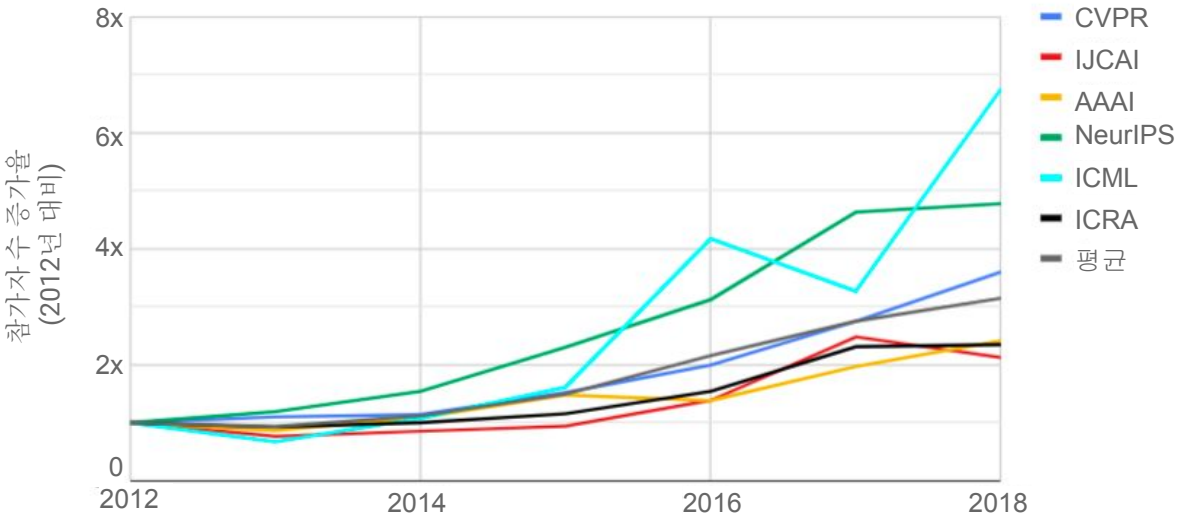
NeurIPS(구 NIPS), CVPR 및 ICML이 가장 높은 참가자수를 보이는 컨퍼런스이며 2012년 이후 가장 높은 참가자 증가율을 보였다. NeurIPS 및 ICML이 2012년 참가자 수 대비 각 4.8배, 6.8배로 가장 높은 증가율을 보였다. 이는 인공지능의 하위 분야로서 머신러닝에 대한 관심을 대변한다. 반면 상징 추론 분야에 초점을 맞춘 컨퍼런스는 상대적으로 낮은 증가율을 보였다.

데이터 및 방법론은 부록 8 참고.

대규모 인공지능 컨퍼런스 참가자 수(1984–2018)
출처: 컨퍼런스 제공 데이터



대규모 인공지능 컨퍼런스 참가자 수 증가율(2012–2018)
출처: 컨퍼런스 제공 데이터



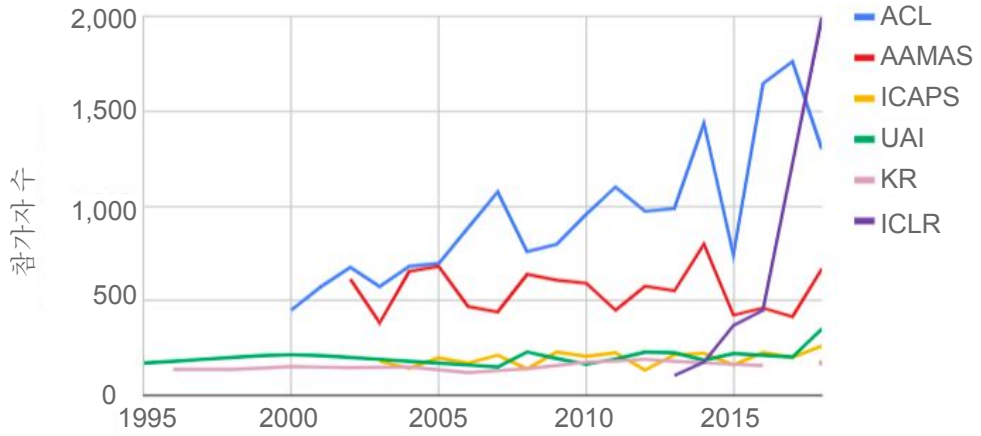
참가: 소규모 인공지능 컨퍼런스

이하 그래프는 소규모 인공지능 컨퍼런스의 참가자 수와 2012년 참가자 대비 참가자 수 증가율을 보여준다. 소규모 인공지능 컨퍼런스는 2017년 2천명 미만이 참가한 컨퍼런스를 의미한다.

ICLR의 2018년 참가자 수는 2012년 대비 20배 증가하였다. 해당 증가세는 인공지능 분야에서 딥러닝 및 강화학습으로 초점이 옮겨간 결과일 수 있다.

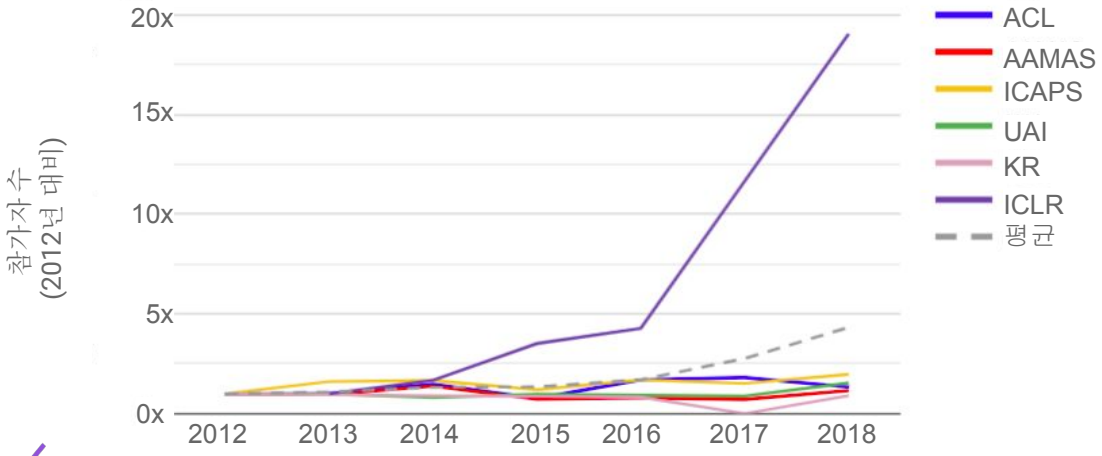
데이터 및 방법론은 부록 8 참고.

소규모 컨퍼런스 참가자 수(1995–2018)
출처: 컨퍼런스 제공 데이터



주석: 2018년 KR에서 첫 컨퍼런스를 개최하였다. 일관성을 위해 컨퍼런스 참가자에 KR 참가자 수는 포함하지 않는다.

소규모 컨퍼런스 참가자 수 증가율(2012–2018)
출처: 컨퍼런스 제공 데이터



ICLR의 2018년 참가자 수는 2012년 대비 20배 증가

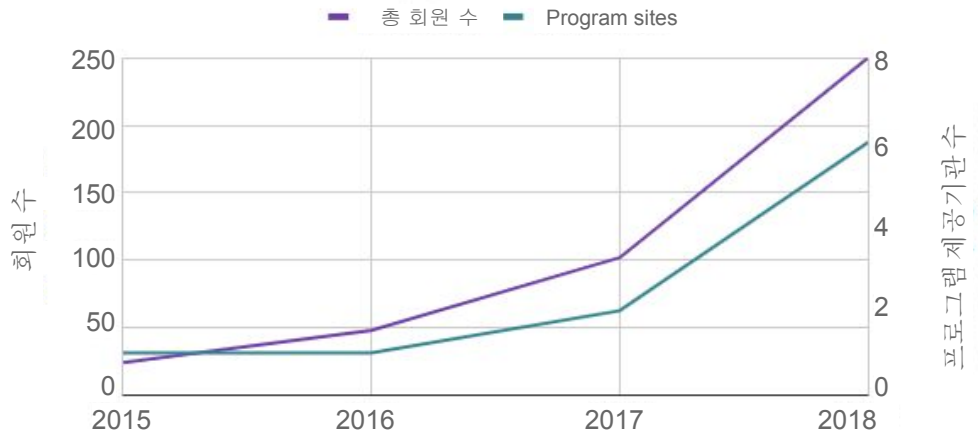
참가: 다양성 관련 기관

이하 그래프는 머신러닝 분야의 여성 모임(Women in Machine Learning, WiML) 이 개최하는 연례 워크숍 참가자 수와 AI4ALL의 회원 수(alumni)를 보여준다. WiML은 머신러닝 분야의 여성을 지원하는 단체이다. AI4ALL은 인공지능 분야의 다양성 및 포용성을 증진하기 위해 고안된 인공지능 교육 이니셔티브이다.

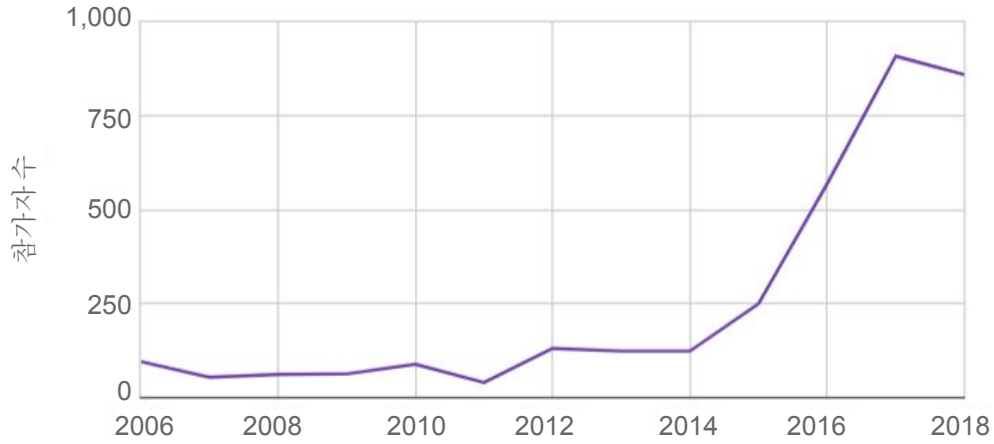
WiML 워크숍 및 AI4ALL의 참가자 수는 지난 몇 년 간 증가하였다. WiML 워크숍의 경우 2014년 대비 600% 증가하였으며 AI4ALL의 경우 회원수가 2015년 대비 900% 증가하였다. 이러한 경향은 인공지능 분야에서 여성 및 소수자 집단을 포함시키기 위한 지속적인 노력을 보여준다.

데이터 및 방법론은 부록 8 참고.

AI4ALL 회원 수 및 프로그램 제공기관 수(2015–2018)
출처: AI4ALL



WiML 워크숍 참가자 수(2006–2018)
출처: WiML



주석: WiML 워크샵 등록이 2017년에 살짝 부풀려졌다. 그 이유는 기존에 하루 일정으로 진행된 것이 이틀 일정으로 바뀌었기 때문이다.

"...본 데이터는 WiML 커뮤니티에서 최근 목도한 증가율을 반영한다. 우리는 이러한 경향이 인공지능 분야에서 의도적으로 다양성을 증진하는데 더 초점이 맞추어 지는 것과 연관이 있다고 본다. 그러나 보통 WiML과 같은 장소에서 열리는 NeurIPS 컨퍼런스 참가자 수가 수천명인 반면 WiML 행사의 참가자 수가 여전히 수백명에 그치기 때문에 아직 갈 길이 멀다."
WiML 이사회

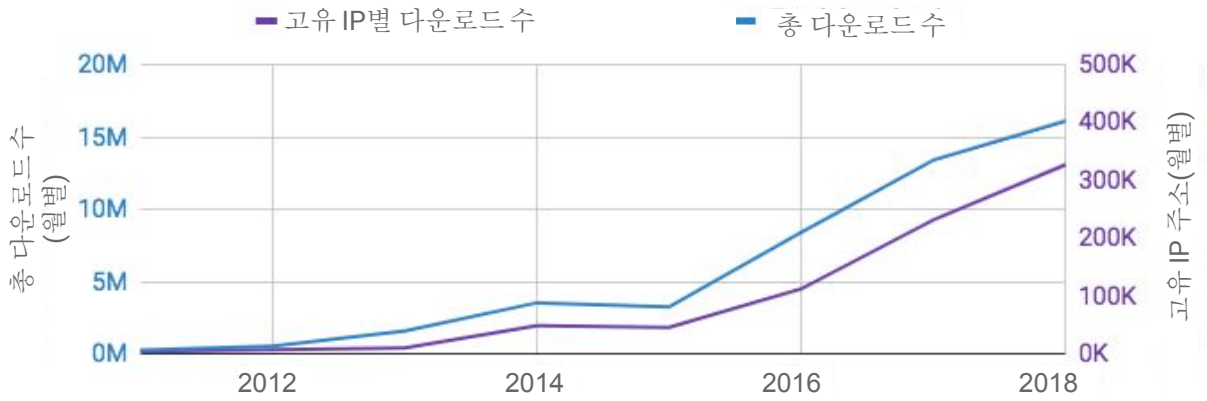
로봇 소프트웨어 다운로드

이하 그래프는 ROS.org에서 로봇 운영 체제(ROS)의 바이너리 패키지(binary packages) 다운로드 횟수를 나타낸다. ROS는 로봇을 위한 오픈소스 소프트웨어 스택으로 널리 사용된다. ROS는 많은 상업적 제조사 및 학계 연구원들이 사용한다. 왼쪽 축은 월 평균 총 다운로드 횟수를 나타내며 오른쪽 축은 고유(unique) IP 주소별 월 평균 다운로드 횟수를 나타낸다.

2014년 이래로 총 다운로드 횟수 및 고유 IP별 다운로드 횟수가 각 353%, 567% 증가하였다. 이는 로봇공학 및 로봇 시스템 사용에 대한 증대된 관심을 나타낸다. 고유 IP별 다운로드 횟수의 성장률이 총 다운로드 횟수의 성장률보다 높다. 이를 통해 ROS가 더 자주 사용되고 있을 뿐만 아니라 더 많은 ROS 사용자가 있다는 것을 추론하 수 있다.

데이터 및 방법론은 부록 9 참고.

ROS 패키지 다운로드(2011—2018)
출처: ROS 지표 보고서



2014년부터 고유 IP별 ROS 패키지 다운로드 횟수 567% 증가

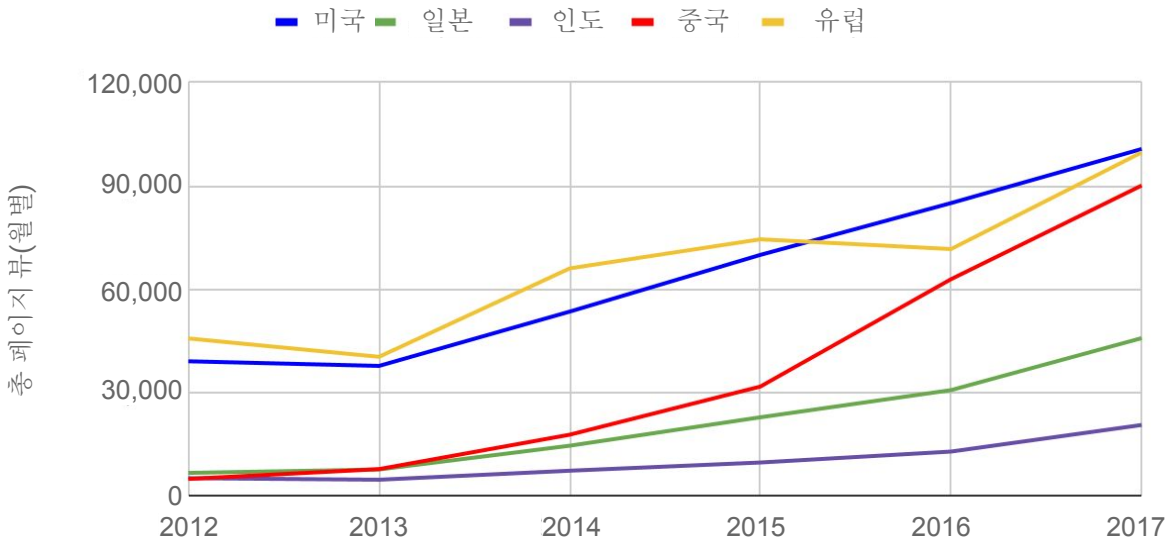
로봇 소프트웨어 다운로드

이하 그래프는 2012년 이래로 ROS.org 페이지 뷰가 가장 높은 상위 5개 지역을 표시한다.

미국 및 유럽이 가장 높은 ROS 페이지 뷰를 기록했다. 중국은 어떠한 지역보다 높은 증가율을 보인 후 근소한 차이로 3위를 기록했다. 2017년 중국의 ROS.org 페이지 뷰는 2012년 대비 18배 증가하였다. ROS.org는 중국에서의 페이지 뷰 증가는 자연스러운 증가이며, 더 많은 홍보나 중국의 다른 자원 때문은 아니라고 밝혔다.

데이터 및 방법론은 부록 9 참고.

ROS.org 페이지 뷰(2012—2017)
출처: ROS 지표 보고서



중국에서의 2017년 ROS.org 페이지 뷰, 2012년 대비 18배 증가

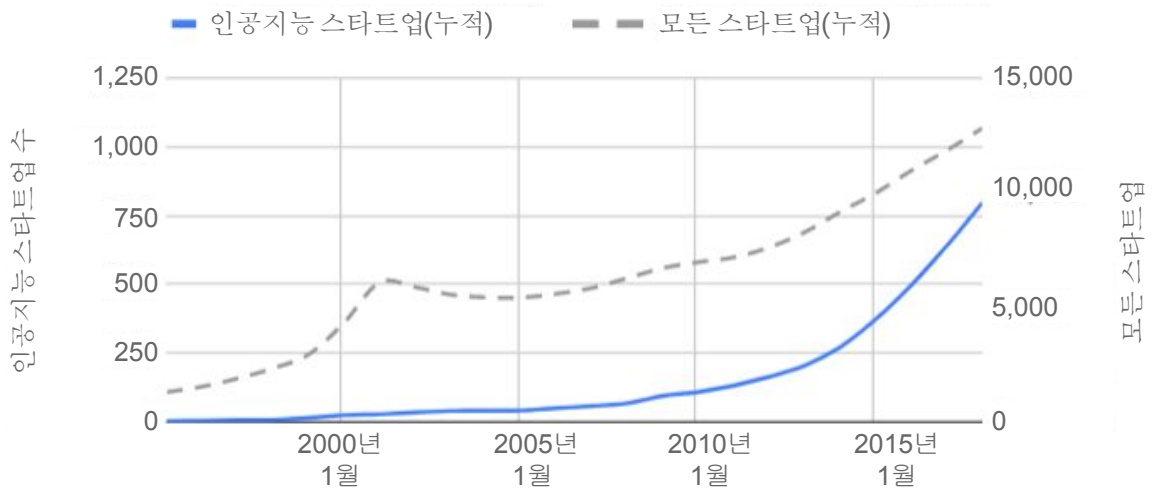
스타트업/투자: 인공지능 스타트업

이하 그래프는 연도별 미국의 벤처 투자를 받고 운영 중인 민간 스타트업의 수를 나타낸다. 파란색 그래프(왼쪽 축)은 인공지능 스타트업만을 보여주는 반면 회색 그래프(오른쪽 축)은 벤처 투자를 받은 인공지능 스타트업을 포함한 모든 스타트업을 가리킨다. 그래프는 매해 1월 해당 스타트업 수를 기록하였다. 데이터셋에서 제외된 스타트업은 제외하였으며(추가 정보 부록 참고) 스타트업 숫자는 매년 누적되는 방식으로 기록하였다.

2015년 1월부터 2018년 1월까지 운영중인 인공지능 스타트업 수는 113% 증가한 반면 모든 운영중인 스타트업 수는 28% 증가하였다. 대부분의 경우에 모든 운영중인 스타트업 수의 증가율을 상대적으로 완만한 반면 인공지능 스타트업의 수는 기하급수적으로 증가하였다.

데이터 및 방법론은 부록 10 참고.

인공지능 스타트업(미국, 1995년 1월 –2018년 1월)
출처: Sand Hill Econometrics



주석: 위 그래프의 인공지능 스타트업 대부분은 인공지능 시스템을 개발한다. 일부는 인공지능을 사업의 중요한 요소로 사용하지만 자체적으로 인공지능 시스템을 개발하지는 않는다. 추가 내용 부록 참고.

2015년부터 2018년까지 운영중인 인공지능 스타트업 수 113% 증가

반면 운영중인 스타트업 수는 전체적으로 28% 증가

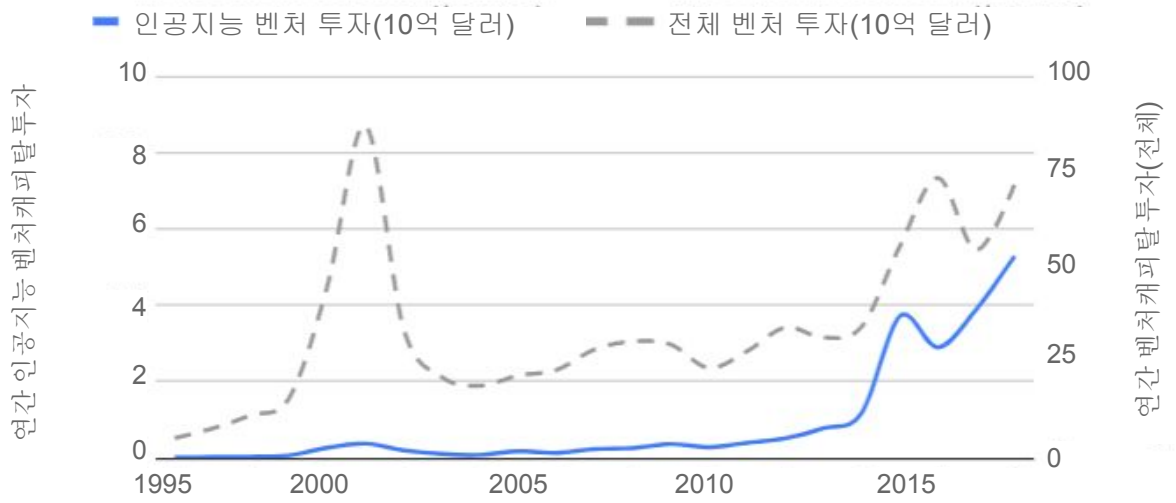
스타트업/투자: 벤처캐피탈 투자

이하 그래프는 모든 투자 단계에 걸쳐 미국 스타트업을 대상으로 한 벤처캐피탈의 연간 투자금액을 나타낸다. 파란색 선(왼쪽 축)은 인공지능 스타트업을 대상으로 한 연간 투자금액을 나타내며 회식 선(오른쪽 축)은 인공지능 스타트업을 포함한 벤처캐피탈에 투자 받은 모든 스타트업을 나타낸다. 데이터는 연간 투자금을 나타내며 이전 페이지의 그래프와는 다르게 투자금이 매년 누적되는 것은 아니다.

2013년부터 2017년까지 미국의 인공지능 스타트업을 대상으로 한 벤처캐피탈 투자는 350% 증가한 반면 총 벤처 투자금은 100% 증가하였다. 1997년 – 2000년까지 있었던 모든 벤처캐피탈 투자 붐은 닷컴 버블로 설명할 수 있다. 2014, 2015년에 있었던 소규모 호황은 상대적으로 경제가 크게 성장한 것을 반영한다.

방법론은 부록 10 참고.

