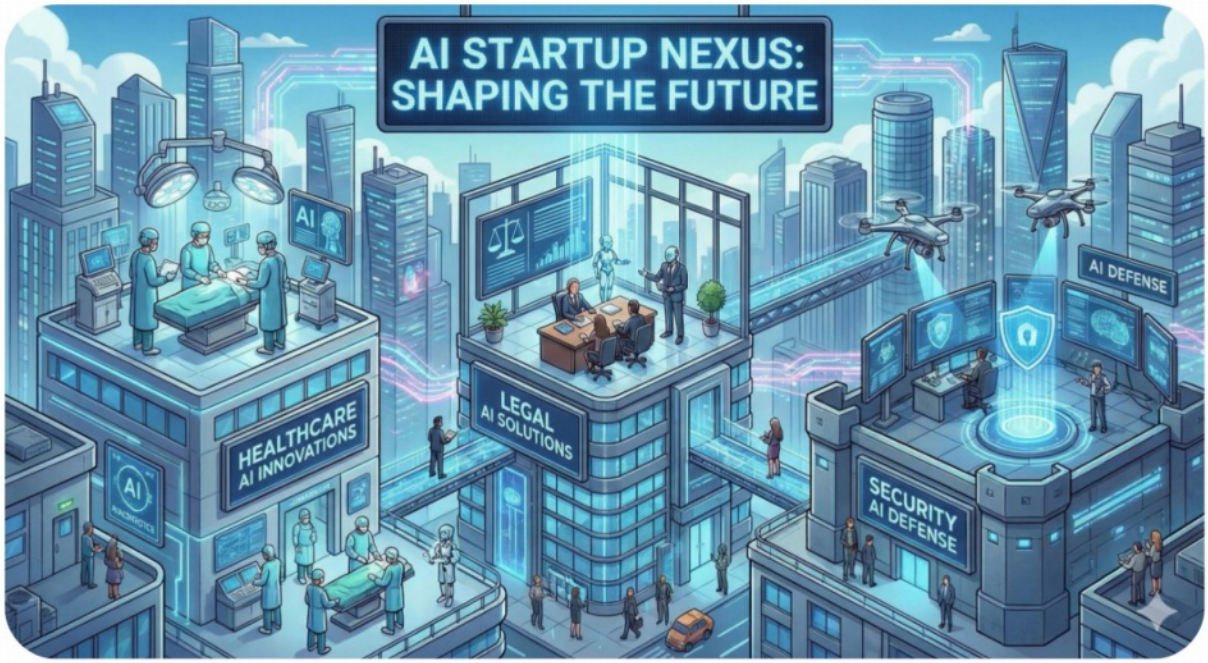
2026-01-20

주목받는 미국 AI 스타트업, 전문성과 기술의 접목

저자 : 안은하

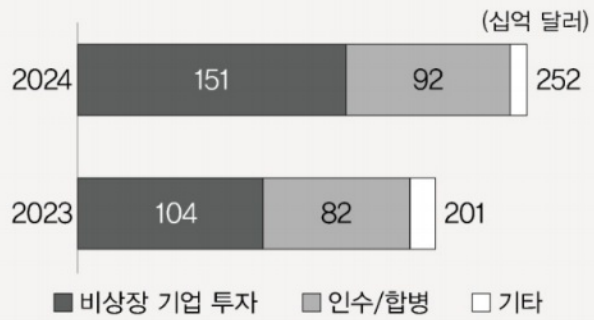


도메인 전문성과 탄탄한 기술력을 갖춘 AI 스타트업들이 주목을 받고 있다. (Gemini로 이미지 생성)

- 어브리지(Abridge): 데이터, 특화 모델, 시스템의 유기적 연결
- 하비(Harvey): 데이터 기반 모델 평가·조율
- 사이에라(Cyera): AI를 통한 데이터 맥락 파악
- 문제 정의로 시작해 도메인 전문성과 독자 기술로 완성

2022년 말 ChatGPT가 공개되고 만 3년이 지났다. ChatGPT가 불을 지핀 AI에 대한 관심은 해를 거듭할수록 더욱 확장되는 모습이다. 2024년 AI 기업에 대한 투자는 전년 대비 약 25% 증가한 2,520억 달러를 기록했다.[1] 그 중에서도 비상장 AI 기업에 대한 투자가 돋보였는데, 투자 총액이 전년 대비 45% 증가한 1,510억 달러에 이르렀으며, 건당 평균 투자액은 약 44% 증가한 4,540만 달러를 기록했다.[2] 투자 건수가 늘어나기 보다, 건당 투자 규모가 커지면서 투자 총액이 증가한 셈이다.