인공지능 스타트업을 대상으로 한 연간 벤처캐피탈 투자액(미국, 1995 – 2017)
출처: Sand Hill Econometrics



2013년부터 2017년까지 미국의 인공지능 스타트업을 대상으로 한 벤처 투자 350% 증가

반면 운영 중인 모든 스타트업을 대상으로 한 총 벤처 투자금은 100% 증가

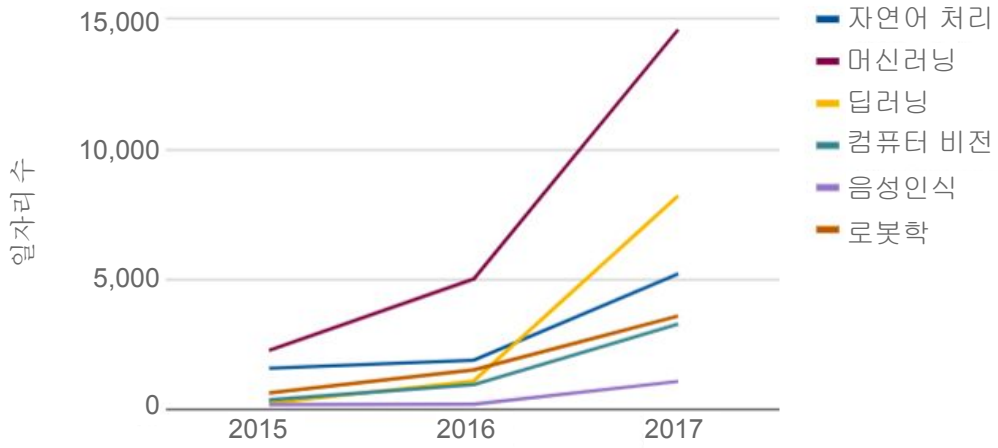
일자리: 인공지능 관련 일자리 증가세

이하 첫번째 그래프는 인공지능 관련 능력을 요구하는 일자리 수의 연간 수치를 보여주며 두번째 그래프는 요구되는 인공지능 관련 기술별 일자리의 상대적 증가율을 나타낸다. 인공지능 관련 기술은 상호 배타적이지 않다.

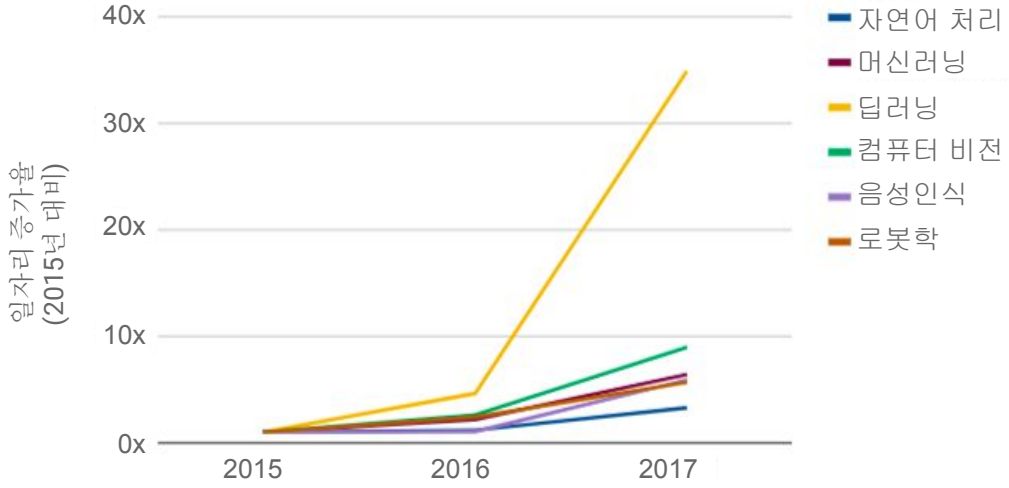
이 중 머신러닝이 가장 많이 요구되는 기술이었으며 딥러닝이 요구되는 일자리가 가장 빠르게 증가하는 경향을 보였다 - 2015년부터 2017년 딥러닝 기술을 요구하는 일자리의 수는 34배 증가하였다.

방법론은 부록 11 참고.

인공지능 기술이 요구되는 일자리 수(2015 – 2017)
출처: Monster.com



인공지능 기술이 요구되는 일자리 증가율(2015 – 2017)
출처: Monster.com



주석: 인공지능 관련 일자리가 전반적으로 증가하는 추세를 보이지만 여전히 컴퓨터 공학자를 위한 일자리 중 매우 적은 수에 불과하다.

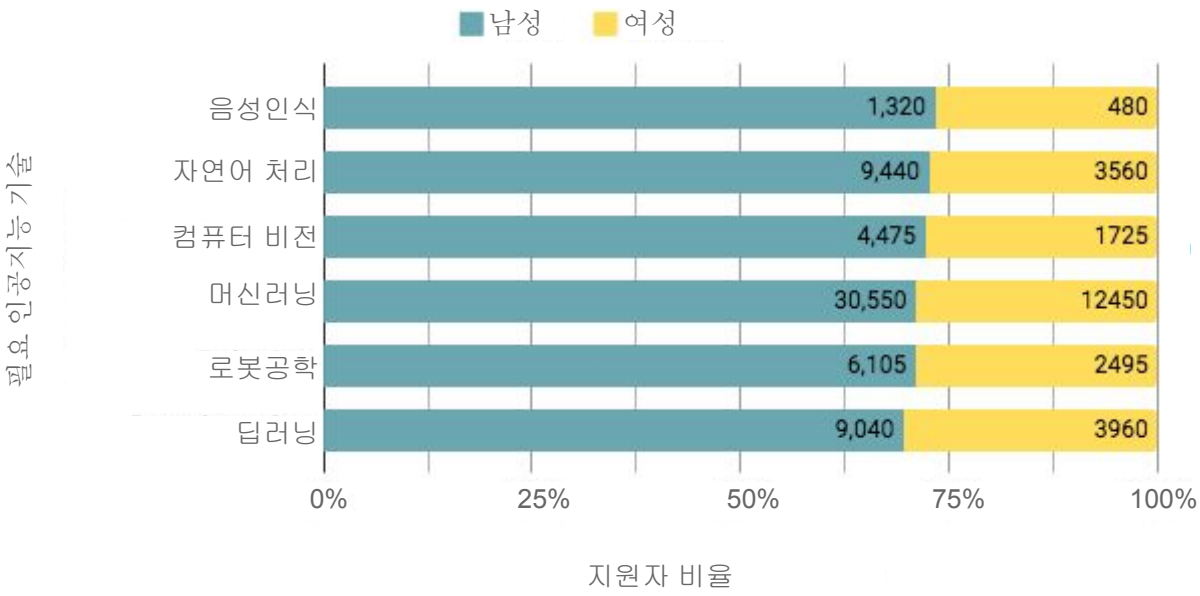
일자리: 지원자의 성별 다양성

이하 그래프는 2017년 인공지능 관련 일자리에 지원한 남녀 지원자 수를 나타낸다. 데이터는 요구되는 기술별로 분류하였으며 상호 배타적이지 않다. 지원자 수가 고용 또는 산업 전체에서 각 성별 비율을 대변하는 것은 아니다.

평균적으로, 남성이 미국에서 인공지능 관련 일자리 지원자의 71%를 차지한다. *머신러닝*이 가장 많은 지원자 수를 보이기 때문에 해당 평균치는 *머신러닝* 기술이 요구되는 일자리 지원자 수의 영향을 많이 받았다. *머신러닝*과 더불어 *딥러닝* 및 *로봇공학*이 다른 기술에 비해 상대적으로 높은 성별 다양성을 보였다.

데이터 및 방법론은 부록 11 참고.

성별 일자리 지원자(2017)
출처: Gartner TalentNeuron



평균적으로 남성이 미국에서 인공지능 관련 일자리 지원자의 71% 차지

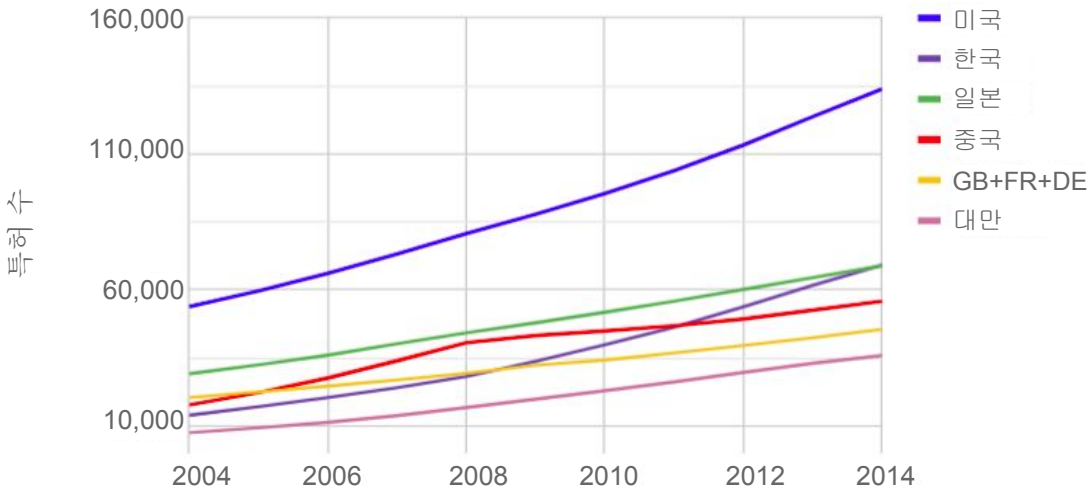
특허

이하 그래프는 (발명가가 속한) 국가별 인공지능 관련 특허 수 및 성장율을 나타낸다. 인공지능 관련 특허 수는 *인식과 의미 이해(Cognition and meaning understanding)* 및 *휴먼 인터페이스 기술(Human-interface technology)* 분야에 속하는 IPC 코드를 이용하여 수집하였다. 특허 개수를 추적하는 것은 쉽지 않다. 본 수집 방법과 관련한 메모 및 뉘앙스는 부록 참고.

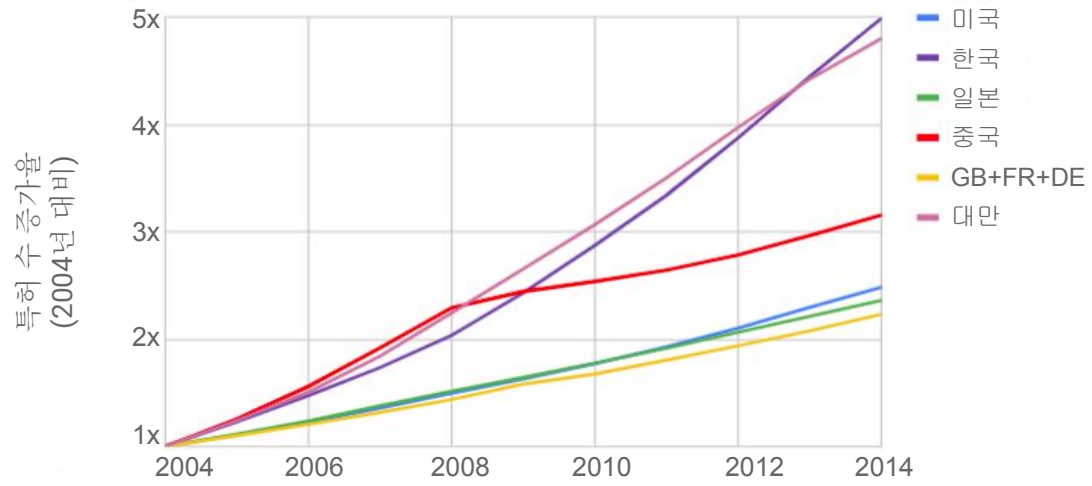
2014년에 인공지능 관련 특허의 30%가 미국에서 등록되었으며 대한민국과 일본이 그 뒤를 이었다. 대한민국 및 일본이 전체 인공지능 특허 중 약 16%를 보유하고 있다. 상위 특허 등록 지역 중 대한민국과 대만이 2004년 대비 2014년 인공지능 특허 수가 5배 증가하여 가장 높은 성장세를 보였다.

데이터 및 방법론은 부록 12 참고.

발명가 소재 지역별 인공지능 관련 특허 수(2004–2014)
출처: amplified



발명가 소재 지역별 인공지능 관련 특허 증가율(2004–2014)
출처: amplified



주석: GB + FR + DE는 영국, 프랑스, 독일의 수치의 합을 의미한다.

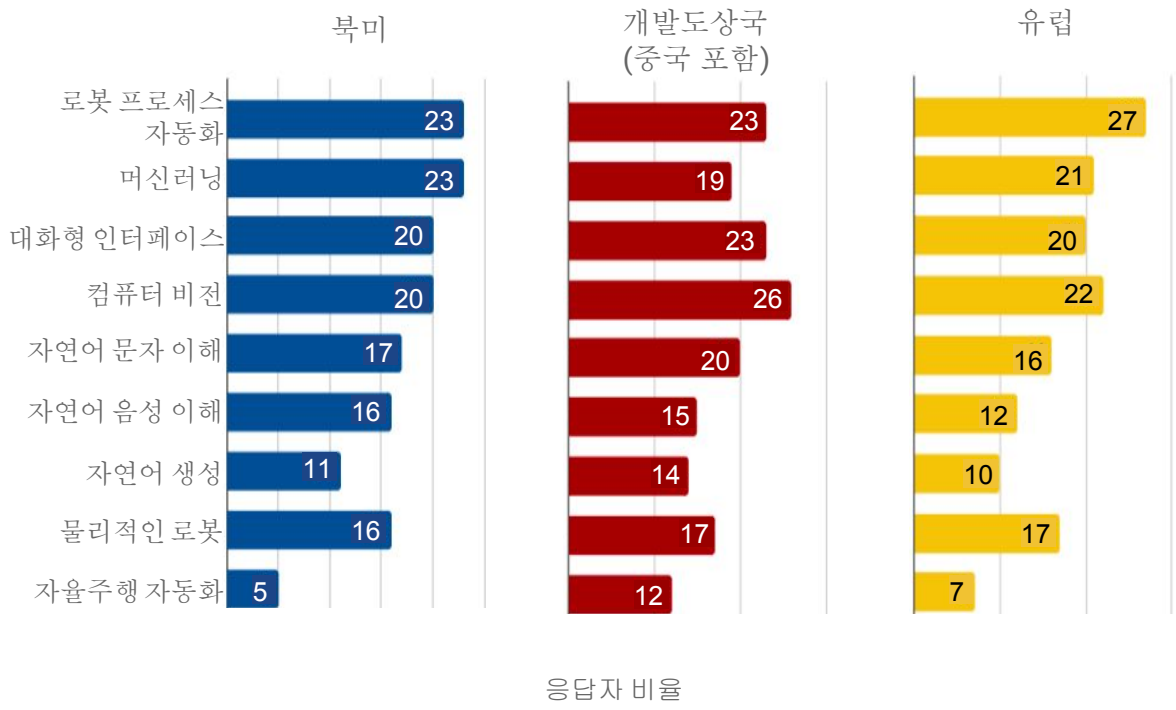
인공지능 도입: 지역별 역량

이하 그래프는 2,135명을 대상으로 한 맥킨지 앤 컴퍼니(McKinsey & Company)의 설문 조사 결과를 나타낸다. 각 응답자는 자신이 속한 기관을 대표하여 응답 하였다. 그래프는 자신이 속한 기관이 인공지능 기술을 적어도 한가지 이상의 분야 또는 사업부문에 도입하였다고 응답한 응답자의 비율을 나타낸다. 응답자는 다수의 인공지능 기술을 선택할 수 있었다. *아시아 태평양, 인도, 중동 및 북아프리카*와 *라틴아메리카* 지역은 다음 페이지 참고.

일부 지역이 타 지역에 비해 특정 기술을 더 도입한 경우는 있지만, 인공지능 기술이 지역별로 도입된 비율은 상대적으로 유사했다. 우리는 앞으로 기업들이 어떻게 인공지능 기술을 도입하는지 지속적으로 추적할 것이다.

데이터 및 방법론은 부록 13 참고.

한가지 이상의 사업부문에 도입된 기술(2018)
출처: 맥킨지 앤 컴퍼니



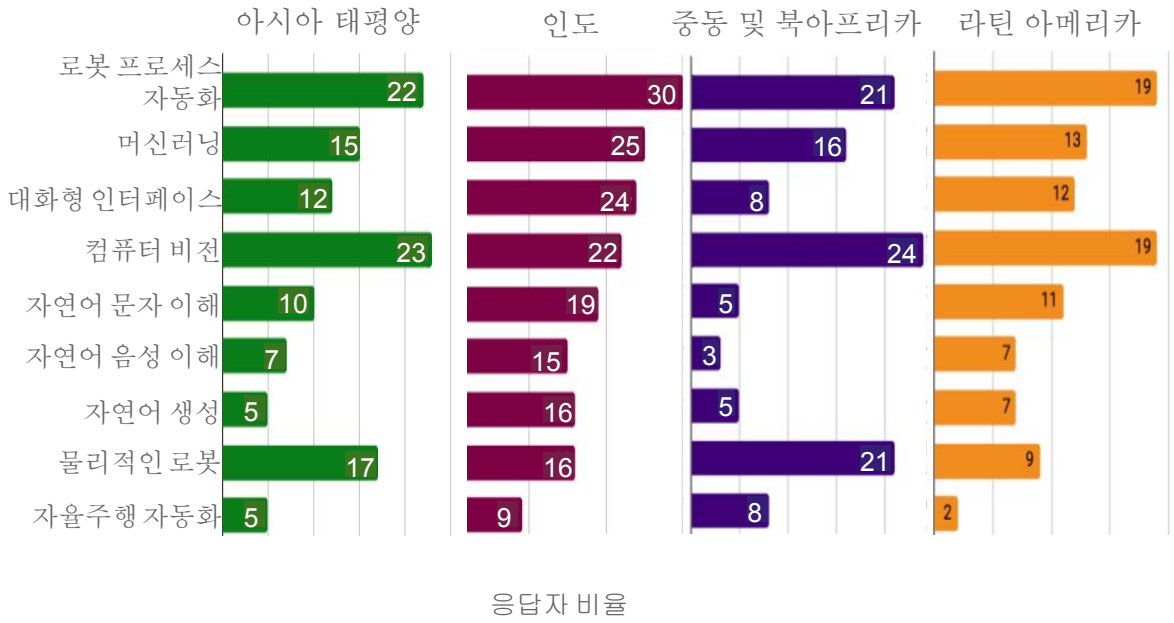
주석: 바의 길이는 각 지역별로 도입된 기술 비중의 상대적인 정도를 나타낸다; 북미: N = 479; 개발도상국(중국 포함): N = 189 (중국 N = 35); 유럽: N = 803

"...우리는 다양한 분야, 사업 부문 및 지역별로 인공지능 기술이 광범위하게 도입된 것을 발견하였다. 전체 기업의 약 절반이 인공지능 기술을 업무 프로세스에 도입하였다. 하지만, 아직 초기 단계이며 대부분의 기업에서 인공지능 기술을 활용하여 대규모로 가치를 창출하는데 필요한 상호 보완적인 기능을 아직 도입하지 않고 있다...."
- 마이클 추이, 맥킨지 앤 컴퍼니, 맥킨지 글로벌 연구소

인공지능 도입: 지역별 역량 (계속)

북미, 개발도상국(중국 포함) 및 *유럽*에 관한 내용 및 데이터 이전 페이지 참고.

한가지 이상의 사업부문에 도입된 기술(2018)
출처: 맥킨지 앤 컴퍼니



주석: 바의 길이는 각 지역별로 도입된 기술 비중의 상대적인 정도를 나타낸다; 아시아 태평양: N = 263; 인도: N = 197; 중동 및 북아프리카: N = 77; 라틴 아메리카: N = 127

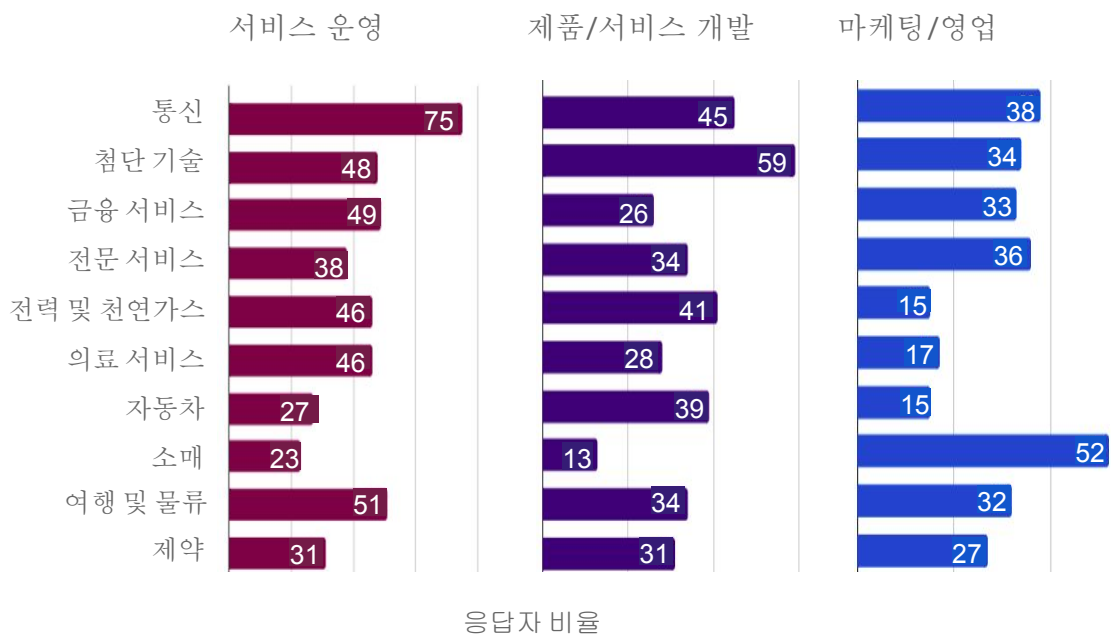
인공지능 도입: 산업 및 부문

본 페이지 및 다음 페이지 그래프는 2,135명을 대상으로 한 맥킨지 앤 컴퍼니의 설문 조사 결과를 나타낸다. 각 응답자는 자신이 속한 기관을 대표하여 응답하였다. 그래프는 자신이 속한 기관이 인공지능 기술을 특정 사업부문에 시범 도입하거나 도입하였다고 응답한 응답자의 비율을 나타낸다. 응답자는 제시된 기능 중 다수를 선택할 수 있었다. *제조(Manufacturing)*, *공급망 관리(Supply-chain management)* 및 *리스크(Risk)* 관련 설문 조사 결과는 다음 페이지 참고.

기관에서는 각자가 속한 산업에서 가장 높은 가치를 제공하는 부문에 인공지능 기술을 적용하는 경향이 있다. 예를 들어, *금융 서비스 분야(Financial services)*에서는 *리스크 분야*에 인공지능을 적용한 경우가 많은 반면, *자동차 산업(Automotive)*에서는 *제조 분야*에 인공지능을 적용하였다고 응답한 비율이 가장 높았고 *소매업(Retail)*에서는 *마케팅/영업 분야*를 선택한 비율이 가장 높았다. 이러한 결과를 통해 특정 기능(예를 들어, *제조*)에서 인공지능 기술의 발전 속도는 해당 기능이 특히 중요하게 활용되는 분야에서 인공지능이 사용되는 정도와 연관되어 있다.

데이터 및 방법론은 부록 13 참고.