글로벌 Private Investor의 AI 기업 투자 현황



국가별로 살펴보면, 비상장 기업에 대한 투자 대부분이 미국에서 발생했음을 확인할 수 있다. 2024년 미국에서는 약 1,090억 달러가 비상장 AI 기업에 투자되며, 2위인 중국(90억 달러)과 3위인 영국(44억 달러)을 크게 앞섰다.[3] 미국의 이러한 기조는 2025년에도 꾸준히 유지되어 온 것으로 보인다. 라운드 당 1억 달러 이상의 높은 투자금을 유치한 미국 스타트업이 2024년 73개에서 2025년 123개로 약 두 배가량 증가했기 때문이다.[4] 생성형 AI의 원년이라 할 수 있는 2022년부터 지난 4년간 수차례에 걸쳐 대규모 투자를 유치한 미국 AI 스타트업은 총 열한 개였으며, 그 중 작년에만 두 번 이상 대규모 투자 유치에 성공한 곳은 네 군데였다.[5] 이에 설립된 지 3년 미만의 스타트업을 제외하고, 특정 산업 영역에서 두각을 나타내고 있는 세 개의 스타트업을 분석함으로써 이들이 왜 투자자의 선택을 받았는지 핵심 차별점을 찾아보았다.

어브리지(Abridge): 데이터, 특화 모델, 시스템의 유기적 연결

헬스케어는 AI를 접목한 다양한 산업 중 대규모 투자의 증가율이 두드러지는 분야다. 1억 달러 이상 투자 유치에 성공한 스타트업은 2022년 다섯 개에서 2025년 열한 개로 증가했으며, 총 투자금 또한 8억 달러에서 27억 달러로 늘어났다. 어브리지(Abridge)는 의료·헬스케어 분야에서 유일하게 세 번 이상 대규모 투자 유치에 성공한 스타트업으로, 임상 문서 자동화 서비스로 시작해 현재 헬스케어 산업 내에서 점진적으로 사업을 확장 중이다.

어브리지의 공동 창업자이자 CEO 시브 라오(Shiv Rao)는 의사 시절, 하루 평균 2시간 이상을 문서 작업에 할애하며 업무 스트레스가 가중되는 상황에 문제의식을 가졌다. 실제로 미국 의료 시스템은 매년 1조 5,000억 달러의 행정 비용 대부분을 임상 문서 작성에 투입하고 있기도 했다. 인적으로나 물적으로나, 분명 낭비가 심한 부분이었지만 문제를 해결할 수 있는 명확한 서비스는 부재한 상황이었다. 의사와 환자 간의 대화를 텍스트로 옮겨주는 일반적인 자동 음성 인식(Automated speech recognition, 이하 ASR) 기술로는 역부족이었다.