산업 및 사업부문별 인공지능 도입(2018)
출처: 맥킨지 앤 컴퍼니

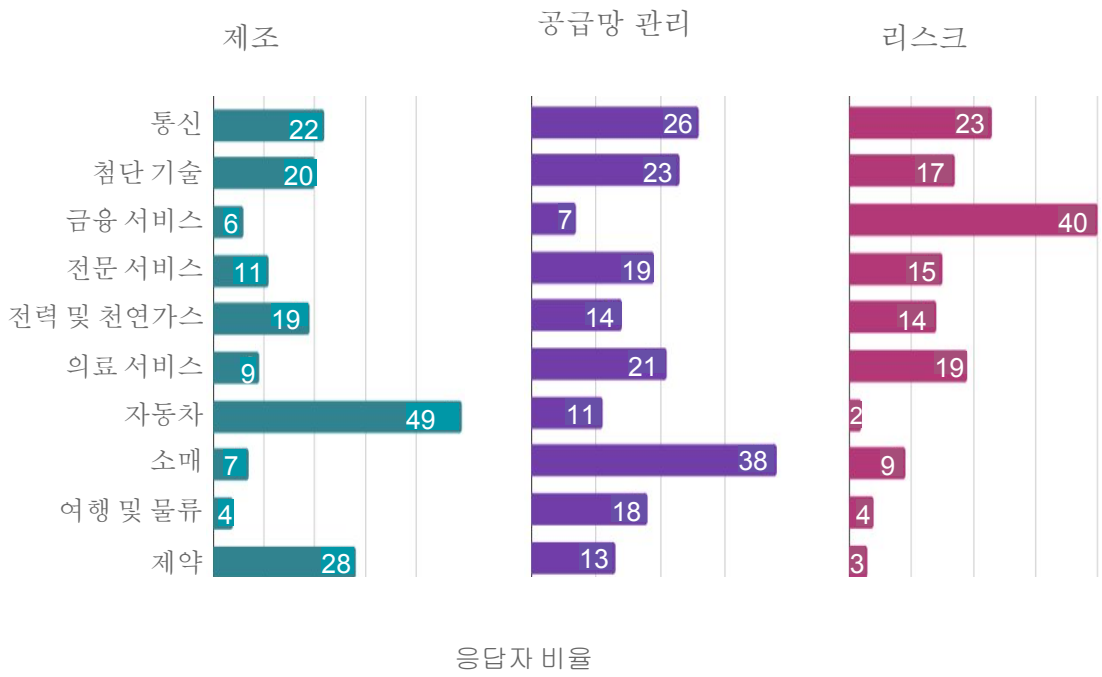


주석: 바의 길이는 각 산업에서 기능별로 인공지능을 도입한 비중의 상대적인 정도를 나타낸다. 통신: N = 77, 첨단 기술: N = 215, 금융 서비스: N = 306, 전문 서비스: N = 221, 전력 및 천연 가스: N = 54, 의료 시스템 및 서비스: N = 67, 자동차 및 조립: N = 120, 소매: N = 46, 여행, 운송 및 물류: N = 55, 제약 및 의료상품: N = 65

인공지능 도입: 산업 및 부문 (계속)

서비스 운영, 제품/서비스 개발 및 마케팅/영업 관련 내용 및 데이터 이전 페이지 보기.

산업 및 사업부문별 인공지능 도입(2018)
출처: 맥킨지 앤 컴퍼니



주석: 바의 길이는 사업부문별로 각 산업에서 인공지능을 도입한 비중의 상대적인 정도를 나타낸다; 통신: N = 77; 첨단 기술: N = 215; 금융 서비스: N = 306; 전문 서비스: N = 221; 전력 및 천연가스: N = 54; 의료 시스템 및 서비스: N = 67; 자동차 및 제조: N = 120; 소매: N = 46; 여행, 운송 및 물류: N = 55; 제약 및 의료 제품: N = 65.

기관에서는 각자가 속한 산업에서 가장 높은 가치를 제공하는 부문에 인공지능 기술을 적용

특정 부문에 관한 인공지능 기술의 발전 속도는 해당 부문이 특히 중요하게 사용되는 분야에서 인공지능이 활용되는 정도와 관련되어 있다.

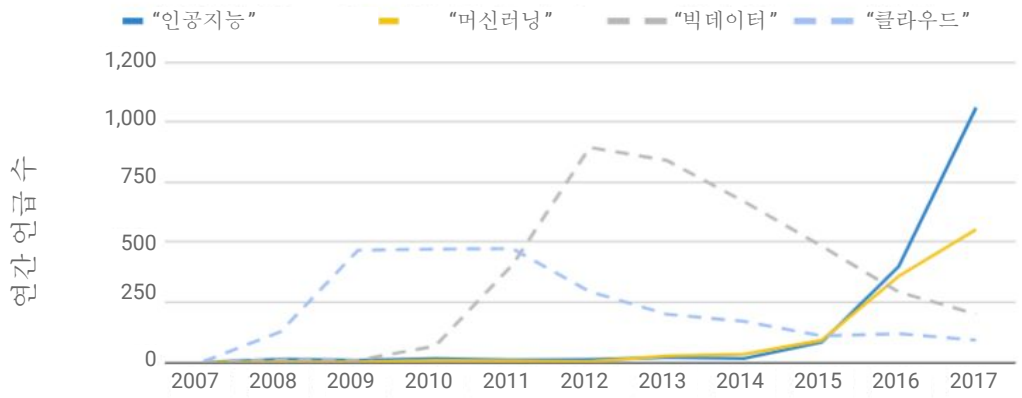
실적 발표 언급 빈도

이하 그래프는 인공지능(AI) 및 머신러닝(ML)이 기업 실적 발표에서 언급된 빈도를 산업별로 나타낸다. 첫번째 그래프는 IT 산업 기업들의 실적 발표에서 인공지능(AI) 및 머신러닝(ML)이 언급된 빈도를 나타낸다. 두번째 그래프는 IT 이외 산업 기업들의 실적 발표에서 인공지능(AI) 및 머신러닝(ML)이 언급된 빈도를 나타낸다. 빅데이터 및 클라우드는 인공지능/머신러닝이 언급된 빈도의 상대적 정도 비교를 위해 추가하였다. 본 분석은 뉴욕증권거래소에서 공개 거래되는 기업들의 데이터만 사용하였다.

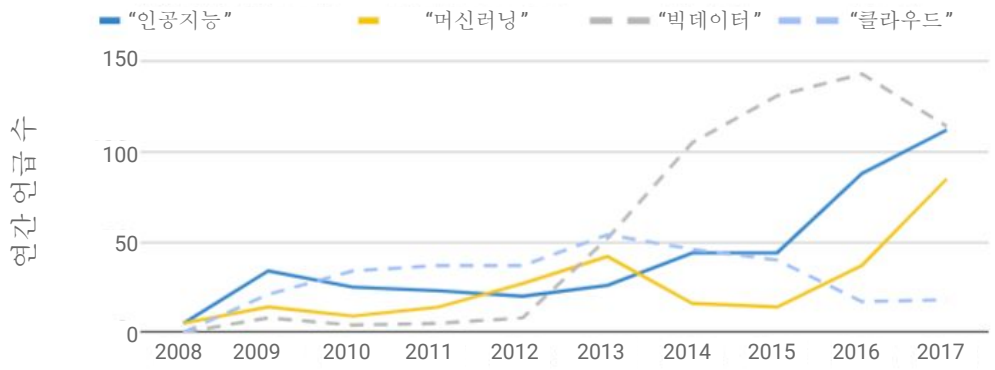
2015년 IT 산업에서 인공지능 및 머신러닝의 언급 빈도는 증가하였다. 대부분의 여타 산업에서 인공지능 및 머신러닝의 언급 빈도는 2016년부터 시작하였다. IT, 소비자 재량(Consumer Discretionary), 금융(Financial), 의료 서비스(Health Care) 산업 기업의 실적 발표에서 인공지능이 가장 많이 언급되었다.

데이터 및 방법론은 부록 14 참고.

기업 실적 발표 언급 빈도 – IT 기업(2007–2017)
출처: Prattle



기업 실적 발표 언급 빈도 – IT 산업 외 기업 언급 빈도 합계(2008–2017)
출처: Prattle



Top 10 회사 – AI 언급 (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 회사 – ML 언급 (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)

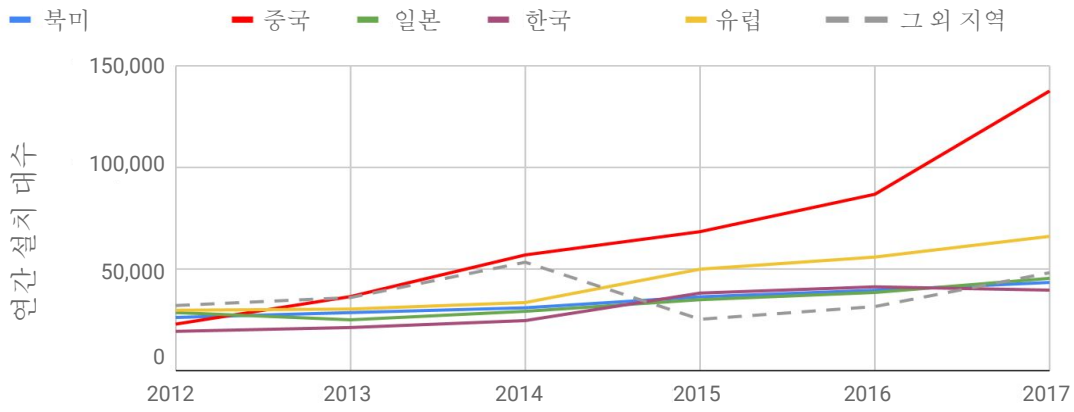
로봇 설치

이하 그래프는 지역별 연간 산업용 로봇 설치 대수를 보여준다. 첫번째 그래프는 로봇 설치 대수에서 상위를 차지하는 5개 지역을 보여주며 두번째 그래프는 5개 지역을 제외한 '나머지 지역'의 수치를 나타낸다.

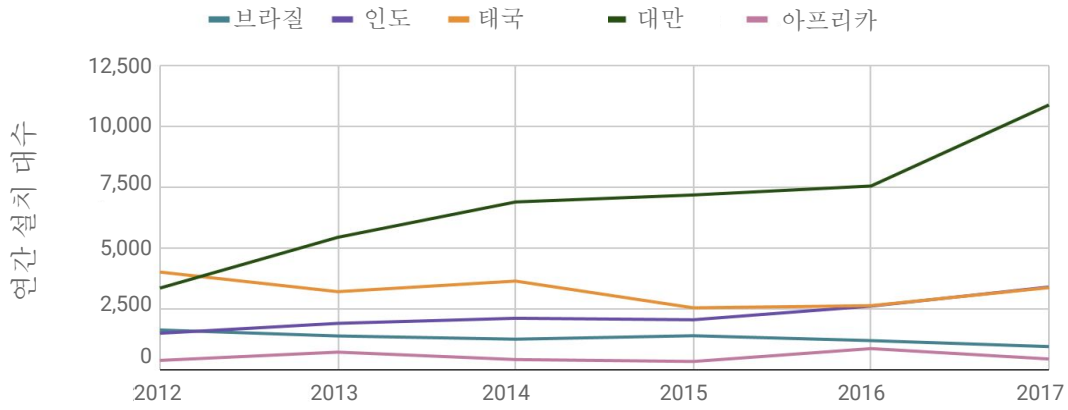
중국의 경우 2012년 이래로 연간 로봇 설치 대수가 500% 증가한 반면 한국 및 유럽은 각 105%, 122% 증가하였다. 그보다 적은 대수를 설치한 대만의 경우 2012년 – 2017년 가장 많이 증가하여 현재 사상 최고의 연간 설치 대수를 기록했다.

데이터 및 방법론은 부록 15 참고.

로봇 설치 – 상위 지역(2012–2017)
출처: ifr.org



로봇 설치 – 그 외 지역(2012–2017)
출처: ifr.org



주석: 설치 대수는 국내 생산 대수에 수입량을 더한 값에서 수출량을 뺀 값이다.

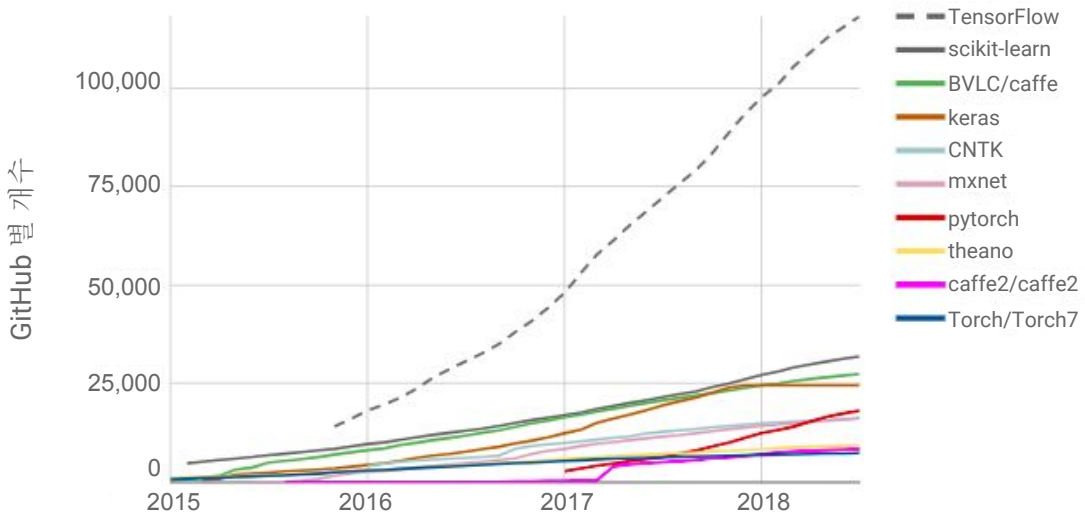
깃허브 별 개수 (GitHub stars)

이하 그래프는 인공지능 및 머신러닝 소프트웨어 패키지가 GitHub에서 받은 별의 개수를 나타낸다. 별의 개수는 다양한 인공지능 프로그래밍 프레임워크의 인기를 가능할 수 있는 대략적인 척도가 될 수 있다.

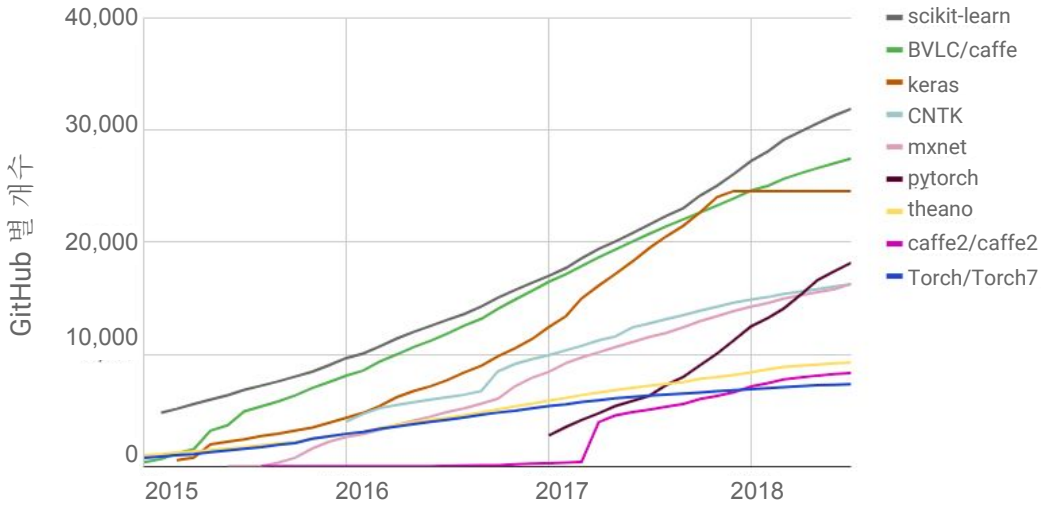
최근의 두가지 트렌드는 주요 기업들이 지지하는 프레임워크 *Tensorflow(Google)*, *Pytorch(Facebook)*, *mxnet(Amazon)*)의 인기가 높아지고 있다는 것과, 여타 언어와 비교하여 확연히 높은 TensorFlow의 인기이다.

데이터 및 방법론은 부록 16 참고.

인공지능 라이브러리별 누적 GitHub 별 개수(2015—2018)
출처: GitHub



인공지능 라이브러리별 누적 GitHub 별 개수—TensorFlow 제외(2015—2018)
출처: GitHub



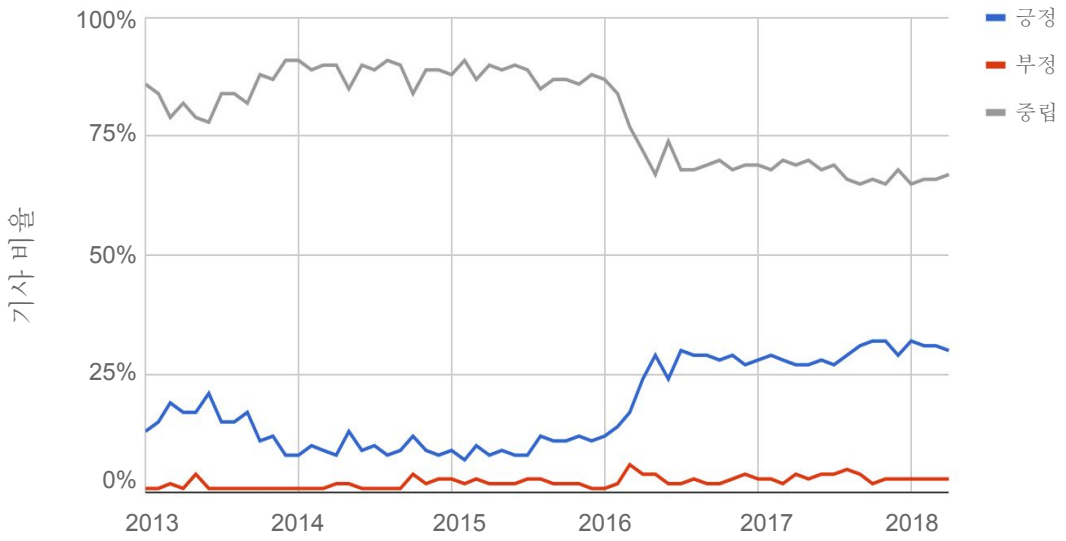
언론 보도 동향

이하 그래프는 *인공지능(Artificial Intelligence)*이라는 단어를 포함한 유명 언론 기사를 *긍정*, *부정*, *중립적* 기사로 분류하여 비율을 표시한 것이다.

특히 2016년 초부터 인공지능이라는 단어가 포함된 기사 중 중립적 기사의 비율을 감소하였고 긍정적인 기사의 비율이 증가하였다. 2016년 1월 긍정적 기사의 비율이 12%였다가 2016년 7월 30%로 증가하였다. 이후 긍정적 기사의 비율이 30%대를 유지하였다.

데이터 및 방법론은 부록 17 참고.

인공지능을 언급한 언론 보도 동향(2013–2018)
출처: TrendKite



2016년 – 2018년 인공지능에 관한 긍정적 기사 1.5배 증가

정부 언급

미국 의회 기록(U.S. Congressional Record과 캐나다와 영국 의회 의사록(Hansards)에 *인공지능*과 *머신러닝*이 언급된 횟수를 보여준다.

미국, 캐나다, 영국 의회의 기록에 따르면 *인공지능*과 *머신러닝*이 언급된 횟수가 2016년부터 증가하였다. 추가적으로 위 3개국에서 2016년 전까지는 머신러닝이 언급된 횟수가 매우 적었으며, 이후에도 여전히 *인공지능*에 비해 상대적으로 적게 언급되었다.

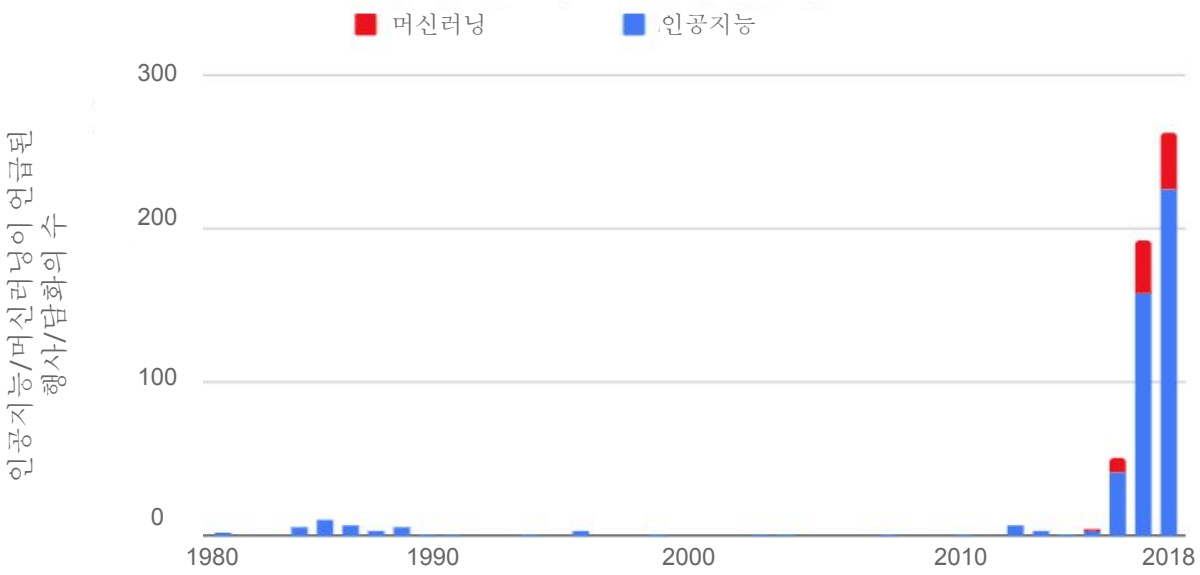
데이터 수집 방법의 차이로 인해 국가별 비교는 힘들다. 예를 들어, 미국 의회 기록의 경우 *머신러닝* 또는 *인공지능*이 특정 행사나 담화에서 한번 이상 언급되었을 경우 한번 언급된 것으로 계산한다. 반면, 영국 의회 의사록에 경우 *머신러닝* 또는 *인공지능*이 특정 논평에서 한번 이상 언급되었을 경우 한번 언급된 것으로 친다. 그러므로 국가별로 비교하기보다는 한 국가 내에서의 경향을 볼 것을 제안한다.

본 페이지의 그래프는 미국 의회에서 언급된 횟수를 나타내고 다음 페이지 그래프에서 캐나다와 영국의 경우를 다룬다. 데이터 및 방법론은 부록 18 참고.

미국 의회에서 인공지능 및 머신러닝 언급 횟수(1995–2018)
출처: 미국 의회 기록 웹사이트, 맥킨지 글로벌 연구소 분석

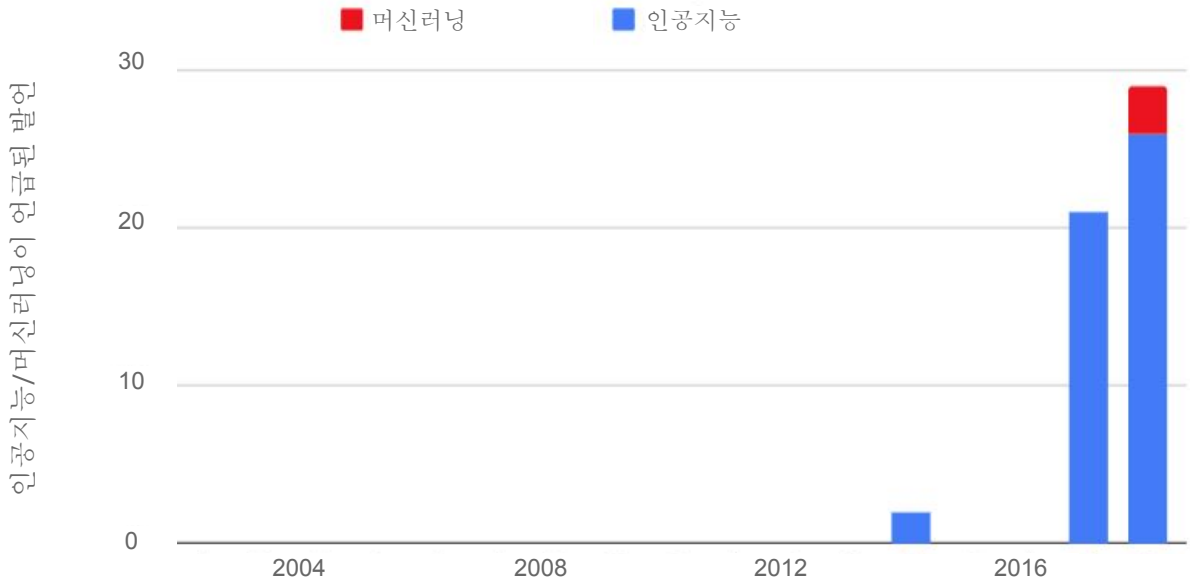
주석: 본 분석에서 미국 상원과 하원의 데이터를 모두 포함하고 있다. 2018년 데이터는 2018년 11월 20일까지의 데이터이다.

영국 의회에서 인공지능 및 머신러닝 언급 횟수(1980 –2018)
출처: 영국 의회 웹사이트, 맥킨지 글로벌 연구소 분석



주석: 영국 하원, 상원, 영국 국회의사당 및 위원회 데이터가 본 분석에 모두 포함되어 있다. 2018년 데이터는 2018년 11월 20일까지의 데이터이다.

캐나다 의회에서 인공지능 및 머신러닝 언급 횟수(2002–2018)
출처: 캐나다 의회 웹사이트, 맥킨지 글로벌 연구소 분석



출처: 본 분석에서는 캐나다 하원의 데이터만 다루고 있다. 2018년 데이터는 2018년 11월 20일까지의 데이터이다. 본 보고서는 비상업적 사용을 위한 비공식 자료이다. 본 그래프를 사용하고자 하는 분들은 하원 의장의 데이터 사용에 관한 허가가 금전적 이익을 위한 상업적 목적에는 적용되지 않는다는 것을 참고하길 바란다.

기술적 성능

객체 인식: ImageNet

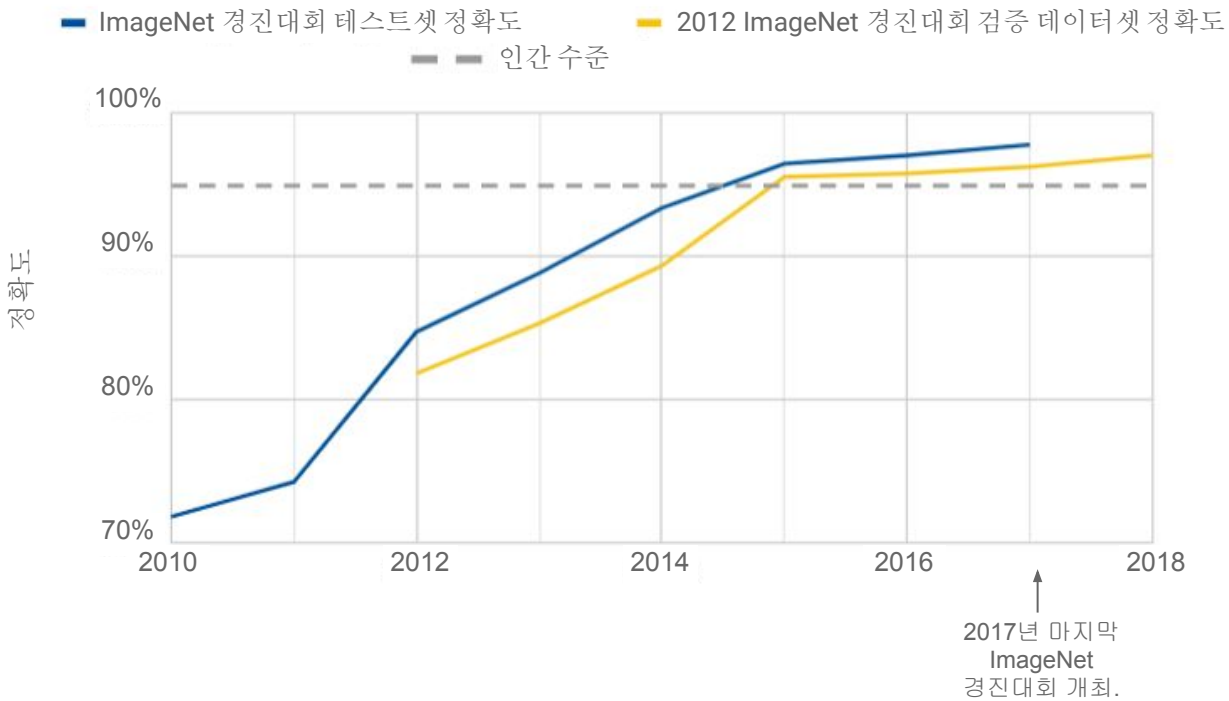
이하 그래프는 2010년부터 2018년까지 ImageNet의 정확도 점수 변화를 나타낸다. ImageNet 경진대회는 2017년까지 개최되었으며 지속되었던(held-out 대회에 특정된(competition-specific) '테스트' 데이터셋의 모델의 정확도를 측정했다. 더 이상 경진 대회가 개최되지 않기 때문에 AI 인덱스는 ImageNet 정확성 점수를 2012년 ImageNet 코퍼스의 검증 데이터셋(validation set from the ImageNet 2012 corpus)을 사용한 논문을 통해 추적하였다.

우리가 대회가 더 이상 개최되지 않는 상황에서 검증(validation) 데이터셋을 이용하여 결과를 측정하였기 때문에 더 이상 모델을 평가할 수 있는 지속된(held-out) 데스트 데이터셋은 존재하지 않는다. 발전의 정도를 분석하기 위해 우리는 2012 – 2018동안 2012년 ImageNet 코퍼스의 검증 데이터셋(validation set from the ImageNet 2012 corpus)을 사용한 주요한 결과들에 주목했다. 그래프에서 알 수 있듯이 ImageNet의 정확도는 지속적으로 높아졌다.

본 지표는 인공지능의 발전 추이를 모델링하는 것의 내재적 어려움을 보여준다. 만약 특정 연구 지표가 특정 대회를 중심으로 만들어졌다면 더 이상 해당 대회가 열리지 않을 경우 진정한 발전 정도를 측정하기 어렵게 된다. 그러나 개방된 데이터셋의 가용성으로 인해 약간의 데이터 처리를 통해 지속적으로 추적 가능하다.

데이터 및 방법론은 부록 19 참고.

ImageNet(2010 –2018)
출처: ImageNet; 부록 참고



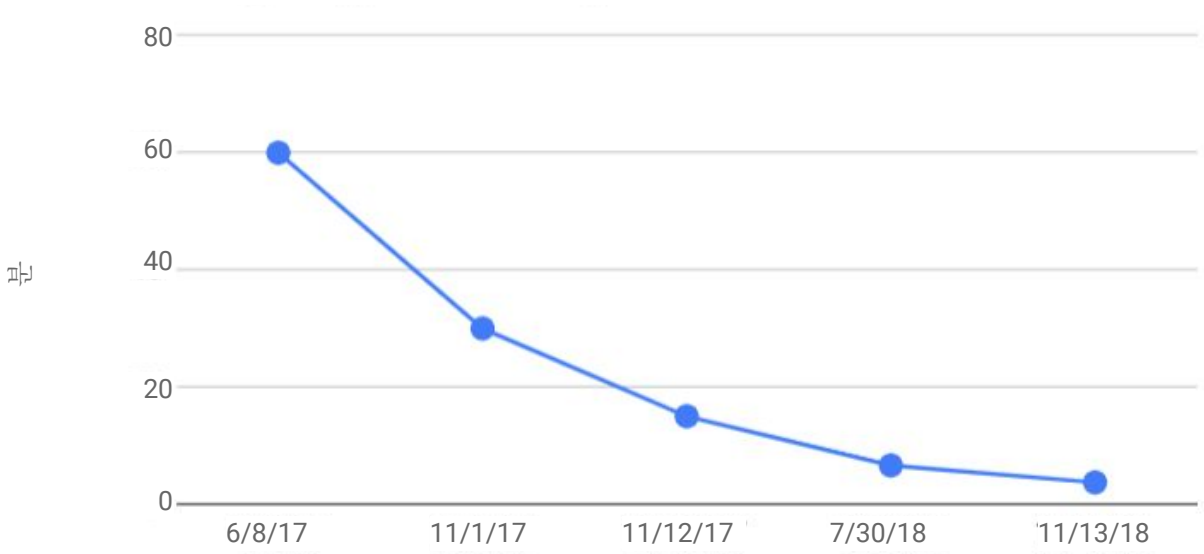
객체 인식: ImageNet 학습 시간

이하 그래프는 ImageNet 코퍼스(corpus)(이미지 데이터베이스)에서 높은 정확도로 이미지를 분류할 수 있도록 네트워크를 학습시키는데 필요한 시간을 나타낸다. 본 지표는 인공지능 분야에서 많은 자원을 가진 사람들이 이미지 분류(classification)와 같은 인공지능 기능을 수행하도록 대규모 네트워크를 훈련시키는데 드는 대략적 시간을 나타낸다. 이미지 분류가 상대적으로 일반적인 지도학습(supervised learning) 활동이기 때문에, 본 지표의 향상은 여타 인공지능 활용 기능의 학습 시간도 빨라지는 것과 연관되어 있다. 1년 반 사이에 네트워크를 훈련시키는데 필요한 시간이 한 시간에서 4분으로 단축되었다.

ImageNet 훈련시간 지표는 인공지능 연구의 산업화를 반영한다. ImageNet 훈련 시간 단축에 기여한 요소는 알고리즘의 혁신 및 기반시설 투자(예를 들어, 시스템을 훈련시키는데 사용되는 기반 하드웨어에 대한 투자, 또는 하드웨어를 연결하는 소프트웨어에 대한 투자)를 포함한다.

데이터 및 방법론은 부록 19 참고.

ImageNet 훈련시간(2017년 6월 — 2018년 11월)
출처: arXiv.org; 저자는 부록 참고



ImageNet 훈련 시간은 2017년 6월부터 2018년 11월까지 16배 단축

인스턴스 객체 세분화: COCO

이하 그래프는 컨텍스트 내의 공통 객체 경진대회(COCO; Common Object in Context Challenge)의 정확도를 보여준다.

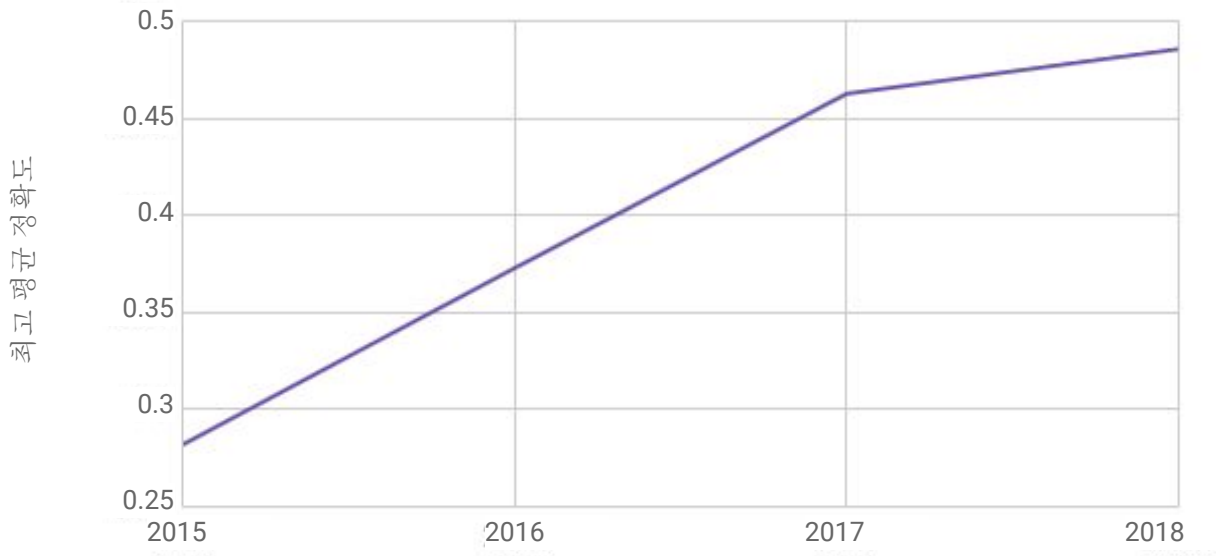
2017년 ImageNet 대회는 컴퓨터 비전 알고리즘이 ImageNet에서 제시한 객체 인식(object detection) 및 이미지 분류(image classification)에서 높은 성능을 보인 후로 중단되었다. 그 후 학계는 더 어려운 컴퓨터 비전 작업으로 옮겨갔다. 픽셀 수준의 정확도의 객체 위치 파악(localizing object)(객체 인스턴스 세분화(object instance segmentation)이라고 불림)와 같은 복잡한 논리를 요구하는 컴퓨터 비전 작업. 또한 풍경을 픽셀 수준의 정확도로 구역별로 나누는 작업(시맨틱 세분화(semantic segmentation)이라고 불림)으로 학계의 초점을 옮겼다.

COCO 지표는 위 언급한 구체적인 컴퓨터 비전 작업에 대한 성능을 보여준다. 본 분석의 목적은 각 원하는 객체의 해당 위치를 알아내고 해당 객체의 테두리를 정확히 파악하는 것이다.

2015년 이래로 COCO 대회에서 기록한 가장 높은 평균 정확도(precision)는 0.2 포인트 또는 72% 증가하였다.

사용된 데이터는 여기 참고, COCO 경진대회에 관한 추가 정보 참고.

COCO 객체 인스턴스 세분화 경진대회(2015—2018)
출처: COCO 리더보드



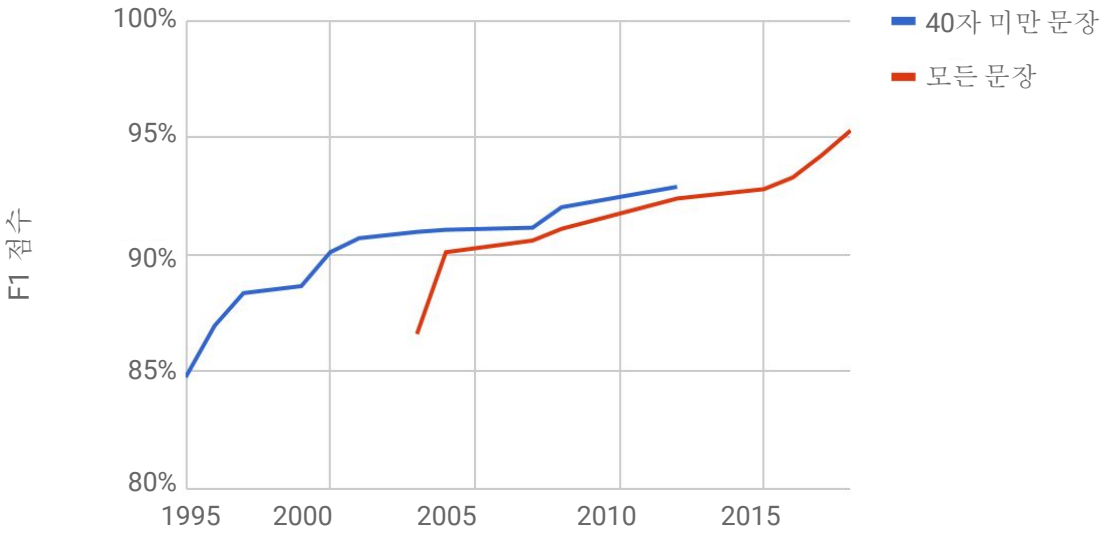
파싱 (Parsing)

이하 그래프는 인공지능 시스템의 문장 구문 구조(syntactic structure) 분석 성능을 나타낸 것이다.

파싱 지표는 질의응답과 같은 특정 작업에서 자연어를 이해하기 위한 첫 단계이다. 기존에 프로그래밍 언어를 파싱할 때 사용되는 알고리즘과 유사한 방식으로 이루어지던 것이 현재는 거의 일반적으로 딥러닝을 통해 파싱이 이루어지고 있다. 2003년이래 모든 문장의 F1 점수는 9 퍼센트 포인트(또는 10% 증가) 증가하였다.

데이터 및 방법론은 부록 20 참고.

문장 구성성분 분석 — Penn Treebank(1995—2018)
출처: 부록 참고



2003년부터 2018년까지 문장 구성성분 분석 성능 10% 향상

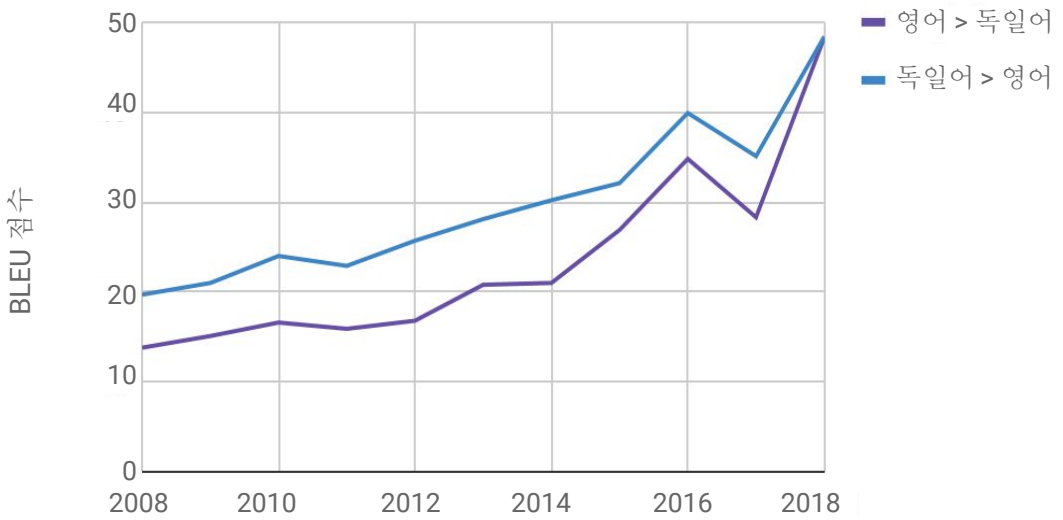
기계 번역

이하 그래프는 인공지능 시스템의 영어-독일어, 독일어-영어 뉴스 기사 번역 성능을 나타낸 것이다.

2008년과 비교하여 현재 영어-독일어 번역 BLEU 점수는 3.5배 증가하였다. 독일어-영어 번역은 동기간 2.5배 증가하였다. 매해 다른 테스트 세트를 사용하기 때문에 연도별 점수를 완벽히 비교할 수는 없다(우리는 이것이 2017년 점수가 하락하는데 영향을 미쳤다고 본다 – 부록에서 추가 정보 참고). 여전히 BLEU 점수는 기계 번역이 향상되었음을 보여준다.

데이터 및 방법론은 부록 21 참고.

뉴스 번역 — WMT 대회(2008—2018)
출처: EuroMatrix



2018년도 영어-독일어 번역 BLEU 점수 3.5배 향상

"최고 성능을 가진 시스템이 잘하는 몇가지 작업이 있다. 이 시스템들은 트랜스포머 아키텍처(transformer architecture)를 신경망 기계번역(Neural MT)에 이용(exploit)하고, 데이터를 더 보다 효율적으로 이용(exploit)한다. 트랜스포머 아키텍처를 통해 이전의 신경망 기계번역 모델(RNNs 기반)보다 향상된 결과를 얻을 수 있지만 이는 보다 주의 깊은 하이퍼 파라미터 선택과 최적화를 필요로 한다. .."
—베리 헤도우, 에든버러 대학교

상업적인 기계 번역 성능
상업적인 기계 번역 성능에 대해 더 알아보기.

질의응답: ARC

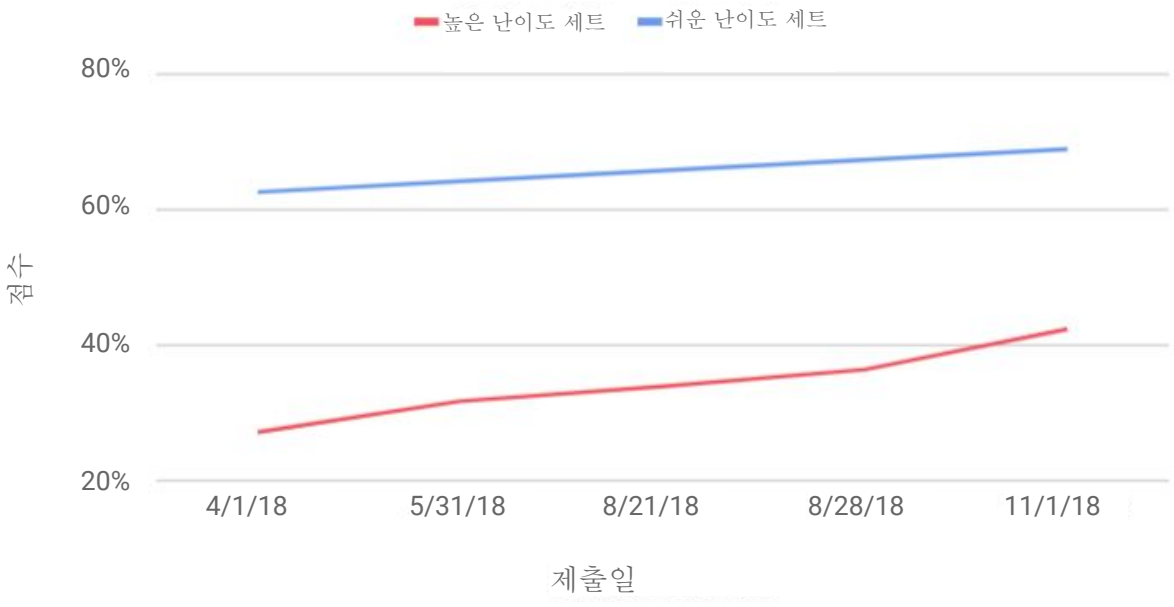
이하 그래프는 AI2 추론 경진대회(ARC; AI2Reasoning Challenge) 결과를 보여준다. ARC 데이터셋은 고난이도 질의응답 연구 촉진을 위하여 만들어진 7,787개의 초등학교(미국 3학년 - 9학년) 수준의 객관식질문을 포함하고 있다. 해당 질문은 높은 난이도 세트(Challenge Set, 2590개 질문)과 쉬운 난이도 세트(Easy Set, 5197개)질문으로 나뉘어져 있다.

높은 난이도 세트는 검색 기반(retrieval-based) 알고리즘과 단어 동시 출현(word co-occurrence) 알고리즘 모두 틀리게 답변한 질문들로 구성되어 있다. 질문은 글로만 이루어져 있으며 영문 버전 시험 문제는 파일에 표시된 것처럼 다양한 학년 수준의 질문으로 구성되어 있다. 각 질문은 객관식(보통 4지선다)이다. 질문은 ARC 코퍼스(corpus)를 동반한다. ARC 코퍼스는 1400만 개의 불규칙적인 질문으로 ARC와 관련된 내용을 포함한 과학 관련 문장이다. 질문에 대한 답을 코퍼스에서 찾을 수 있는 지는 확실하지 않다.

ARC 벤치마크는 2018년 4월에 발표되었다. 2018년 성능은 쉬운 난이도 세트의 경우 63%에서 69%로 높아졌고 높은 난이도 세트의 경우 27%에서 42%로 높아졌다.

데이터 및 방법론은 부록 22 참고.

ARC 리더보드(2018년 4월—2018년 11월)
출처: 앨런 인공지능 연구소



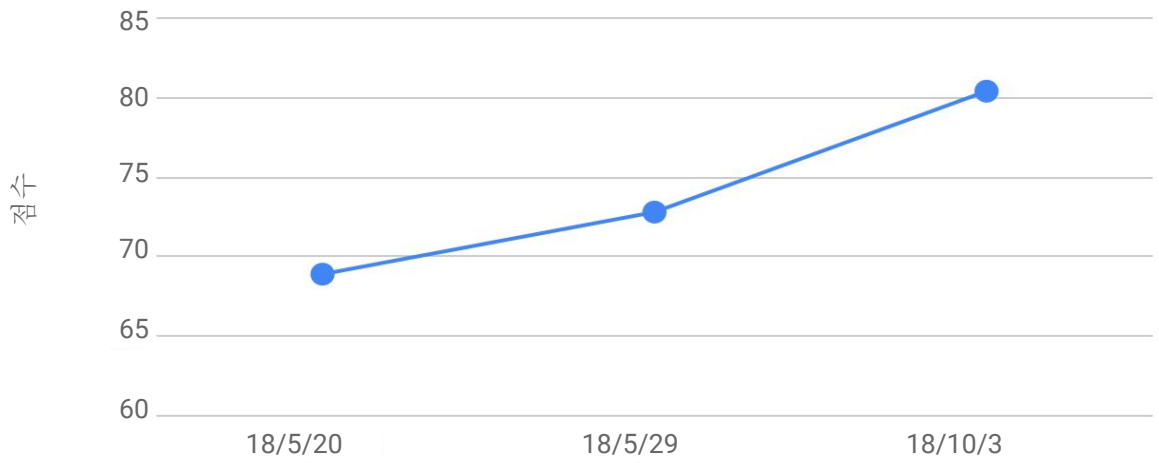
주석: 본 그래프는 제출된 답안 중 상위 답안을 추세선으로 연결한 것이다. 제출일과 관련한 추가 정보는 부록에서 확인 가능하다.