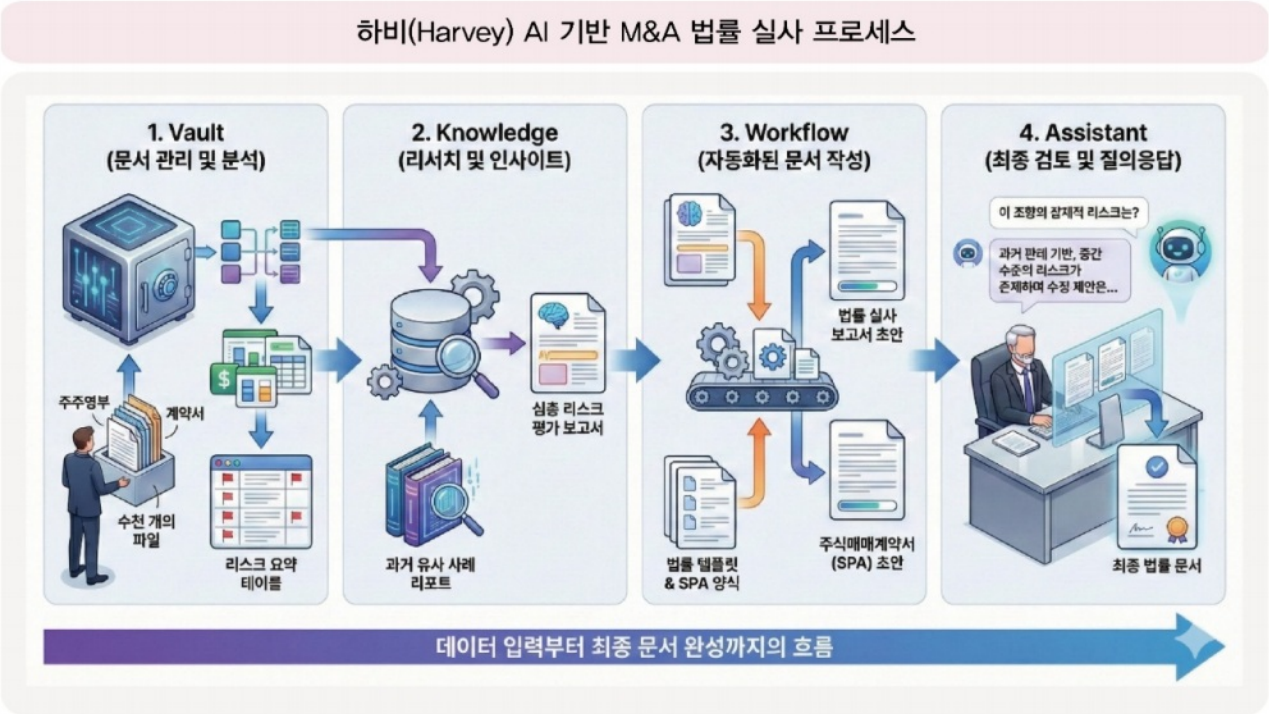
이에 어브리지는 자체 특화 모델을 만들기로 결정했다. 빅테크 모델을 사용하더라도 평균 이상의 결과물 정도는 얻을 수 있었으나, 신뢰도가 중요한 의료계에서 '평균 이상'은 만족할 만한 수준이 될 수 없었다. 어브리지는 150만 건의 비식별 데이터를 기반으로 의료 특화 ASR 모델을 구축해 일반 모델 대비 단어 오류를 24%, 신약명 오류를 83% 감소시키는 한편, 병원에서 널리 사용하는 표준 문서 양식[6]을 학습시킨 임상 문서 작성 모델을 구축함으로써 의사들이 결과물을 추가 수정해야 하는 상황을 최소화했다. 두 모델을 포괄하는 어브리지의 AI 시스템은 이어스(Ears)라고 불리는데, 2025년에는 해당 시스템에 과거 진료 기록 통합, 질환 심각도 코드[7] 매칭, 보험사 청구 문서 작성 등의 기능을 갖춘 상황 추론 엔진을 추가하여 병원 사업 관리(Revenue cycle management) 영역에도 진출하고 있다.

의료 영역을 깊게 파고든 결과, 어브리지는 스타트업 중 유일하게 미국 전자 의료 기록 시스템 점유율 1위인 EPIC의 주요 협력사로 지정되어 150개 이상의 의료 기관에 서비스 중이며, 의료진들이 문서 작성에 투입하는 시간을 60% 줄이고, 번아웃 비율도 40% 감소시키는 성과를 거두었다. 현재 어브리지의 연간 매출은 1억 달러 이상으로 추정된다.