질의응답: GLUE

이하 그래프는 GLUE 벤치마크 리더보드의 결과를 보여준다. 일반적인 언어 이해 평가(GLUE; General Language Understanding Evaluation)는 자연어 이해(NLU) 시스템을 다양한 작업에서 시험하고 특정 작업에 맞추어 만들어진 것이 아닌, 시스템 개발을 독려하기 위하여 만들어졌다. GLUE는 9개의 하위 작업으로 구성되어 있다. 그 중 2개의 하위 작업은 하나의 문장(언어학적인 용인성(linguistic acceptability)과 분위기(sentiment) 측정을 위함), 3개의 하위 작업은 유사성과 다른 말로 바꾸는 것 그리고 4개의 하위 작업은 Winograd Schema Challenge를 포함한 자연어 추론 작업으로 구성되어있다. 코퍼스(corpora)의 크기는 1,000 미만에서 400,000 이상으로 다양하다. 본 지표는 정확도/F1과 Matthews Correlation Coefficient를 포함한다.

해당 벤치마크가 2018년 5월에 발표되었지만 이미 처음 발표한 베이스 라인(baseline)과 비전문가의 인간의 수준 추정 값인 90% 사이의 격차의 약 절반에 해당하는 향상이 이루어졌다.

데이터 및 방법론은 부록 23 참고.
GLUE 벤치마크 리더보드(2018년 5월—2018년 10월)
출처: Gluebenchmark.com



"우리는 아직 GLUE와 관련된 대규모 커뮤니티는 발견하지 못했지만 두개의 중요한 답안이 문장 이해를 위한 표현 학습(representation learning)에서 중요한 마일스톤으로 여겨진다. Google의 BERT는 최신 기술 수준을 보여주며, 제출된지 한달이 좀 넘었을 뿐인데 이미 8차례나 피인용되었다. Google BERT는 최근 열린 NLP 컨퍼런스인 EMNLP 에서 참가자들 사이에서 뜨겁게 논의된 주제였으며 조만간 선 학습된 표현(pretrained representation)에 관한 작업인지의 여부를 떠나 모든 언어 이해 문제에 있어 BERT가 기준(standard baseline)이 될 것으로 보인다."
—샘 보우만, NYU

기타 질의응답 데이터셋/작업
질의응답 데이터셋의 급격한 증가가 있었다. Squad 2.0; OpenBookQA; HotpotQA; MultiRC; NarrativeQA; DrQA; CoQA; QuAC; CWQ (rc setting)는 2018년 AI 인덱스 보고서에 포함되지 않았다.

기타 지표

파생 지표

*파생 지표*는 앞서 다룬 트렌드 사이의 관계를 분석한다.

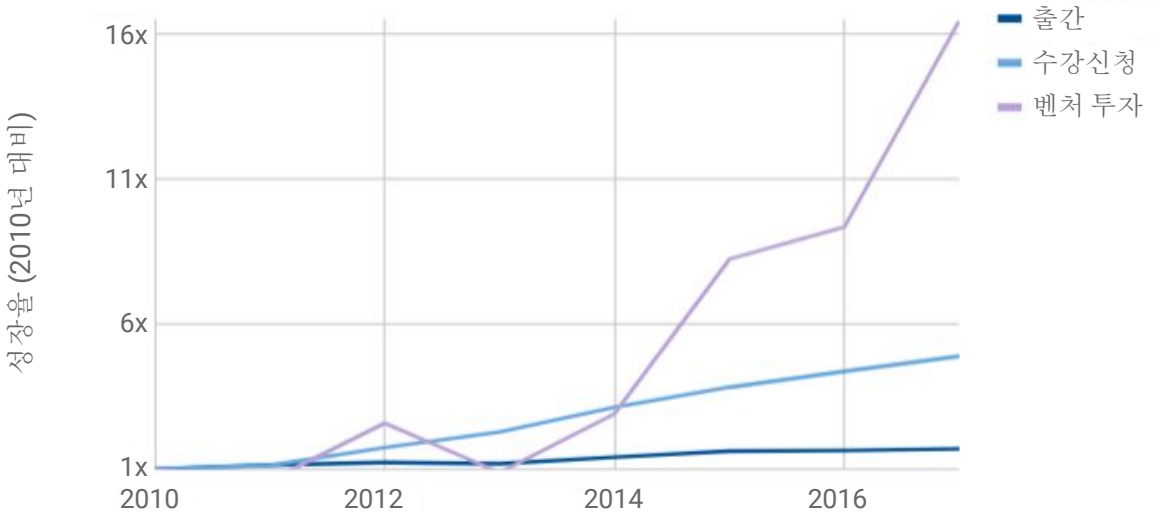
첫번째 지표인 *학계-산업계 관계(Academic-Industry Dynamics)*는 특정 학계 지표와 특정 산업계 지표를 비교하였다. 두번째 지표인 *AI 활성 지수(AI Vibrancy Index)*는 학계 지표와 산업계 지표를 하나로 통합한 지표이다.

학계-산업계 관계(Academia-Industry Dynamics)

학계의 인공지능 관련 활동과 산업계의 인공지능 관련 활동 간의 관계를 분석하기 위해 우리는 먼저 앞서 다룬 지표 중 일부를 활용하였다. 우리는 Scopus 인공지능 관련 논문 게재율, 몇몇 미국 대학교의 인공지능 및 머신러닝 개론 수업 수강신청율 및 인공지능 관련 스타트업 대상 벤처캐피탈 투자 지표를 활용하였다.

위 지표들은 직접적으로 비교할 수 없는 지표이다. 해당 지표간 관계를 분석하기 위하여 각 지표를 2010년부터 시작하는 것으로 정규화(normalize)하였으며 지표별 절대값 대신 지표별 증가율을 비교하였다.

학계-산업계 관계(2010 – 2017)
출처: Sand Hill Econometrics, Scopus, 대학교 제공 데이터



주석: 2017년 보고서는 스탠포드 대학교의 수강신청 지표만 사용하였으며 올해 수강신청 지표에 포함된 대학교 명단은 이 곳을 참고 바람.

AI 활성 지수 (The AI Vibrancy Index)

AI 활성 지수는 세가지 학계-산업계 지표(출간, 수강신청, 및 벤처캐피탈 투자)를 하나의 지표로 통합하였다. AI 활성 지수는 하나의 분야로서 인공지능의 활력(liveliness)을 정량적으로 측정한다. AI 활성 지수는 학계-산업계 관계처럼 2010년도 이후 지표값을 정규화하였다.

벤처캐피탈 투자가 다른 두 지표와 비교하여 가장 크게 성장하였기 때문에 AI 활성 지수는 크게 벤처캐피탈 투자 주도로 구성되어 있다(이전 페이지 참고). 3가지 지표는 모두 동일한 가중치로 사용되었다. 독자 여러분이 본 보고서의 데이터를 다운받아 지표 및 가중치를 조절하여 각자의 활용 목적에 맞는 활성 지수를 만들기 바란다.

출간: 출간 관련 부록 보기
수강신청: 수강신청 관련 부록 보기
벤처캐피탈 투자: VC 투자 관련 부록 보기

AI 활성 지수(2010 – 2017)
출처: Sand Hill Econometrics, Scopus, 대학교 제공 데이터

최근 정부 정책

이하 열거된 자료는 유럽연합 집행위원회의 합동 연구 센터(Joint Research Center) 및 팀 버튼(Tim Dutton)의 2018년 6월 국가별 AI 전략 개요(Overview of National AI strategies)에 의해 마련되었다.

미국	유럽 *(EU, 영국, 프랑스 정보만 포함. 여타 유럽 국가들의 정책은 포함되어있지 않음)*	중국
2016년 10월: 오바마 행정부는 국가적인 AI R&D 전략(national AI R&D strategy) 을 제안하였다. 해당 전략의 목적은 연구 투자 목적, 인간과 AI간 상호 협력 방법 개발, AI와 관련한 안전, 인종, 법적 사회적 함의에 대한 논의, AI 학습과 관련한 공공 데이터셋 마련, 기준과 벤치마크를 활용한 AI 기술 평가하려는 취지로 만들어졌다. 미국 정부는 AI에 관한 첫 정책 보고서인 "인공지능의 미래를 준비하며"(Preparing for the Future of Artificial Intelligence를 발간하였다. **2018년 5월**: 트럼프 행정부는 AI회의(*Summit on AI*) 에서 4가지 목표를 발표하였다. (1) AI 분야에서 미국의 선도적 지위 유지 (2) 미국 근로자 지원 (3) 공공 R&D 촉진 (4) 혁신을 방해하는 장애물 제거. AI R&D 관련 우선순위에 관해 백악관에 자문을 제공하고 학계 및 산업계 사이의 파트너십 체결을 고려하기 위하여 인공지능 특별 위원회 (*The Select Committee on Artificial Intelligence*) 를 구성했다.	***EU*** **2018년 4월**: EU 회원국은 인공지능에 관한 협력 선언 (*Declaration of Cooperation on AI*) 에 서명함으로써 함께 협업하기로 동의하였다. EU 집행위원회에서는 (1) EU의 연구 투자 촉진 (2) 사회경제적 변화에 대한 대비 (3) 윤리적이고 법적인 기본 틀의 설계를 목표로 인공지능에 관한 커뮤니케이션(*Communication on AI*)을 발표하였다. 커뮤니케이션 에서는 또한 2018년부터 2020년까지 인공지능 연구를 지원하기 위해 1.9조원(1.5B 유로)를 투자하였다. 해당 투자금은 Horizon 2020 프로그램에서 투자한 3.3조원(2.6B 유로)와는 별개이다. EU 집행위원회의 목표는 앞으로 10년동안 25.5조원 (20B 유로)을 투자하는 것이다. **2018년 6월**: the EU 집행위원회에서 2021년-2027년 예산을 11.7조원(9.2B 유로) 으로하는 디지털 유럽(Digital Europe) 프로그램을 제안하였다. 해당 프로그램은 고등 인공지능 기술 및 경제 및 사회 전반에 걸친 인공지능의 사용에 초점을 두고 있다. EU 집행위원회는 또한 모두가 사용할 수 있는 공통의 유럽 알고리즘 라이브러리 (European libraries of algorithms) 개발을 제안하였다.	2014년 이래로 중국 정부는 2018년까지 16.5조원($14.7B) 규모의 인공지능 시장을 중국에서 양성하고 중국이 2030년까지 세계에서 가장 앞선 인공지능 분야 국가가 되기 위한 목적으로 주요 인공지능 정책을 발표하였다. **2015년 7월**: *Internet+* 정책은 중국의 지능형 생산시스템에 초점을 맞춘 정책이다. 인공지능과 관련된 목표는 (1) 인공지능 개발에 관한 대중적 지지 확대 (2) 스마트홈, 스마트 자동차 및 로봇과 같은 분야에서 인공지능의 사용 촉진 (3) 음성, 이미지, 비디오, 지도, 연구를 포함하는 대규모 훈련 데이터베이스 개발 (4) 컴퓨터 비전, 언어 처리 및 인간-컴퓨터 상호작용과 같은 주요 인공지능 기술 개발 및 산업화를 포함한다. **2016—2020:** 로봇 산업 개발 계획(robot industry development plan)은 중국에서 지능형 산업 및 서비스 로봇 개발을 위한 계획을 제시한다.

주석: 추가 지표 다음 페이지 참고.

최근 정부 정책 (계속)

이하 열거된 자료는 유럽연합 집행위원회의 합동 연구 센터(Joint Research Center) 및 팀 버튼의 2018년 6월 국가별 AI 전략 개요(Overview of National AI strategies)에 의해 마련되었다.

미국	유럽 *(EU, 영국, 프랑스 정보만 포함. 여타 유럽 국가들의 정책은 포함되어있지 않음)*	중국
2018년 9월: DARPA에서 인공지능 기술의 한계를 극복하기 위한 목적으로 2조원 이상의 투자($2B+ investment plan)를 발표 하였으며 AI 넥스트(AI Next) 프로그램이 개시하였다. 미국 하원의 감독 및 정부개혁 위원회(the U.S. House Committee on Oversight and Government Reform) 산하 IT 소위원회 (Subcommittee on Information Technology)에서 인공지능 및 인공지능이 정책에 미치는 영향에 대한 백서 : "기계의 역습: 인공지능과 이의 증가하는 미국의 정책에 대한 영향력 (Rise of the Machines: Artificial Intelligence and its Growing Impact on U.S. Policy)을 발표하였다. 민간기업은 미국의 인공지능 개발/투자에서 핵심적인 역할을 한다. 2017년 아마존 및 알파벳과 같은 민간 기술 기업은 R&D에 각 18.1조원 ($16.1B)와 15.6조원($13.9B)을 투자하였다. 해당 투자 금액을 이해하기 쉽게 설명하자면 2019년 예산에서 미국 국립과학재단(NSF)의 예산 총액과 미국 방위고등연구계획국(DARPA)과 미국 교통부(DOT) 가 자율•무인 시스템 연구에 투자한 금액을 합해 5.9조원 ($5.3B)이었다.	**<u>프랑스</u>** **2018:** 프랑스에서는 최근 3편의 보고서를 발표하였다: (1) 연구, 교육 및 혁신 전략을 제안하는 프랑스 인공지능 계획(The French AI plan), (2) 과학 기술 의회 사무실(the Office of the Parliament for Science and Technology)에서 발표한 사회 및 규제와 관련한 문제에 초점을 맞춘 보고서, (3) 두뇌 유출을 줄이기 위한 정책에 초점을 맞추고 인공지능 연구에서 다양성의 중요성을 언급한 Mission Villani 보고서. **<u>영국</u>** **2016:** 영국 상원 인공지능 특별 위원회(The U.K. Parliament House of Lords Select Committee)에서 인공지능 정책 보고서(AI policy reports)를 발표하였다. **2018년 4월:** 정부에서는 연구, 교육 및 영국의 데이터 기반시설을 위해 1.35조원($1.2B)을 투자하는 *AI Sector Deal*을 발표하였다.	**2017년 7월:** 새로운 인공지능 시대 개발 계획(*New Generation AI Development Plan*)은 2030년까지의 인공지능 개발 전략을 제시한다. 해당 계획은 2020년까지 중국의 인공지능 기술이 세계에서 가장 발전된 수준에 이르도록 하는 계획을 담고 있다. 또한 2025년까지 중국이 인공지능 이론 분야에서 비약적 발전을 이루고, 산업 고도화 및 경제 구조 조정에 있어 인공지능이 원동력을 제공할 것이라고 밝혔다. 2030년까지는 중국이 세계에서 인공지능 혁신의 주요 중심지가 될 것이라고 하였다.

인간 수준의 성능 마일스톤

인간 수준의 성능 마일스톤

2017년 발표한 첫 AI 인덱스 보고서에서 인공지능이 인간 수준을 넘어선 사건의 타임라인을 다룬 바 있다. 해당 타임라인은 게임 성적, 정확한 의학적 진단, 및 기타 일반적이지만 고도의 작업에서 인공지능이 인간 또는 초인간적인 성능을 보인 경우에 대해 다루고 있다. 올해 우리는 2017년 타임라인에 4가지 새로운 성과를 추가했다. 이러한 결과를 확대 해석하지 않는 것이 중요하다. 이하 제시한 사건들은 매우 구체적이고 해당 예시들은 매우 놀랍지만 기타 작업 처리에 있어 시스템의 일반적인 성능에 대해서는 알려주는 바가 없다.

1980

오델로

1980년대에 리카이푸(Kai-Fu Lee)와 샌조이 마하잔(Sanjoy Mahajan)은 베이지안(Bayesian) 학습 기반 시스템을 이용한 오델로 보드 게임 프로그램 BILL을 개발하였다. 1989년 BILL은 컴퓨터 선수를 위한 전미 대회에서 우승하였으며 가장 높은 순위의 미국 선수 브라이언 로즈와의 경기에서 56대 8로 승리하였다. 1997년 로지스텔로(Logistello)라는 프로그램이 오델로 세계 챔피언과의 경기에서 6전 6승하였다.

1995

체커

1952년 아서 사무엘스(Arthur Samuels)는 자가 대국을 통해 성능을 향상시키는 일련의 체커 게임 프로그램을 만들었다. 그러나 1995년이 되어서야 체커 프로그램 치눅(Chinook)이 세계 챔피언을 제쳤다.

1997

체스

1950년대 일부 컴퓨터 과학자들은 1967년이면 컴퓨터가 인간 체스 챔피언을 능가할 것이라고 하였지만, 1997년이 되어서야 IBM의 딥블루(DeepBlue system) 시스템이 체스 챔피언 게리 카스파로프(Gary Kasparov)와의 경기에서 승리했다. 오늘날 스마트폰에서 구동되는 체스 프로그램은 그랜드 마스터 수준이다.

2011

제퍼디! (Jeopardy!)

2011년 IBM 왓슨(IBM Watson) 컴퓨터 시스템은 유명 퀴즈 쇼인 제퍼디!(Jeopardy!)1에 참가하여 전 챔피언인 브레드 러터(Brad Rutter), 켄 제닝스(Ken Jennings)와의 대결에서 우승하여 100만 달러 상금을 거머쥐었다.

2015

아타리(Atari) 게임

2015년 구글 딥마인드(DeepMind) 내 팀은 강화 학습 시스템(reinforcement learning system)을 사용하여 47개의 아타리 게임을 학습시켰다. 해당 시스템은 몬테수마의 복수(Montezuma's Revenge) 등과 같은 일부 게임에선 고전을 면치 못하고 있지만, 브레이크아웃(Breakout)을 포함한 대부분의 게임에서는 인간 수준의 수행능력을 달성할 수 있게 되었다.

인간 수준의 성능 마일스톤

2016

이미지넷(ImageNet)의 객체 인식

이미지넷(ImageNet)의 객체 인식에서 자동 라벨링(labeling) 오류율(error rate)이 2010년 28%에서 2016년 3% 미만으로 감소하였다. 동일 작업에서 인간의 수행능력(인간의 오류율)은 5%이다.

2016

바둑

2016년 3월 구글 딥마인드(DeepMind)에서 개발한 알파고(AlphaGo) 시스템은 세계 최고 바둑 선수 중 한 명인 이세돌 9단과의 경기에서 4승 1패로 승리하였다. 이후 2017년 3월에 딥마인드(DeepMind)는 알파고 마스터(AlphaGo Master)를 선보였고, 세계 랭킹 1위 커제 9단을 3대 0으로 꺾었다. 2017년 10월 구글 딥마인드(DeepMind) 팀은 영국의 과학 할술지 <네이처>지에 실은 논문에서 최신 버전인 알파고 제로(AlphaGo Zero)에 대해 상세히 기술하였으며, 알파고 제로는 이세돌 9단을 4대 1로 이긴 알파고와의 대결에서 100전 100승을 거뒀다.

2017

피부암 분류

2017년 <네이처>지에 따르면 앙드레 에스테바(Andre Esteva) 등은 2,032개의 질병에 대한 129,450건의 임상 이미지를 이용하여 인공지능 시스템을 훈련시켰고, 21명의 피부과 전문의들의 진단 능력과 비교를 수행했다. 그 결과 인공지능 시스템은 피부암 진단에 있어서 피부과 전문의에 견줄 수 있는 수준을 보였다.

2017

스위치보드(Switchboard) 테스트에서의 음성인식

2017년 Microsoft 와 IBM은 제한된 스위치보드(NIST Switchboard) 테스트에서 '인간 수준(human parity)'에 근접한 음성 인식률을 달성하였다.

2017

포커

2017년 1월 카네기멜론대학교(CMU)의 프로그램인 리브라투스(Libratus)가 네 명의 세계적 포커 선수들은 상대로 두 명이 벌이는 헤즈 업 노 리밋2 텍사스 홀덤(Head's Up No-Limit Texas Hold'em)이라는 포커 게임을 하였으며 총 12만번의 승부 끝에 리브라투스의 승리로 마무리되었다. 2017년 2월 캐나다 알버타 대학교의 인공지능 프로그램인 딥스택(DeepStack)은 11명의 프로 도박사들과 개인당 3천 회가 넘는 게임을 하였으며 딥스택은 도박사들과의 실력 격차가 '통계학적으로 유의미한' 수준으로 우위임을 드러냈다.

2017

미스 팩맨(Ms. Pac-Man)

마이크로소프트가 인수한 딥러닝 팀인 말루바(Maluuba)는 미스 팩맨 아타리 2600 버전의 최고 점수 999,990점을 얻을 수 있는 방법을 익힌 인공지능 시스템을 개발하였다.

인간 수준의 성능 마일스톤

2018

중국어 – 영어 번역

Microsoft의 기계 번역 시스템은 뉴스 기사를 중국어에서 영어로 번역하는데 있어 인간 수준의 정확도를 이루어냈다. 해당 테스트는 일반적으로 기계 번역 대회에서 사용되는 데이터셋인 newstest2017을 이용하여 이루어졌다.

2018

깃발 뺏기(Capture the Flag)

딥마인드가 Quake III Arena Capture the Flag(인기 3D 멀티플레이어 1인칭 비디오 게임)의 변형 버전에서 인간 수준의 성적을 기록했다. 딥마인드 에이전트는 길찾기, 따라가기, 방어와 같은 인간과 유사한 행동을 보였다. 학습된 에이전트는 같은 팀원 및 상대팀의 역할에서 인간 수준의 높은 성적을 넘어서는 승률을 기록하며 기존의 최신 시스템들을 능가했다.

2018

도타 2

오픈AI의 다섯개 신경망 네트워크로 이루어진 팀인 오픈AI Five가 도타2(제약 있음) 에서 아마추어 인간 팀과의 승부에서 승리했다. 오픈AI Five는 매일 180년어치에 달하는 게임을 자가 경기(self-play)을 통해 학습했다(오픈AI Five는 인간 프로 팀과의 경기에서는 패했기 때문에 아직 인간을 넘어선 수준은 아니다).

2018

전립선암 분류(Prostate Cancer Grading)

구글은 적출술 표본(prostatectomy specimen)을 통한 전립선암 분류에서 70%의 정확도를 보였다. 미국내 면허를 가진 일반 병리학자의 평균 정확도는 61%였다. 더불어 검증 셋의 모든 샘플을 평가한 상위 10명의 일반 병리학자의 정확도와 비교하면 딥러닝 시스템이 이 중 8명보다 높은 정확도를 보였다.

보고서에서 다루지 못한 내용

보고서에서 다루지 못한 내용

AI 인덱스는 여전히 진화하고 있으며 인덱스가 인공지능의 현 상황을 정확히 진단하기 위해서는 많은 추가적인 요소가 필요하다. 많은 요소를 추가하는 작업에는 인공지능 커뮤니티의 지속적인 참여와 협력이 필수적이다. AI 인덱스가 확대될 수 있는 방향에 관해 3명의 인공지능 전문가의 의견을 청취했다.

프란세스카 로시(FRANCESCA ROSSI)
IBM, 파도바 대학교*(휴가 중)*

상식추론 및 자연어 이해.
인공지능 기반의 질의응답 시스템은 매우 유용하지만 대화를 지속하는 능력에서 또한 부족한 점이 있다. 딥러닝 자연어 이해와 더불어 상황추론 능력(Scene reasoning capability) 여전히 쉽지 않은 작업이다. 자연어이해와 상황추론 능력을 모두 측정할 수 있는 지표가 대화 기반(dialog-based) 인공지능 시스템 발전 및 인공지능의 많은 부분의 발전에 도움이 될 것이다. AI 인덱스는 해당 기능과 관련한 테스트 및 도전에 있어 성과를 측정하는데 있어 추가적인 지표를 사용할 수 있다. AI2의 위노그래드 스키마 경진 대회(Winograd Schema Challenge)및 아리스토 프로젝트(The Aristo project)와 이와 관련된 과학 시험 관련 캐글 경진 대회(Kaggle competitions) 및 대화 기반 경진 대회인 IBM 토론가(IBM Debater)를 예로 들 수 있다.

인간과의 협력
만약 우리가 인공지능이 인간의 지능을 대신하기보다 향상시키기를 바란다면 인간vs기계에서 인간+기계 환경으로 바꾸는 것이 중요하다. 해당 분야의 발전을 측정할 수 있는 지표로는 인간이 최종 결정자이나 기계의 도움을 받는 비자율(non-autonomous) 시스템의 수를 이용할 수 있다.

AI 추론 + 학습
머신러닝은 대규모 데이터에서 숨겨진 연관관계를 찾을 수 있기 때문에 많은 응용 분야에서 성공적으로 활용되고 있다. 머신러닝을 통해 매우 정확한 예측을 할 수 있다. 그러나 많은 분야에서 예측을 넘어 개입을 통해 미래를 만들어 가기 위해서, 우리는 연관관계를 찾는 것을 넘어 인과관계를 찾을 수 있어야 한다. 이와 관련한 지표로 인공지능 컨퍼런스(머신러닝에만 국한된 컨퍼런스만이 아닌 AAAI나 IJCAI와 같은 AI 컨퍼런스 포함)에서 인과관계와 관련한 논문 수를 활용할 수 있다.

논문 제목을 검색하는 것만으로도 충분히 해당 지표를 만들 수 있다. 또한 머신러닝과 상징적 인공지능(Symbolic AI)을 결합한 연구를 모니터링하는 것도 중요하다. 관련 지표로는 머신러닝과 상징적 인공지능을 함께 연구한 논문 수를 활용할 수 있다.

보고서에서 다루지 못한 내용(계속)

로드니 브룩스(RODNEY BROOKS)
MIT

인공지능을 활용한 로봇
필자는 로봇 수출입(shipment) 지표에 관해 언급하고자 한다. 로봇 수출입 지표의 연관성은 출처에 달려있다. 많은 출처에서 인공지능이 거의 활용되지 않은(또는 전혀 활용되지 않은) 산업용 로봇 수출입을 다루고 있으며 이러한 데이터를 활용하는 것은 인공지능의 발전 정도를 측정하는 지표로 적절하지 않다. Rethink Robotics가 필자가 아는 예외적인 출처이다.

인공지능을 활용한 (SLAM과 다른 인공지능 알고리즘을 사용하는) 드론과 같은 로봇을 살펴볼 수도 있다. 룸바(Roomba)와 같은 가정용 로봇 또한 인공지능을 활용하고 있다. 룸바와 같은 경우에는 가장 많이 수출입 된 로봇이기 때문에 수치를 살펴보기에 흥미로운 로봇이다. 최근 가정용 소셜 로봇의 실패 사례가 몇몇 있었지만 해당 분야에서의 성공 사례는 인공지능에 크게 의존하게 될 것이다.

토비 월시(TOBY WALSH)
뉴사우스웨스트 대학교 시드니 및 베를린 공과대학교

정부 투자 및 군사적 사용
지난 2년간 인공지능 분야에서 가장 크게 성장한 것 중 하나가 바로 정부 및 산업계에서 발표한 관련 투자액일 것이다. 예를 들어, 영국 정부의 경우에는 10억 파운드(GBP)를, 프랑스의 경우에는 150억 유로를 그리고 독일의 경우에는 30억 유로를 투자하기로 약속한 바 있다. 동시에 알리바바와 같은 기업은 150억 달러를 투자할 계획이며 소프트뱅크의 비전 펀드는 1000억 달러에 달하는 투자금의 대부분을 인공지능과 관련해 투자하고 있다.

AI 인덱스는 인공지능 관련 투자액 증가를 측정하는 정량적 지표를 활용할 수 있다. 중국과 같은 국가에서 인공지능을 활용하여 경제적, 군사적 우위를 점하려고 하는 야심을 보이는 것을 우리는 종종 “AI 경주”(AI race)라고 부른다. 인공지능 연구와 개발의 초점이 어떻게 변화하고 있는지 측정하기 위해서 해당 지표를 사용하여 투자를 국가별로 분석할 수 있을 것으로 보인다. 많은 우려가 제기되는 또 다른 분야는 인공지능의 군사적 사용이다. 지뢰 제거와 같이 인공지능의 활용이 환영받는 분야도 있겠지만 완전히 자동화된 무기를 개발하는 것과 같은 분야에 대해서 관련 학계 및 산업계에서 심각한 우려를 제기하고 있다. AI 인덱스는 전세계 군사 분야에서 인공지능의 활용 및 개발을 측정할 수 있는 지표를 추가할 수 있다.

감사의 말

감사의 말

AI 인덱스는 인공지능에 대한 백년간 연구(One Hundred Year Study on AI (AI100). 에서 고안된 보고서이다. AI 인덱스는 AI100 산하에서 독립 프로젝트로 탄생했으며 현재 스탠포드 대학교의 인간중심 AI 연구소(Human-Centered AI Institute (HAI)) 에 소재해있다. AI100와 HAI의 지속적인 협력을 기대한다.

이와 더불어 구글, 마이크로소프트 바이트댄스(Toutiao) 및 AI 저널에서 스타트업 투자금을 제공하였다.

보고서에서 다루지 못한 내용 부분에 의견을 제공해주신 이하 분들께 감사드리며 해당 의견은 추후 발간될 AI 인덱스 보고서에 반영될 것이다:

프란세스카 로시, 로드니 브룩스, 토비 월시

또한 2018년 AI 인텍스 보고서에 포함된 지표와 관련하여 의견을 제공한 분들에게도 감사드린다:

폴 긴스파그, WiML 위원회, 마이클 추이, 베리 헤도우, 샘 보우만, AAAI,
제임스 A. 랜데이

시사점을 주는 데이터셋과 시각 자료를 제공한 Elsevier에 특별히 감사드린다. 많은 도움을 주신 이하 Elsevier 분들께 감사드린다:

앤 개브리엘, 클리브 마스틴, 사라 허커트, 마크 시베르트, 조그 헬위그

AI 도입에 관한 조사의 데이터를 제공하였으며 정부에서 인공지능 및 머신러닝에 대해 언급한 것을 분석하든데 도움을 준 맥킨지 앤 컴퍼니와 맥킨지 글로벌 연구소에게도 감사드린다. 많은 도움을 주신 이하 맥킨지 분들께 감사드린다:

헤더 한셀만, 다니엘라 세일러, 제이크 실버그

데이터 제공 및/또는 지표에 대한 설명을 제공하는데 각고의 노력을 해주신 분들께도 감사드린다:

수잔 우즈워드, 툴리 푸트, 샘 데이비스, 라스 코토프, 콘스탄틴 새벤코브, 아니타 황, 캘빈 르가식

2018년 보고서 발간에 핵심적인 역할을 한 이안 댄포스(테크니컬 리드)과 마이클 장(그래픽 디자인 및 커버 디자인 담당)에게도 감사드린다.

감사의 말

마지막으로 2018 AI 인덱스 보고서에 필수적인 데이터를 제공한 이하 기관, 대학교 및 컨퍼런스에 감사드린다:

기업 / 기관

arXiv, ROS.org, Sand Hill Econometrics, Monster.com, amplified, 맥킨지 앤 컴퍼니, 맥킨지 글로벌 연구소, Prattle, IFR.org, TrendKite, 앨런 인공지능 연구소, 유럽연합집행위원회 공동연구개발센터

대학교

UC 버클리, 스탠포드 대학교, 일리노이대학교 어바나-샴페인 캠퍼스, 워싱턴 대학교-시애틀, 카네기 멜론 대학교

칭화 대학, 국립천문학연구소, 광전자기술연구소, 브리티시 컬럼비아 대학교, 토론토 대학교, 에든버러 대학교, 중국과학기술대학, 상해교통대학교, 칠레 가톨릭 대학교, 빈 공과대학교, 히브리 대학교, 로잔 연방 공과대학교(EPFL), 몬트리올 대학 인공지능 연구소(MILA)

학회 및 단체

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

부록

부록 1
Scopus에 등재된 논문

주제별로 Scopus에 등재된 논문 관련 페이지로 돌아가기(9 페이지)
사용된 데이터 보기

출처
Elsevier 출판사에서 구축한 전 세계의 우수 학술논문 인용지수 Scopus에 7천만 여종의 논문이 등재되어 있다.
Scopus 에 대한 추가 정보 보기.

방법론
본 페이지에 서술한 방법론과 다음 페이지에 서술한 방법론의 차이를 이해하는 것이 중요하다. AI 인덱스 팀은 Scopus 쿼리 툴을 이용하여 *인공지능(Artificial Intelligence)*와 *컴퓨터 과학(Computer Science)* 키워드가 태그된 논문 수를 수집했다.

Scopus는 등재된 논문을 키워드, 출간일, 관련 국가 및 기타 지표로 태그해놓는다. 연도 및 지역에 따라 인공지능 관련 논문을 조사하기 위해 'opensearch:totalResults' 툴을 사용하였으며 이하 질의(query)를 사용하였다:

국가별 인공지능 관련 논문 질의
title-abs-key(artificial intelligence) AND PUBYEAR = {} AND AFFILCOUNTRY({})

주제별 연간 출간된 논문에는 다음 질의가 사용되었다:

인공지능 관련 논문 질의
title-abs-key(artificial intelligence) AND PUBYEAR AFT {} AND PUBYEAR BEF {}
컴퓨터 과학 관련 논문 질의
SUBJAREA(COMP) AND PUBYEAR AFT {} AND PUBYEAR BEF {}
Scopus에 등재된 모든 논문에 관한 질의
PUBYEAR AFT {} AND PUBYEAR BEF {}

위의 모든 질의는 1996년부터 2017년까지 연도별로 만들었다.

뉘앙스
- Scopus 시스템은 소급적으로 업데이트 된다. 결과적으로 Scopus 시스템에 등재된 논문 수는 시간이 지남에 따라 증가할 수 있다.
- Elsevier사에서는 1995년 후에 출간된 논문부터 질의가 가장 신뢰 할만한 결과값을 내놓는다고 조언했다(그러므로 Scopus에서 1996년 이후로 출간된 논문에 관한 데이터만 수집하였다).
- Scopus 논문 수는 연관성이나 논문 품질과 무관하게 인공지능과 관련된 모든 논문 수를 포함한다.

지역별 인공지능 관련 논문 관련 뉘앙스
- 논문이 복수의 지역에 태그된 경우 중복으로 계산되었다. 해당 이유로 특정 연도의 수치가 작년 AI 인덱스 보고서에 표시된 수치와 상이할 수 있다.
- "기타" 지역은 Scopus에 등재된 인공지능 관련 논문을 출간한 모든 기타 국가를 포함한다.

주제별 논문 관련 뉘앙스
- 2017 AI 인덱스에서는 컴퓨터 과학 범주에 속한 인공지능 관련 논문만 수집하였다. 올해는 컴퓨터 과학 분야에 속하는지와 무관하게 인공지능이 테그된 모든 논문 수를 수집하였다.
- Scopus에 등재된 인공지능 관련 논문이 일반적으로 넓은 컴퓨터 과학 범주에 포함되기 때문에 컴퓨터 과학 논문 수의 증가가 인공지능 관련 논문 수의 증가를 포함한다.
- "모든 논문"의 범주는 모든 인공지능 및 컴퓨터 과학 관련 논문을 포함한다.

부록 2
Elsevier의 Scopus 논문 분석 방법론

논문 관련 페이지로 돌아가기(10 페이지)
사용된 데이터 보기

출처
Elsevier 출판사에서 구축한 전 세계의 우수 학술논문 인용지수 Scopus에 7천만 여종의 문서가 등재되어 있다. 해당 데이터는 Elsevier 출판사에서 수집하였다. Scopus 에 관한 추가 정보 보기. Elsevier에 대한 추가 정보 보기.

논문 인덱싱, 소속 정보, 지리적 범위 및 제목에 관한 상세 분석방법은 Scopus 컨텐츠 커버리지 가이드에서 확인할 수 있다.

방법론
본 페이지에 서술된 분석방법과 이전 페이지에 서술된 분석방법의 차이를 이해하는 것은 중요하다. 본 페이지에서는 인공지능 관련 논문을 약 800개의 키워드를 사용하여 상향식 방법으로 분석하는 Elsevier의 방법에 대해 다루었다. Elsevier의 최신 AI 보고서 에서 상세한 내용을 확인할 수 있다.

인공지능에 대한 정의
Elsevier는 인공지능과 관련된 약 800개의 키워드를 정리했다. 해당 키워드는 Scopus에 등재된 60만편 이상의 인공지능 관련 논문을 분석하는데 사용되었으며 Elsevier의 AI 보고서의 기초가 되었다. 해당 키워드는 이하의 방법으로 만들어졌다:

먼저 Elsevier는 핑거프린트 엔진(Fingerprint Engine)을 활용하여 인공지능 관련 책, 강의 계획서, 특허 및 신문 기사를 수집하였다. 핑거프린트 엔진으로 수집된 문서를 분석하여 가장 많이 사용되는 관련 키워드를 정리하였다. 해당 작업을 통해 2만개 관련 용어를 추려냈다. 해당 용어들을 사람이 직접 검토하여 797개의 개별 키워드를 선택하였다. Elsevier는 몇 달 안에 해당 키워드를 공개할 것으로 보인다. 샘플 리스트 확인하기.

그러나 797개의 인공지능 관련 키워드 중 하나 또는 그 이상을 포함한 문서 중 실제로는 인공지능 관련 문서가 아니다. 이는 잘못된 긍정(false positive) 때문인데 잘못된 긍정이란 예를 들어 *도덕적 가치*와 같은 용어는 너무 광위하여 인공지능 관련 문서가 아닌 경우에도 포함할 수 있다. 이와 유사하게 *신경망 네트워크*란 용어는 생물학과 같은 비인공지능 분야의 용어이다.

이러한 잘못된 긍정을 해결하기 위해 Elsevier에서는 무엇이 코어 AI(core AI)이며 무엇이 응용 AI(applied AI)인지에 관한 1,500편의 문서를 사람이 직접 선별하여 만든 학습 셋을 이용하여 전문가 가이드라인을 만들었다. Elsevier에서는 해당 전문가 가이드라인을 통제된 분류기(supervised classifier)를 이용하여 잘못된 긍정을 제거하고 Scopus에 등재된 60만편의 문서를 분석하였다.

결과적으로, Elsevier의 방법은 복수의 관점과 키워드를 이용하고 통제된 분류기의 정확도를 더해 인공지능 분야를 광범위하게 분석하는 목적을 가지고 있다.

소속 국가 정보(Country Affiliation)
저자의 소속 국가 정보는 해당 저자의 주 소속 기관에 따라 결정된다. 해당 정보는 저자가 논문에 기재한 것을 바탕으로 한다. 국제적 기관의 경우, 저자가 소속 기관 설명에 구체적인 국가를 명시하지 않으면 해당 기관의 본부가 위치한 국가가 소속 국가 정보가 된다. 예를 들어, 소속 기관을 “구글” 이라고 한 저자의 경우 소속 국가 정보가 미국이 되고 소속 기관을 “구글 스위스” 라고 한 경우 유럽이 소속 국가가 된다.

여러 지역의 저자가 공저한 논문의 경우 중복 계산되었다. 예를 들어 하버드 대학교 소속 저자와 옥스포드 대학교 저자가 공저한 논문의 경우 미국과 유럽에 모두 포함되었다.

유럽은 EU44로 정의된다.

위 서술한 기준은 분야 소속 정보에도 적용된다.

부록 2
Elsevier의 Scopus 논문 분석 방법론 (계속)

인공지능 세부분야별(그래프 보기)
본 페이지에서는 Scopus에서 인공지능 하위 분류를 정하는 Elsevier의 방법에 관한 정보를 다루었다. Elsevier의 방법에 대한 추가 정보는 Elsevier 최신 AI 보고서.

광범위한 하위 분류를 정하기 위해서 Elsevier에서는 루벤 방법(Louvain method)이라고 불리는 비지도 클러스터링 기법을 활용하였다. 해당 방법은 문서에서 동시출연(co-occurance)하였는지를 바탕으로 키워드를 클러스터로 분류하여 연결 고리를 표시한다. 동시출연은 해당 클러스터가 독립된 것이 아니며 서로 간에 긴밀히 연관되었다는 것을 보여준다. 컴퓨터 비전 관련 문서에 나타난 신경망 네트워크를 예로 들 수 있다. Elsevier에서는 인공지능 분야가 이하 분야에 클러스터 된 것을 발견했다:

확률적 추론, 신경망 네트워크, 컴퓨터 비전, 검색 및 최적화, 자연어 처리 및 지식 표현, 퍼지 논리, 플래닝 및 의사결정

Elsevier AI 리소스 센터는 클러스터에 관한 반응형 그래프를 만든 바 있으며 유저들이 개별 연결 고리와 클러스터를 지역별 시간별로 볼 수 있게 하였다.

상대적 활동 초점(그래프 보기)
논문은 각 논문의 범위와 내용에 따라 주제 분야를 분류하는 전문색인가들에 의해 주제 분류가 이루어졌다.

뉘앙스

- 모든 데이터는 Elsevier에서 내부적으로 수집하였으며 AI 인덱스에서는 키워드 결정 과정이나 관련 논문 수를 세는 작업에 관여하지 않았다.
- 인공지능의 하위 분야에 속하는 논문 수는 다른 분석 방법이 사용된 2017년 AI 인덱스 보고서에 포함된 결과에 약 두배에 달하는 수치이며 중국 인공지능 개발 보고서에 보고된 인공지능 관련 논문 수에 비하면 절반 정도이다. 중국 인공지능 개발 보고서는 Web of Science 데이터를 기반으로 만들어졌다.
- 음성 인식, 컴퓨터 비전, 로봇 공학, 사이버 보안, 생물정보학, 의료 서비스 등의 다양한 분야에서 인공지능이 급속도로 많이 활용되고 있기 때문에 인공지능의 범주를 설정하는 것을 어려운 작업이다. 인공지능의 범주를 설정하는 것이 어려운 또 다른 이유는 인공지능이 논리, 확률 및 통계, 최적화, 사진측량, 신경과학 및 게임이론 등의 분야와의 방법론적 관련성 때문이다. 인공지능 계량서지학(AI bibliometrics)에 관한 커뮤니티의 관심을 고려할 때, 관련 연구를 하는 연구자들이 특히 다양한 문헌 데이터베이스에서 결과를 재현할 수 있는 방법에 대해 투명성을 높이는 것이 가치 있을 것으로 보인다.

부록 3
arXiv 논문

arXiv에 수록된 논문 관련 페이지로 돌아가기(12 페이지)
사용된 데이터 보기

출처
arXiv.org는 물리학, 수학, 컴퓨터 과학, 정량적 생물학, 통계학, 전기 공학 및 시스템 과학, 경제학 등의 분야 논문을 모아놓은 온라인 아카이브이다. arXiv는 코넬대학교에서 소유 운영하고 있다. arXiv.org 에 대한 추가 정보 보기.

방법론
분석에 사용된 데이터는 arXiv.org로부터 제공받았다. 키워드를 선별하였으며 키워드가 속한 카테고리는 아래와 같다:

인공지능 - Artificial intelligence (cs.AI)
컴퓨테이션 및 언어 - Computation and language (cs.CL)
컴퓨터 비전 및 패턴 인식 - Computer vision and pattern recognition (cs.CV)
머신러닝 - Machine learning (cs.LG)
신경망 및 진화 컴퓨팅 - Neural and evolutionary computing (cs.NE)
로봇학 - Robotics (cs.RO)
통계학에서의 머신러닝 - Machine learning in stats (stats.ML)

대부분의 카테고리에서 arXiv는 1999년부터 2017년까지의 데이터를 제공하였다. AI 인덱스 팀은 통계학에서의 머신러닝을 포함하기 위해 2010년 자료부터 분석하였다. 통계학에서의 머신러닝 카테고리는 2010년 이전에는 arXiv에 존재하지 않았다.

카테고리별 arXiv 제출율은 arXiv.org 논문 제출 통계 참고

뉘앙스
- 카테고리는 논문 저자에 의해 지정되며 주요(Primary) 카테고리로 지정된 항목이 해당 논문의 카테고리가 된다. 그러므로 한가지 간소화된 카테고리 지정 절차가 있는 것은 아니다. 더불어 인공지능 또는 머신러닝 카테고리가 다른 하위 분야/키워드로 분류될 수도 있다.
- arXiv 팀에서는 arXiv에 참여하는 것이 더 많은 참여를 낳을 수 있다고 하였다. arXiv에 하위 카테고리가 늘어나는 것이 특정 커뮤니티의 과한(over-indexed) 참여로 이어질 수 있다.
- arXiv에 저장된 논문 수가 늘어나는 것이 해당 주제에 관한 논문 수 증가를 반영하는 것은 아니다. arXiv의 논문 수를 늘리기 위한 노력 덕분에 또는 인공지능 커뮤니티에서 지식 전파의 중요성이 커진 것 등이 논문 수 증가로 이어졌을 수 있다.

부록 4
국가별 AAAI 논문

국가별 AAAI 논문 관련 페이지로 돌아가기(20 페이지)
사용된 데이터 보기

출처
전미인공지능학회(AAAI)는 연례 AAAI 컨퍼런스를 포함하여 매년 컨퍼런스를 주최한다. 2018 AAAI 논문 제출/채택율에 관한 국가별 원본 데이터(raw data)는 AAAI에서 제공하였다. AAAI 컨퍼런스에 관한 추가 정보 보기.

방법론
AAAI의 국가별 논문 제출/채택율 정보를 수집하였다. AAAI가 AI 인덱스 팀에서 얻을 수 있는 가장 상세한 정보를 제공한 유일한 곳이었다. AI 인덱스에서는 추후 보고서에 AAAI 컨퍼런스 외에도 다른 컨퍼런스 관련 상세 데이터를 포함시키길 희망한다.

뉘앙스
- 본 분석에 포함된 국가는 AAAI 컨퍼런스에 10편 이상의 논문을 제출한 국가이다.
- 데이터는 2018년 2월 루이지애나 주, 뉴올리언즈에서 개최된 2018년 컨퍼런스 데이터이다. 제출/채택된 논문 경향이 이전 년도들과 상이할 수 있다.
- 논문 채택은 제한된 학회지 공간 때문에 한정적일 수 밖에 없었다. 2019 AAAI 컨퍼런스는 학회지의 공간적인 제약에 관한 AAAI의 이하 발언을 참고하기 바란다.
- “올해(1/27/19–2/1/19 호놀룰루, 하와이) AAAI에 7,700편 이상의 논문이 제출되었다. 이 중 7,095편의 논문을 검토하였으며 학회지의 공간적인 제약 때문에 오직 1,150편의 논문만이 채택되었으며 채택율을 16.2%였다:

  “올해(2019년) AAAI에 7,700편 이상의 논문이 제출되었다. 이 중 7,095편의 논문을 검토하였으며 학회지의 공간적인 제약 때문에 오직 1,150편의 논문만이 채택되었으며 채택율은 16.2%였다. – AAAI

부록 5
인공지능 및 머신러닝 수강신청(미국)

미국 수업 수강신청 관련 페이지로 돌아가기(21 페이지)
사용된 데이터 보기

출처
수강신청 데이터는 각 대학교를 통해 수집하였다. 전체 학생 수는 각 학교 아카이브(주로 교학과 사이트에 위치)를 통해 수집하였다. 본 분석에는 이하 대학교가 포함되었다:

UC 버클리, 스탠포드 대학교, 일리노이대학교 어바나-샴페인 캠퍼스, 워싱턴 대학교-시애틀, 카네기 멜론 대학교

방법론
AI 인덱스 팀에서는 미국의 유수 컴퓨터 과학 분야 대학교의 인공지능 및 머신러닝 개론 수업 수강신청 수를 요청했다. 수강신청에 관한 충분한 역사적 자료가 없거나 데이터의 뉘앙스가 클 경우 해당 대학교는 본 분석에 포함하지 않았다.