하비(Harvey): 데이터 기반 모델 평가·조율

앞서 어브리지가 의료진과 병원의 고민을 해결해줬다면, 하비(Harvey)는 변호사와 로펌의 고민을 해결하는데 주안점을 뒀다. 변호사 출신 윈스턴 와인버그(Winston Weinberg)는 로펌 1년차에 반복되는 리서치에 시달리며, 적어도 5년은 지나야 반복 업무에서 벗어나 본인이 꿈꿔왔던 전략적 업무를 할 수 있는 상황에 답답함을 느꼈다. 게다가 주니어 변호사가 작성한 문서를 상사가 평가하는데 다시 상당한 시간이 들어가는 구조였는데, 이는 미국 변호사의 시간당 평균 임금이 352달러임을 고려한다면 굉장히 비효율적이었다. 그는 AI를 통해 초반부터 고품질의 문서를 생성할 수 있다면 법률계의 비효율성이 어느정도 해소될 수 있을 것이라 내다봤다.

법률은 판례를 비롯해 많은 데이터가 개방되어 있다. 이에 하비는 직접 파운데이션 모델을 구축하기 보다, 오픈AI와의 파트너십을 통해 GPT-4를 베이스로 약 100억 토큰 규모의 미국 판례법을 학습시켜 법률 문서에 최적화된 파인튜닝 모델을 개발했다. 하지만 법학 서적과 판례를 모두 암기한다고 해서 훌륭한 변호사가 될 수 있는 것은 아니기에, 도메인 전문성을 가진 변호사들을 적극 활용해 차별성을 갖추기 시작했다. 400여 명의 임직원 중 약 20%가 변호사 출신으로, 이들은 '응용 법률 연구(Advanced Legal Researchers)' 팀에서 실제 사건에 적용되는 법률 판단과 사고 구조를 설계하는 역할을 담당하고 있다. 그 결과 하비는 파인튜닝 모델에서 한단계 더 나아가 법률 사고 체계를 갖춘 맞춤학습 모델[8]을 구축하는 한편, 다양한 빅테크 모델의 법률 관련 역량을 평가할 수 있는 빅로우 벤치(BigLaw Bench)를 개발할 수 있었다.



① 매수자 측 주관 법률대리인이 주주명부 등 수천여 개의 파일을 'Vault'에 업로드하면, 구조화된 실사 데이터, 계약별 리스크 요약 테이블 등이 산출 ② 해당 로펌은 자사의 'Knowledge'에서 과거 유사 인수합병의 실사 리포트 등을 참고하여 리스크를 평가 ③ 'Workflow'에 등록된 템플릿으로 법률 실사 보고서, 주식매매계약서 등을 작성 ④ 'Assistant'에서 시니어 혹은 파트너 변호사가 다양한 질문을 던지며 최종적으로 문서를 마무리

모델 관점에서 살펴보면, 하비는 사용자가 명령한 작업의 종류에 따라 일반 모델, 파인 튜닝 모델, 맞춤학습 모델 중 가장 적합한 모델을 배분하는 체계를 갖추고 있다. 멀티 모델 전략 기반의 법률 서비스를 통해 하비는 현재 63개국 700개 이상의 고객사에 서비스를 제공하며, 변호사들의 월 평균 업무 시간을 13~25시간 단축시키고, 연간 매출 1억 달러 이상을 창출하고 있다.