일부 학교에서는 학기별 수강신청 정보를 제공하였으며 일부에서는 학년도별 정보를 제공하였다. 각 학교의 수강신청 정보를 비교하기 위하여 우리는 학기별 수강신청 데이터를 학년도별로 변환하였다. 더불어, 일부 학교에서는 다양한 수업을 개론 수업으로 분류하는 반면 일부 학교에서는 개론 수업이 한 개인 경우가 있었다. 적절하고 연관성이 있을 경우 여러 수업을 합치는 방법으로 인공지능 개론 수업 수강신청 경향을 표시했다.

학부생의 수강신청 비율 계산을 위하여 각 년도의 인공지능/머신러닝 수업 수강신청 학생수를 동일 연도의 학부생 수로 나누었다. 해당 방법으로 계산한 이유는 각 학교별 수강신청 경향을 동일선 상에서 비교하기 위함이다.

뉘앙스
- 거의 대부분의 학교에서 특히 최근 수강신청은 학생들의 수요 보다는 공급에 의한 것이라고 언급하였다. AI 인덱스의 데이터는 성공적으로 수강신청한 학생 수를 나타내며 대기 리스트(waitlist)나 기타 요구 지표(demand metrics)는 고려하지 않았다.
- 수업들은 주로 학부생에게만 열려 있으며 일반적으로 전공자와 비전공자 모두 수강할 수 있다. 일부 수업은 수업명이 바뀌었으며 AI 인덱스의 데이터는 2017년도 기준 수업명을 표시하고 있다. 또한 AI 인덱스 보고서는 대학교 관계자가 수강신청 데이터에 관해 제공한 추가적인 세부 사항이나 뉘앙스를 포함하고 있다.
- 대학교 관계자가 언급하지 않은 뉘앙스는 포함되어 있지 않다.

인공지능 수업:
Berkeley CS 188
Stanford CS 221
UIUC CS 440
- UIUC에서는 수강신청율 증가의 원인을 더 넓어진 교실 및 학생들의 추가 수요를 충족시키기 위해 추가 섹션을 개강한 것 때문으로 보았다.

UW CSE 415 / 416 (비전공) & CSE 473(전공)
- CSE 416는 2017학년도에 새로 개강한 수업이며 2017년도 수강신청율 증가의 일부 원인이다.

머신러닝 수업:
Berkeley CS 189
- 버클리 대학교 관계자는 수강신청 증가가 참신함(novelty), 과목에 대한 흥미, 머신러닝 개론 수업으로 필수과목을 이수할 수 있게 한 전공의 증가가 복합적으로 작용한 것으로 보았다.

Stanford CS 229
UIUC CS 446
UW CSE 446
CMU 10-701

부록 6
국제적 수강신청 추세

국제적 수강신청 관련 페이지로 돌아가기(23 페이지)
사용된 데이터 보기

출처
수강신청 데이터는 각 대학교에서 직접 수집하였다. 본 분석에는 이하 대학교가 포함되었다:

칭화 대학(중국), 국립천문학연구소, 광전자기술연구소(멕시코), 브리티시 컬럼비아 대학교(캐나다), 토론토 대학교(캐나다), 에든버러 대학교(스코틀랜드), 중국과학기술대학(중국), 상해교통대학교(중국), 칠레 가톨릭 대학교(칠레), 빈 공과대학교(오스트리아), 히브리 대학교(이스라엘), 로잔 연방 공과대학교(EPFL, 스위스), 몬트리올 대학 인공지능 연구소(MILA, 캐나다)

방법론—부록 5 참고.

Nuance

- 거의 대부분의 학교에서 특히 최근 수강신청은 학생들의 수요 보다는 공급에 의한 것이라고 언급하였다. AI 인덱스의 데이터는 성공적으로 수강신청한 학생 수를 나타내며 대기 리스트(waitlist)나 기타 요구 지표(demand metrics)는 고려하지 않았다.
- 미국 소재 대학교와 달리, 해외 소재 대학교의 경우 수업이 학부생 에게만 제공되는지의 여부가 많은 차이를 보였다.
- 페이지 23에 위치한 첫번째 그래프는 인공지능 및 머신러닝 수업 수강신청 합계를 보여준다. 두번째 그래프는 인공지능 수업 수강신청 정보만을 담고 있다. 동일 그래프에 표시된 대학교들의 데이터 성질을 동일하게 하기 위해 그래프를 두개로 나누었다. 일부의 경우에 추가적인 데이터가 존재하는 대학교도 있었지만 모든 학교의 정보를 동일 선에서 비교하기 위해 추가 데이터는 표시하지 않았다. 추가 데이터는 우측 상단 사용된 데이터 보기를 눌러 확인할 수 있다.
- 일부 수업은 수업 명이 바뀌었으며 AI 인덱스의 데이터는 2017년도 기준 수업 명을 표시하고 있다. 또한 AI 인덱스 보고서는 대학교 관계자가 수강신청 데이터에 관해 제공한 추가적인 세부 사항이나 뉘앙스를 포함하고 있다. 대학교 관계자가 언급하지 않은 뉘앙스는 포함되어 있지 않다.

INAOE — 수업: C141 (AI) and C142 (computational learning)
주석: INAOE의 AI / ML수강신청은 INAOE 대학원 프로그램에 합격한 학생들의 전체에 매우 큰 영향을 받는다. INAOE의 담당자가 말하길 감소하는 INAOE 학생수가 AI / ML 수강신청에 영향을 준다고 하였다.
USTC — 수업: USTC는 몇가지의 개론적인 AI / ML수업을 다양한 부서에서 제공을 한다. 부서: Department of Computer Science and Technology, The Department of Automation, the Department of Information Science and Technology and the Department of Data Science.
에든버러 대학교— 수업: Intro applied ML (학부 및 대학원생 대상) and Informatics 2D — Reasoning and Agents(오직 학부)
SJTU — 수업: CS410 (undergraduate intro to AI)
PUC — 수업: Intro to AI
2017년 이전까지 해당 수업은 학기당 한가지 수업 밖에 없었지만 2017년에 2018년 대비 수강신청이 높았던 이유는 수업 개수가 하나에서 둘로 늘었기 때문이다.
칭화 대학교 — 수업: AI (60240052 & 40250182) 및 ML (00240332 & 70240403 & 80245013 & 80250943)
학부 및 대학원생 대상
토론토 대학교 — 수업: AI (CSC384) 및 ML:(CSC411)
2016에 최초로 여름학기에 인공지능 수업이 열렸다. 증가한 수강신청 수요 때문에 2015년에는 두 학기 모두 머신러닝 수업이 열렸다
브리티시 컬럼비아 대학교 — 수업: AI 및 ML (CPSC 322, CPSC 340, CPSC 422)
빈 공과대학 — 수업: Intro to AI (학부) 및 ML (대학원)
두 수업 모두 컴퓨터 과학 전공 필수 과목은 아님
히브리 대학교 — 수업: Intro to AI (67842) 및 Intro to ML (67577)
EPFL — 수업: Artificial Intelligence (학부) 및 Intelligent Agents (대학원)
EPFL에서는 학사일정 조정과 자원 배분의 이유로 매해 열리는 수업이 상이하다고 밝혔다.
MILA — Machine Learning (대학원생 대상, 전공생 대상 수업 아님)
MILA는 대학원생을 대상으로 하는 프로그램이다. MILA에서는 머신러닝 수업은 머신러닝을 연구하는 학생들이 수강하는 것이 아니라 머신러닝을 배우고 싶은 타 전공 학생들이 수강한다고 밝혔다. 머신러닝을 연구하는 학생들은 상위 과정을 수강한다. 해당 수업은 2014년에는 열리지 않았다.

부록 7
교수 다양성

교수 다양성 관련 페이지로 돌아가기(25 페이지)
사용된 데이터 보기

출처
교수의 다양성에 대한 정보는 2018년 9월 21일 기준 인공지능 학과 웹사이트에서 직접 수집하였다. 세계 유수 컴퓨터 과학 대학교 중 쉽게 접근할 수 있는 교수 명단이 있는 학교들의 정보가 포함되어 있다.

방법론
인공지능 교수의 성별 다양성을 수집하기 위해 각 대학교 웹사이트에서 교수진 이름을 수집하였으며 이름과 사진에 성별 분류를 하였다(첫번째 뉘앙스 참고). 각 학교별 세부 정보는 이하 참고:

UC 버클리 – 교수진 정보 보기
조교수, *Primary*, *Secondary* 교수 포함
스탠포드 대학교 – 교수진 정보 보기
교수, 연구자(Research Scientists), 협력교수(*Affiliated Faculty)* 포함
일리노이대학교 어바나-샴페인 캠퍼스 – 교수진 정보 보기
컴퓨터과학 교수, 협력교수*(Affiliate Faculty)* 포함
카네기 멜론 대학교 – 교수진 정보 보기
교수진 페이지에 표시된 모든 교수 포함
유니버시티 칼리지 런던 – 교수진 정보 보기
인물*(People)* 페이지에 표시된 모든 교수 포함
옥스포드 대학교 – 교수진 정보 보기
교수 명단(*Faculty* section)에 포함된 인물만 포함
취리히 공대 – 교수진 정보 보기
박사(“*Dr.*”) 타이틀이 사람들만 포함

뉘앙스

- AI 인덱스 팀에서는 교수 이름과 사진을 이용하여 각 인물에 성별을 지정하였다. 이를 통해 AI 인덱스 팀에서 성별을 분류하는데 있어 오류가 있을 수도 있다. 우리는 이분법적이지 않은(non-binary) 성별 카테고리를 분석에 포함하지 못한 것을 아쉽게 생각한다. 그렇지만 우리는 본 분석이 인공지능 관련 학계의 성별 다양성을 개괄적으로 보여주는데 유용할 것이라고 생각하며 미래 보고서에 추가적인 젠더를 포함시키고자 한다.
- 대학교 데이터는 2018년 9월 21일 기준으로 수집하였다. 2018년 AI 인덱스 보고서가 발표될 즘에는 교수진 정보가 기준일과 달라졌을 수 있다.
- 데이터는 각 학교의 인공지능 교수진 명단을 통해 수집하였으며 방문 교수나 타 학과 소속 교수는 포함하지 않았다. 또한 인공지능 교수진의 일부로 활발히 활동하는 교수의 경우에는 타 학과 소속 교수의 경우에도 포함하였다.
- 학계의 성별 다양성이 산업계의 성별 다양성을 대변하지는 않는다(사실 인공지능 업계에 종사하는 여성의 비율이 더 낮을 수 있다).
- 본 지표는 오늘날 인공지능 학계 성별 다양성에 대한 정보를 보여주며 과거에 비해 다양성이 개선되었는지에 관한 정보는 포함하지 않고 있다. 스탠포드 대학교 인공지능 교수인 Dr. 제임스 A. 랜데이 교수는 이하와 같이 언급하였다:

  “우리는 더 다양한 교수진을 끌어들이려고 하고 있다. 교수진 중 대부분의 여성 교수들은 지난 2, 3년 사이에 채용된 것을 고려하면 교수진의 다양성이 개선되고 있다고 볼 수 있다”
  - 랜데이 박사, 스탠포드 대학교

부록 8
컨퍼런스 참가

컨퍼런스 참가 관련 페이지로 돌아가기(26 페이지)
사용된 데이터 보기

출처
컨퍼런스 참가자 데이터는 컨퍼런스 및 기관 담당자로부터 직접 수집하였다. 분석에 포함된 컨퍼런스는 이하와 같다:
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 workshop

방법론
2018년에 2000명 이상이 참가한 컨퍼런스를 대규모 컨퍼런스로 정의하였으며 2000명 미만이 참가한 컨퍼런스를 소규모 컨퍼런스로 정의하였다. 본 분석에 포함된 컨퍼런스는 인공지능 연구에 있어 선도적인 컨퍼런스이며 연도별 참가자 정보를 제공한 컨퍼런스이다.

AI4ALL과 WiML은 두 기관의 인공지능 분야의 포용성 개선에 대한 노력과 가용 데이터를 이유로 포함되었다. 미래 발간된 보고서에 많이 소수(underrepresented) 집단에 초점을 맞추는 기관/컨퍼런스를 더 포함시키고자 한다.

뉘앙스
컨퍼런스 관련 뉘앙스
- 일부 컨퍼런스 주최 측에서는 대략적 참가자 데이터만 가지고 있었으며 우리는 대략적 통계도 정확한 통계로 간주했다.
- 일부 컨퍼런스는 연례적으로 열리지 않으며 일부 컨퍼런스의 경우에는 개최하지 않은 해도 있었다.
- 일부 컨퍼런스 주최 측에서는 컨퍼런스 장소가 일년 앞서 결정되기 때문에 참가 가능한 인원에 한계가 있는 경우도 있다고 밝혔다. 그러므로 참가자 수가 참가 수요를 정확히 대변하는 것은 아니다.

AI4ALL/WiML과 관련된 뉘앙스
- AI4ALL/WiML 외에도 인공지능의 포용성을 확대하기 위한 여러 공식/비공식 프로그램이 존재한다.
- 참가자 수 증가가 해당 분야에서 여성 및 소수 집단의 참여가 늘었다는 것을 정확히 의미하는 것은 아니다.

부록 9
로봇 소프트웨어 다운로드

로봇 소프트웨어 다운로드 관련 페이지로 돌아가기(29 페이지)
사용된 데이터 보기

출처
ROS.org 페이지 및 소프트웨어 다운로드 데이터는 ROS.org 커뮤니티 지표 보고서를 바탕으로 했다. 연도별 PDF 보고서 보기. ROS.org 에 대한 추가 정보 보기.

뉘앙스
- ROS.org 페이지 뷰는 페이지 뷰 수가 높은 국가의 수치만 보여줄 뿐 전체 페이지 뷰를 보여주지는 않는다. 90개 이상의 하위권 국가들은 포함되어있지 않다.
- 2018년 지표는 ROS.org에서 볼 수 있다. 그러나 2018년 자료는 약간 다른 방법이 사용되었기 때문에 AI 인덱스 보고서에는 포함되지 않았다. 관련 추가 정보 보기.
- 월 평균 페이지 뷰 수치는 매해 7월 데이터를 사용하여 측정하였다(2012년은 6월 데이터 사용, 2013년은 8월 데이터 사용)
- ROS.org 관계자는 데이터가 오직 영문 웹사이트을 바탕으로 하기 때문에 비영어권 국가에서 로봇 소프트웨어 관련 활동은 낮게 평가되었을 수 있다고 밝혔다.
- ROS.org 관계자는 2015년의 다운로드 횟수 하락은 샘플링 및 해당 월에 적은 수의 패치가 출시되었기 때문일 것으로 추정했다. 관련 추가 정보 더 보기.

부록 10
스타트업 / 투자

스타트업 / 투자 관련 페이지로 돌아가기(31 페이지)
가용한 사용된 데이터 없음.

출처
샌드힐 이코노메트릭스(Sand Hill Econometrics)에서 다우존스 벤처소스 데이터베이스를 활용한 스타트업 및 벤처 투자 데이터를 제공하였다. 크런치베이스 API 를 활용하여 샌드힐에서 제공한 데이터베이스 내 인공지능 기업 명단을 만들었다. 샌드힐, 벤처소스, 크런치베이스 관련 추가 정보 보기.

방법론
이하 키워드를 활용하여 크런치베이스 데이터베이스에서 인공지능 관련 기업 명단을 추출했다:

Category Name	Crunchbase Category UUID
인공지능 (Artificial Intelligence)	c4d8caf35fe7359bf9f22d708378e4ee
머신러닝 (Machine Learning)	5ea0cdb7c9a647fc50f8c9b0fac04863
예측분석 (Predictive Analytics)	ca8390d722c65bb5f87022f52f364b1b
지능 시스템 (Intelligent Systems)	186d333a99df4a4a6a0f69bd2c0d0bba
자연어 처리 (Natural Language Processing)	789bbbefc46e1532a68df17da87090ea
컴퓨터 비전 (Computer Vision)	a69e7c2b5a12d999e85ea0da5f05b3d3
얼굴 인식 (Facial Recognition)	0b8c790f03bcb2aba02e4be855952d6d
이미지 인식 (Image Recognition)	af9307c9641372aeaac74391df240dd2
시맨틱 검색 (Semantic Search)	bb1777e525f3b9f2922b0d5defe0d5bb
시맨틱 웹 (Semantic Web)	ac34c4c6e430f66f44aeef8ca45b52bb
음성 인식 (Speech Recognition)	24dad27e6a49ccc1ec31854b73d60d16
언어 분석 (Text Analytics)	e7acbc9932c74c8809ed8fae63eb2559
가상 비서 (Virtual Assistant)	2eb0b56b6896e0f0545fc747a9ba6751
비쥬얼 검색 (Visual Search)	ad8f33a8c5d0b7d786f389aa2954c119

크런치베이스 데이터베이스에는 4,477개의 인공지능 기업 데이터가 포함되어 있었다. 해당 4,477개 기업의 데이터는 벤처소스 데이터베이스에 따른 벤처 투자를 받은 기업 명단과 교차 비교하였다. 크런치베이스의 4,477개 기업 중 1,096개 기업이 벤처소스에 따른 벤처 투자를 받은 인공지능 기업이었다. 샌드힐 이코노메트릭스에서는 해당 1,096개 기업의 숫자와 투자액을 집계하였다.

"모든 스타트업"에 관한 데이터는 오직 벤처소스 데이터베이스에서 수집하였으며 크런치베이스 API는 사용하지 않았다.

뉘앙스
- 벤처소스 데이터는 비공개 자료이며 대중에게 공개되지 못함을 아쉽게 생각한다.
- VentureSource의 데이터 분석은 Sand Hill Econometrics가 담당하였다.
- 공개 상장되었거나, 인수 또는 폐업한 기업의 경우 벤처소스 데이터베이스에서 제외하였다. 그러나 해당 기업이 데이터베이스에서 제외되기 전 투자 자료는 데이터베이스에서 삭제하지 않았다.
- 크런치베이스와 벤처소스 모두 인공지능 시스템을 개발하는 기업, 인공지능 "메이커"와 활용하는 기업, 인공지능 "유저"를 구분하는데 어려움을 겪고 있다. AI 인텍스에서는 109개의 기업을 무작위로 (벤처소스 데이터베이스에 따른 벤처 투자를 받은 기업 중 10%) 조사한 결과 이 중 68개의 기업(또는 62%)가 인공지능 "메이커"였다. 샘플 중 나머지 38%는 인공지능 기업으로 분류되지만 스스로 인공지능을 개발하지는 않았을 수 있다.
- 크런치베이스 데이터베이스에서 인공지능 기업으로 분류된 기업은 현재 인공지능 관련 기술로 기계학습 기술을 선호한다.
- 활발히 운영되고 있는 스타트업의 수는 매년 1월 기준으로 집계하였다.

부록 11
일자리

일자리 관련 페이지로 돌아가기(33 페이지)
사용된 데이터 보기

출처
인공지능 관련 기술을 필요로 하는 신규 일자리에 관한 그래프*(33 페이지)* 와 인공지능 관련 일자리 지원자의 성별 다양성에 관한 그래프*(34 페이지)* 는 구인구직 사이트인 Monster.com에서 제공하였다. Monster.com에서는 온라인 구인시장 정보 포털인 Garner TalentNeuron을 활용하여 데이터를 수집하였다. Garner TalentNeuron 와 Monster.com 에 관한 추가 정보 보기.

방법론
Monster.com
Monster.com에서는 이하 키워드 검색을 위하여 Garner TalentNeuron 데이터베이스를 활용하였다. 요구 기술에 해당 키워드가 포함되었는지를 기준으로 수집하였다. 또한 각각의 키워드는 인공지능이라는 키워드와 함께 검색하였다. 예를 들어, 로봇공학을 요구 기술로 명시한 일자리의 경우 인공지능과 로봇공학을 **둘다** 명시한 경우에 인공지능 관련 신규 일자리로 취급하였다. Monster.com의 일자리는 미국 내 일자리만 제공하고 있다. 아래 키워드 참고:

인공지능 +
자연어 처리, 머신러닝 기술, 딥러닝, 컴퓨터 비전, 음성 인식, 로봇공학

뉘앙스
- 한가지 이상의 인공지능 관련 기술을 요구한 경우 중복으로 계산되었다. 일부 요구되는 기술의 경우 실제로 명시되어 있지 않은 일정 기술을 요구하는 것을 함의하고 있을 수 있다.

인공지능 일자리 지원자의 성별 다양성과 관련한 뉘앙스
- 지원자 수가 실제 고용된 사람의 수나 산업 전체의 상황을 대변하는 것은 아니다.
- 우리는 이분법적이지 않은(non-binary) 성별 카테고리를 분석에 포함하지 못한 것을 아쉽게 생각한다. 그렇지만 우리는 본 분석이 인공지능 관련 학계의 성별 다양성을 개괄적으로 보여주는데 유용할 것이라고 생각하며 미래 보고서에 추가적인 젠더를 포함시키고자 한다.

부록 12
특허

특허 관련 페이지로 돌아가기(35 페이지)
사용된 데이터 보기

출처
특허 관련 데이터는 특허 검색 서비스 amplified에서 제공하였다.

방법론
모든 방법론 및 사용된 데이터에 관한 amplified의 추가 정보 보기.

뉘앙스

- AI 인덱스의 특허 관련 그래프는 2015년에서 2017년까지의 정보는 표시 되어있지 않다. 해당 연도에는 특허 출원에서 등록까지 소요되는 시간이 길어 총 특허 숫자가 제공되지 않기 때문이다.

인공지능 관련 특허를 지속적으로 그리고 지역별로 추적하고자 노력하였으나 기준이 되는 추적 방법이 존재하지 않았다. 결과적으로 AI 인덱스의 특허 데이터는 다른 기관들의 데이터와는 다른 양상을 띈다.

주목할만한 예는 2018년 여름 발간된 칭화대학교의 중국 인공지능 개발 보고서이다. 해당 보고서에 따르면 중국이 가장 많은 특허 등록 건수를 가지고 있었으며 미국과 일본이 그 뒤를 이었다. 그러나 어떠한 특허 코드와 정의를 활용하여 이와 같은 결과를 얻었는지는 불확실하며 지역별 실제 특허 건수는 그래프에 보기 쉽게 표시하기 어렵다.

인공지능 관련 분야가 정의하기 쉽지 않기 때문에 AI 인덱스는 본 분석에 사용된 특허 코드와 키워드를 완전히 투명하게 공개한다. 우리는 독자들께서 우리의 특허 건수 집계 방법에 대한 의견을 제시하여 집계 방법의 정확도가 높아지길 기대한다.

부록 13
인공지능 도입 설문조사

인공지능 도입 설문조사 관련 페이지로 돌아가기(36 페이지)
사용된 데이터 보기

출처
설문조사는 McKinsey & Company(맥킨지 앤 컴퍼니)에 의해 작성, 수집 및 분석되었다. 맥킨지의 인공지능 도입 관련 설문조사에 관한 추가 결과 보기.

방법론
맥킨지에서는 설문조사 방법에 관하여 아래와 같이 밝혔다:
본 설문조사는 온라인 및 오프라인에서 2018년 2월 6일부터 2018년 2월 16일까지 이루어졌다. 총 2,145명이 응답하였으며(설문조사를 요청한 20,466명의 사람들 중) 응답자는 맥킨지의 온라인 임원 패널 내 다양한 지역별, 산업별, 기업 규모별, 특화된 분야 및 임기를 모두 반영하고 있다. 응답자 중 1,657명은 자신들이 속한 기업이 설문조사에 포함된 9가지 분야 중 한가지 이상의 분야 또는 사업부문에서 인공지능 기술을 시범 도입하고 있다고 밝혔다. 모든 응답자는 McKinsey.com에 등록 사용자 30,000명 이상으로 이루어진 온라인 패널의 회원이었다. 온라인 패널은 맥킨지의 자체적인 연구에 참여하기로 한 사람들이며 지위를 막론한 기업 임원, 관리자 및 직원을 대표하는 세계적인 표본이라 할 수 있다. 설문조사 대상에는 108개 국가가 포함되어 있었으며 국가별 응답률 차이를 보완하기 위하여 각 국가의 GDP가 세계 GDP에서 차지하는 비율에 따라 데이터에 가중치를 두었다.

맥킨지에서는 이하 질문에 답한 응답자 비율을 제공하였다:

Q1. 소속 기관의 현재 인공지능 기술 활용 정도는 어떻습니까(로봇 프로세스 자동화, 머신러닝, 가상 에이전트 또는 대화 인터페이스, 컴퓨터 비전, 자연어 텍스트 이해, 자연어 음성 이해, 자연어 생성, 물리적 로봇공학, 무인자동차)? 1가지만 고르시오:
1. 전혀 사용하지 않고 있다
2. 한가지 분야 도는 사업부문에서 시범 도입되었다.
3. 최소 한가지 분야 또는 사업부문의 표준 업무 프로세스에 도입되었다.
4. 다양한 분야 또는 사업부문의 표준 업무 프로세스로 도입되었다.
5. 잘 모르겠다.

Q2. 소속 기관에서 어떠한 사업 분야에 인공지능을 도입하였나요?[임의로 순서를 정함] (해당되는 모든 선택지를 고르시오)
1. 마케팅 및 판매
2. 전략 및 기업 금융(예를 들어 capital-allocation decisions)
3. 리스크(예를 들어 리스크 관리, 사기 및 부채)
4. 인적 자원
5. 제품 및/또는 서비스 개발
6. 공급망 관리
7. 제조
8. 서비스 운영(예를 들어, 오프라인 서비스, 고객 관리, 비영업 부서(back office))
9. 해당되는 것 없음
10. 잘 모르겠다.

첫번째 질문의 경우 우리는 맥킨지에게 응답 결과에 대한 지역별 분류를 요청하였다. AI 인덱스에서는 세번째와 네번째 선택지을 선택한 응답자 수를 합하여 최소 한가지 이상의 분야 또는 사업부문에 인공지능 기술을 도입하였다고 응답한 응답자의 비율을 계산하였다. 두번째 질문의 경우 맥킨지는 지역별 산업별 데이터를 제공하였다. 첫번째 질문과 달리 두번째 질문의 경우 인공지능 기술을 도입 또는 시범 도입하였다고 응답한 응답자를 대상으로 하였다. 관련 응답자 수는 해당 그래프 하단에 명시되어 있다.

뉘앙스
- 설문조사 응답자는 각자의 인공지능 도입에 관한 소속 기관의 상황에 대한 인식에 기반하고 있다는 제약이 있다.

부록 14
기업 이니셔티브: 기업 실적 발표

실적 발표 관련 페이지로 돌아가기(40 페이지)
사용된 데이터 보기

출처
실적 발표에 관한 모든 데이터는 텍스트 분석 및 심리 분석에 특화된 투자연구기업인 Prattle이 제공하였다. Prattle이 사용한 실적 발표 전문은 펙스셋(FactSet)의 데이터를 활용하였다.

방법론
Prattle은 실적 발표에서 인공지능 및 머신러닝이 언급된 횟수를 제공하였으며 추가 비교 및 벤치마킹을 위하여 클라우드와 빅데이터가 언급된 횟수 또한 제공하였다. 기업별 데이터가 제공되었으며 각 기업이 속한 산업 및 기업 규모에 관한 정보도 포함하고 있었다.

실적 발표 전문 에서 인공지능에 관해 언급하는 문자(strings) 및 언어 패턴을 검색하기 위한 텍스트로 Prattle은 다양한 정규표현식(regular expressions, regex라고도 불림)을 사용하였다. 인공지능 및 머신러닝의 축약형인 AI 및 ML로 언급된 것 또한 포함되었다. Prattle은 거짓 긍정(false positive)(예를 들어 Prattle에서는 ML 앞에 숫자가 붙은 경우는 수집하지 않음으로서 밀리미터를 ML로 표시한 것)을 제거하였다.

산업 분류는 S&P에서 사용하는 표준 분류법인 GICS를 사용하였다. 기업 규모 분류는 시가총액을 기준으로 하였으며 Prattle은 발행주식과 주가를 이용하여 시가총액을 계산하였다. Prattle에서 제공한 임곗값(thresholds)은 이하와 같다:

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

모든 데이터 분석은 AI 인덱스에 의해 이루어졌다.

뉘앙스
- 빅데이터와 클라우드에 관한 언급 횟수는 과거에에 높았다가 최근 줄어든 추세를 보였다. 우리는 이러한 추세가 기술의 정규화 때문이라고 해석했다.
- 언급 횟수가 기업의 인공지능에 대한 투자 또는 도입을 간접적으로 말해주는 것은 아니다.
- 공개 상장된 모든 기업이 분석에 포함되었다.

부록 15
로봇 설치

로봇 설치 관련 페이지로 돌아가기(41 페이지)
사용된 데이터 보기

출처
데이터는 국제 로봇학 연맹(International Federation of Robotics, IFR)의 2014, 2015, 2017년 세계 로봇학 리포트(World Robotics Reports)에서 직접 수집하였다. 해당 보고서 링크는 이하 참고. IFR에 관한 추가 정보 보기

2012년 데이터
출처: 2014 World Robotics Report – 전체 개요(Table 1)
2013–14 년 데이터
출처: 2015 World Robotics Report – 전체 개요(Table 1)
2015–16 년 데이터
출처: 2017 World Robotics Report – 전체 개요(Table 4.1)

방법론
제시된 데이터는 국가별 산업용 로봇 설치 대수를 나타낸다. 산업용 로봇은 ISO 8373:2012 기준에 의해 정의된 로봇을 의미한다.

IFR의 방법론 보기.

뉘앙스
- 인공지능으로 분류되는 소프트웨어를 사용하는 로봇의 비율을 어떻게 계산할지는 명확하지 않으며 인공지능이 산업용 로봇 사용량에 얼마나 영향을 미치는 지는 확실하지 않다.
- 2017년 AI 인덱스 보고서에서는 본 지표가 로봇 수입량으로 불렸다.

부록 16

깃허브 별 개수(GitHub Stars)

깃허브 관련 페이지로 돌아가기(42 페이지)
사용된 데이터 보기

출처

우리는 Google BigQuery에 저장된 GitHub archive를 사용하였다.

방법론

그래프는 다양한 Github repository가 받은 별의 수를 보여준다. 관련 Repository는 이하를 포함한다:

apache/incubator-mxnet, BVLC/cafe, cafe2/cafe2, dmlc/mxnet, fchollet/keras, Microsoft/CNTK, pytorch/pytorch, scikit-learn/scikit-learn, tensorflow/tensorflow, Theano/Theano, Torch/Torch7

GitHub archive 데이터는 Google BigQuery에 저장되어있다. AI 인덱스 팀은 구글 BigQuery에서 대상 repository의 'WatchEvents' 수를 계산하였다. 이와 관련하여 2016년 데이터를 수집하는데 사용한 코드 샘플은 아래와 같다:

```
 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
```

다음 페이지에서 관련 뉘앙스 보기

부록 16
깃허브 별 개수 (계속)

뉘앙스
현재 GitHub Archive는 repository에서 사용자가 별을 제거할 경우 이를 반영할 방법은 제공하지 않고 있다. 그러므로 계산된 별의 개수는 약간 과장되었다고도 할 수 있다. 계산된 별의 개수와 실제 GitHub repository의 별 개수와 비교했을 때 숫자는 비슷하였으며 경향에도 변함이 없었다.

GitHub 별 개수를 세는 다른 방법도 존재한다. star-history tool을 이용하여 데이터를 무작위로 검사하였다.

GitHub 프로젝트의 Fork 또한 흥미로운 조사 대상이다. 팀에서는 GitHub 저장소의 별 개수와 Fork가 거의 동일한 경향을 보이는 것을 발견하였다.

부록 17
언론 동향

언론 동향 관련 페이지로 돌아가기(43 페이지)
사용된 데이터 보기

출처
모든 데이터는 TrendKite에서 수집하였다. TrendKite는 언론 기사를 색인하고 기사를 '긍정적', '부정적', '중립적' 등으로 분류하여 동향을 분석한다. 인덱스 팀은 긍정적 또는 부정적으로 분류된 기사의 비율을 표시했다. TrendKite에 대한 정보 더 보기.

방법론
AI 인덱스 팀에서는 이하 질의를 사용하여 인공지능 관련 기사를 판별했다. 부정적인 입장을 가진 관련 없는 기사가 불균형적으로(disproportionate) 많은 매체는 제외하도록 질의를 조정했다. 2017년 AI 인덱스 보고서에서는 중립적 기사의 비율의 경우 분석에는 포함하였으나 그래프에는 표시하지 않았으나 올해 보고서에는 중립적 기사의 비율을 그래프에 포함하였다.

질의
"Artificial Intelligence"
AND NOT "MarketIntelligenceCenter.com's"

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

필터
- 오직 영어로 된 기사만 포함
- 보도자료 제외
- 금융 기사 제외
- 부고 제외

부록 18
정부 언급

정부 언급 관련 페이지로 돌아가기(44 페이지)
사용된 데이터 보기

국가별 출처, 방법론 및 뉘앙스:
데이터 수집 및 분석은 맥킨지 글로벌 연구소(McKinsey Global Institute (MGI))에 의해 이루어졌다.

캐나다(하원):
데이터는 캐나다 의회 웹사이트의 한사드 서치(Hansard search)를 활용하여 수집하였다. MGI에서는 "인공지능" 및 "머신러닝"(따옴표 포함)을 키워드로 검색하였으며 결과는 CSV 형식으로 다운받았다. 데이터의 범위는 "모든 토의(All debates)"로 설정하였다. 데이터는 2018년 11월 20일 기준이며 해당 데이터는 2002년 8월 31일부터 온라인을 통해 공개되었다.

각 언급 횟수는 하원의원 프로시딩(proceedings of the House of Commons)에 언급된 논평 또는 발언에서 인공지능 또는 머신러닝이 언급된 횟수를 나타낸다. 즉, 특정 행사 또는 대화에서 의원이 인공지능이나 머신러닝을 여러 차례 언급한 경우는 한번으로 계산하였다. 그러나, 만약 동일한 행사에서 화자가 인공지능이나 머신러닝을 별도 발언에서 언급하였다면(발언과 발언 사이에 다른 화자의 발언이 있는 경우) 각각 별도로 계산하였다. 인공지능 및 머신러닝은 별도 검색을 통해 따로 계산하였다. AI 또는 ML로 언급된 경우는 포함하지 않았다.

영국(하원, 상원, 국회의사당 및 위원회)
데이터는 영국 의회의 한사드 서치(Hansard search)의 레퍼런스 찾기(Find References) 기능을 활용하여 수집하였다. MGI에서는 "인공지능" 및 "머신러닝"(따옴표 포함)을 키워드로 검색하였으며 검색 결과를 분류하였다. 데이터는 2018년 11월 20일 기준이며 해당 데이터는 1800년 1월 1일 데이터부터 온라인을 통해 공개되었다. Open Parliament Licence v3.0에 따라 허가된 의회 정보를 포함하고 있다.

캐나다의 경우처럼 언급된 논평 또는 발언에서 인공지능 또는 머신러닝이 언급된 횟수를 나타낸다. 즉, 특정 행사 또는 대화에서 의원이 인공지능이나 머신러닝을 여러 차례 언급한 경우는 한번으로 계산하였다. 그러나, 만약 동일한 행사에서 연사사가 인공지능이나 머신러닝을 별도 발언에서 언급하였다면(발언과 발언 사이에 다른 화자의 발언이 있는 경우) 각각 별도로 계산하였다. 인공지능 및 머신러닝 언급 횟수는 별도 검색을 통해 따로 계산하였다. AI 또는 ML로 언급된 경우는 포함하지 않았다.

미국 (상원 및 하원)
데이터는 미국 의회 기록 사이트 의 고급 검색 기능을 이용하여 수집하였다. MGI에서는 "인공지능" 및 "머신러닝"(따옴표 포함)을 키워드로 검색하였으며 결과는 CSV 형식으로 다운받았다. "단어 변형(word variant)" 옵션은 선택하지 않았으며 상원 섹션, 하원 섹션 및 리마크 섹션(Extensions of Remarks)에서 언급된 것을 포함하였지만 데일리 다이제스트(Daily Digest)는 포함하지 않았다. 데이터는 2018년 11월 20일 기준이며 해당 데이터는 104대 의회 데이터부터 온라인을 통해 공개되어있다.

각 언급 횟수는 의회 기록에 포함되어 있는 특정 행사에서 인공지능 또는 머신러닝이 언급휜 횟수이며 법안을 읽는 경우도 포함한다. 만약 연사가 발언에서 인공지능 또는 머신러닝을 여러 차례 언급했거나 여러 연사가 동일 행사에서 인공지능 또는 머신러닝을 언급한 경우 모두 1회로 처리하였다. 인공지능 및 머신러닝 언급 횟수는 별도 검색을 통해 따로 계산하였다. AI 또는 ML로 언급된 경우는 포함하지 않았다.

부록 19
ImageNet

ImageNet 관련 페이지로 돌아가기(47 페이지)
사용된 데이터 보기: (1) 정확도 점수 (2) 학습 시간

ImageNet 정확도 점수

출처
이미지넷 정확도 데이터는 2010년부터 2017년까지 열린 LSVRC 대회 리더보드에서 수집하였다. 2016년부터 2018년까지의 검증 데이터셋의 성능은 arXiv 문헌 검토를 통해 수집하였다.

발전 상황을 추적하기 위하여 이하 논문에 포함된 점수를 참고하였다:
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)

인간 수준의 능력 추정치는 Russakovsky et al, 2015, 2015를 기반으로 한다. LSVRC ImageNet 경진대회 및 ImageNet 데이터셋에 관한 추가 정보 보기.

방법론
이미지넷 대회 관련 데이터는 ImageNet 웹사이트에서 열린 각 LSVRC 대회 리더보드에서 수집하였다. 검증 데이터셋의 성능은 arXiv 문헌 검토를 통해 수집하였다. 매년 가장 높은 점수만 수집하였다.

ImageNet 학습 시간

출처
이미지넷 학습 시간은 arXiv 문헌 검토를 통해 수집하였다. 이하 보고서를 활용하였다:

2017년 6월 8일: 60분
2017년 11월 1일: 30분
2017년 11월 12일: 15분
2018년 7월 30일: 6.6분
2018년 11월 13일: 3.7분

부록 20
파싱 (Parsing)

파싱 관련 페이지로 돌아가기(50 페이지)
사용된 데이터 보기

출처
Penn Treebank에서 월스트리트 저널의 문장으로 구성된 부분은 각 문장을 문장 구성성분 분석트리(constituency-based parse tree) 방식으로 주석을 단 데이터 셋이다. 해당 데이터의 섹션 23은 자동 파서(automatic parsers) 연구를 위한 주요 테스트 셋으로 이용하였다.

방법론
자동 파서(automatic parsers)는 자동으로 생성된 파스(parse)의 문장 구성 성분과 테스트 셋의 GOLD parse 문장 구성 성분을 비교하여 평가한다. 생성된 구성 성분의 정확도(precision)와 재현율(recall)이 반영돼 F1 점수(F1 score)가 산출된다. 보고서에 담긴 내용은 Penn Treebank의 월스트리트 저널 부분의 섹션 23에 있는 문장 파서(parser)의 F1 점수이다. 보고서에서는 F1 점수가 산정 가능한 문장들의 점수와 40글자 미만 문장의 점수를 다루었다. 문장 구성성분 분석트리에 관한 위키피디아 페이지를 통해 더 알아보기.

우리는 영문 문장 구성성분 분석을 위하여 nlppregress.com에 보관된 리더보드 데이터를 사용하였다. 최신 결과 점수 및 검색은 추가적인 문헌 검토를 통해 이루어졌다. 새로운 데이터는 2017년 AI 인덱스 보고서 데이터에 더해졌다. 보고서에 포함된 파서의 F1점수는 단일 모델 뿐만 아니라 앙상블 기법을 포함한다.

뉘앙스
- 언제 새로운 파서가 자동 파싱의 최신 성능을 향상시켰는지 판별하기 위하여 문헌 검토를 하였다. 인덱스 팀은 1995년까지 거슬러 올라가서 발견한 파서의 F1 점수를 수집하였다. 하나의 모델이 아닌 앙상블 기법이 고려되었다.

부록 21
기계 번역(BLEU)

기계 번역 관련 페이지로 돌아가기(51 페이지)
사용된 데이터 보기

출처
데이터는 연례 기계 번역 워크샵(WMT) BLEU 점수 메트릭스를 제공하는 EuroMatrix를 통해 얻었다.

방법론
매년 WMT는 뉴스 번역 작업을 제시하고 새로운 학습 및 테스트 데이터 셋을 제공한다. 참가 팀은 기계번역 작업을 위해 만든 시스템을 제출한다.

WMT에서 사용하는 주요 평가 지표는 경쟁 엔트리의 순위를 매기는 것을 목표로 하며 전년도 결과와 비교하지 않는다. 또한 매우 노동 집약적이다. 인덱스 팀은 BLEU를 사용하였다, BLEU는 인간의 번역과 기계 번역의 결과의 개략적인 비교를 자동으로 수행하는 방법이다. BLEU는 변형된 버전의 정확도(precision)로 0에서 1 사이의 점수를 주며 높을수록 더 정확하다는 것을 의미한다. 기계 번역 시스템이 번역한 번역된 코퍼스(corpus)의 BLEU 평균 점수 또한 계산해 볼 수 있다. 영어-독일어, 독일어-영어 뉴스 번역 작업을 위해 제출한 시스템이 얻은 해당 연도 평균 BLEU 점수 중 최고점을 기록하였다.

EuroMatrix는 2006년부터 제출된 영어-독일어, 독일어-영어 뉴스 번역 작업(task)의 BLEU 점수를 기록해왔다. 인덱스 팀은 각 연도마다 최고 성적을 거둔 BLEU 시스템을 선별하였다. 특히 문장을 토큰화하는(tokenizing) 프로토콜을 정의한 BLEU(11b)를 사용하는 시스템에 더 비중을 두었다. 가능하다면 영어-독일어, 독일어-영어 두 가지 경우 모두에서 높은 BLEU 점수를 얻은 시스템을 그 해 대표 수치로 선택하였다.

뉘앙스
- BLEU는 자동으로 계산될 수 있으며 번역 품질과 인간의 판단의 상관관계를 보여줄 수 있다 여겨졌다. 그러나 BLEU 지표는 코퍼스(corpus) 간 비교에 사용될 수 없으며 시스템 간 BLEU 점수를 비교하는 것도 오해를 야기할 수 있다.
- 그래프는 일반적으로 우상향의 형태를 띄지만, 2016년 점수와 비교하여 2017년 BLEU 점수가 크게 떨어진 것을 통해 해당 지표가 오류가 있음을 알 수 있다(2017년 점수가 여전히 2015년 점수에 비해서는 높다). 2016년과 비교하여 기계 번역의 성능이 저하 되었다가 보다 평가 시스템이 완벽하지 않을 가능성이 있다. 그러나 장기간에 걸친 경향을 보면 BLEU 점수를 통해 기계 번역 분야의 발전이 있었음을 알 수 있다.

부록 22
ARC

ARC 관련 페이지로 돌아가기(52 페이지)
사용된 데이터 보기

출처
AI2 추론 경진 대회(ARC)는 앨런 인공지능 연구소가 주최한다. ARC 데이터는 ARC 리더보드를 통해 수집하였다. 쉬운 셋(easy set) 정보 및 경진 대회 셋(challenge set) 정보 보기.

방법론
참가자들은 앨런 연구소(Allen Institute) 웹사이트를 통해 ARC 데이터셋을 다운받고 리더보드에 제출하였다.

난이도 낮은 발전 코퍼스(Easy development corpus) 질문 예시:

> 가장 최근에 개발된 기술은 무엇인가요? (A) 휴대폰 (B) 텔레비전 (C)냉장고 (D) 비행기 [4학년]
>
> 학생이 조류(algae)가 생산자(producers)라고 가정하였다. 그렇다면 이에 대한 답을 얻기 위해 할 수 있는 가장 좋은 질문은 무엇인가? (A) 조류가 다른 개체를 먹나요? (B) 어떤 개체가 조류를 먹나요? (C) 조류가 먹이를 만들기 위해 햇빛을 사용하나요? (D) 생태계가 조류 없이 살아남을 수 있나요? [8학년]

난이도 높은 발전 코퍼스(Challenge development corpus) 질문 예시:

> 후안과 라케이사가 경사길로 몇 개의 사물을 굴러 떨어뜨렸다. 이 둘은 어떤 사물이 가장 빠르게 구르는지 보고자 한다. 반복하여 실험하기 위하여 무엇을 해야 하는가 (A) 사물을 그룹으로 묶는다. (B) 경사 각도를 변경한다. (C) 다른 사물을 고른다. (D) 실험의 상세 결과를 기록한다. [4학년]
>
> 고압 시스템이 공기가 물이 응결할 수 있는 차가운 곳으로 상승하는 것을 막는다. 고압 시스템이 장기간 한곳에 머무를 경우 생길 수 있는 결과는? (A) 안개 (B) 비 (C) 가뭄 (D) 토네이도 [8학년]

각각의 질문은 1점짜리 질문이었다. 모델들은 여러 선택지를 고를 수 있었다. N개의 답변을 한 모델의 경우 1/N 점을 받았으며 답을 하지 않은 경우에는 0점을 받았다. 총점은 개별 점수의 합의 평균이다.

AI 인덱스는 이전 제출된 답안의 점수보다 높은 점수만을 수집하였다. 만약 제출된 답안의 점수가 이전에 제출된 답안의 점수보다 낮은 경우 그래프에 표시하지 않았다.
규칙 및 제출 가이드라인에 관한 추가 정보 보기.

부록 23
GLUE

GLUE 관련 페이지로 돌아가기(53 페이지)
사용된 데이터 보기

출처
GLUE 벤치마크 데이터는 GLUE 리더보드를 통해 수집하였다. GLUE 벤치마크 데이터에 관한 추가 정보 보기.

방법론
참가자들은 GLUE 웹사이트를 통해 GLUE 작업을 다운받고 리더보드에 제출하였다. 점수는 각 작업별 평가지표를 바탕으로 계산되었다. 모든 평가지표는 100배로 확대 되었다(예를 들어 퍼센트로). 이후 각 점수의 평균 점수를 최종 점수로 하였다. 복수의 평가지표가 있는(MNLI 포함) 작업의 경우에는 평가지표 별 점수의 평균을 최종 점수로 하였다.

리더보드에서 참가자가 제출한 결과 중 최상의 점수에 대해서만 순위가 매겨지고 리더보드에 표시된다. 각 참가자가 제출한 결과 중 최상의 점수를 받은 결과를 제외한 다른 결과의 점수는 각 참가자의 페이지를 확장하여 확인할 수 있다. 경쟁하는 참가자들의 경우 결과를 비공개로 제출하는 경우도 있으며 이럴 경우 해당 결과는 리더보드에 표시되지 않는다. AI 인덱스의 그래프는 비공개로 제출된 결과는 포함하지 않는다. MNLI에 해당하거나 해당하지 않은 경우에는 점수 산정의 편의상 한가지 작업으로 보았다.

AI 인덱스에서는 이전 결과의 점수보다 높은 점수를 받은 결과만을 수집하였다. 만약 제출된 결과물이 이전에 제출된 모든 결과보다 낮은 점수를 받았다면 해당 결과는 AI 인덱스 보고서에 포함되지 않았다.
규칙 및 제출 가이드라인에 관한 추가 정보 보기.