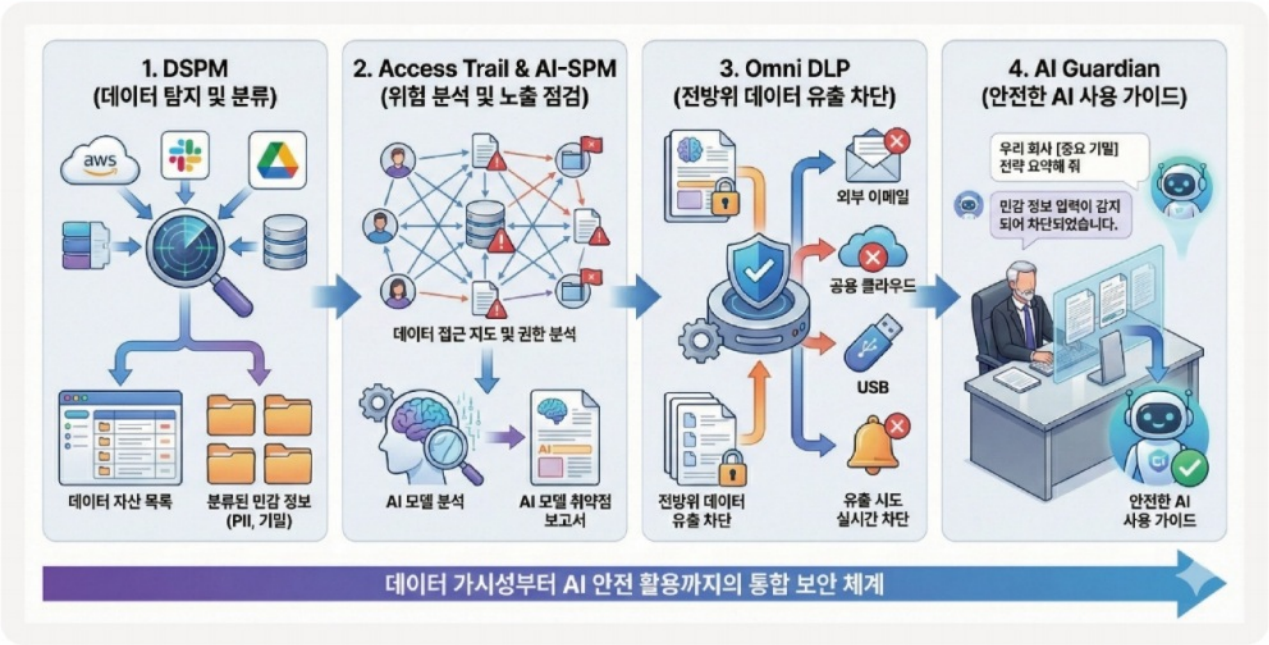
사이에라(Cyera): AI를 통한 데이터 맥락 파악

최근 발생한 국내외 보안 사고뿐만 아니라 AI로 인한 데이터 사각지대가 나타나며 보안 스타트업이 각광을 받고 있다. 사이에라(Cyera)는 이스라엘 사이버 보안 부대 Unit 8200[9] 출신의 요탐 세게브(Yotam Segev)가 설립한 스타트업으로, 그는 군 복무 시절 기존의 수동적 보안 방식이 클라우드, 나아가 AI 환경에 적합하지 않다는 문제의식을 갖고 있었다. 창업 초기 100여 명의 정보보안최고책임자(CISO)를 인터뷰하며, 이들 대다수가 자사의 어떤 데이터가 어디에 있고, 얼마나 안전하고, 누가 접근할 수 있는지에 대해 명확하게 대답하지 못하는 것에서 사업의 가능성을 확인했다.

그동안 관습적으로 사용되던 보안은 '데이터 손실 방지(Data loss prevention, 이하 DLP)' 방식으로, 데이터가 새나가지 않도록 걸어 잠그는 데 집중했다. 이 방식을 택할 경우, 데이터를 기반으로 다양한 작업을 수행하는 클라우드나 AI 환경에서는 보안이 새로운 시도나 발전을 저해할 수 있다. 이에 대한 보완책으로 등장한 '데이터 보안 태세 관리(Data security posture management, 이하 DSPM)' 방식은 데이터를 발견하고 이해하며 위험성을 파악하는데 집중한다.

사이에라는 데이터 손실 방지(DLP)와 데이터 보안 태세 관리(DSPM) 방식이 실시간 AI 작업 과정에도 매끄럽게 구현될 수 있도록, 머신러닝, 개체명 인식(Named entity recognition), 대규모 언어 모델(LLM) 등의 기술을 사용하는 DataDNA를 개발했다. 이를 기반으로 클라우드와 온프레미스 환경에 저장된 데이터를 스캔해 구조화하고, 맥락을 바탕으로 위험도를 평가해 적합한 보안 수준을 부여한다. 예를 들어, 어떤 데이터가 '재무팀 소유이고, 아마존 웹 서비스 서울 권역에 있으며, 민감도가 높고, 유출 시 500만 달러의 위험이 있다'면, '접근 가능한 인원을 3명으로 제한하고, 다중 인증을 필수화해야 한다'고 판단하는 것이다.

사이에라(Cyera) 통합 데이터 보안 플랫폼 워크플로우



① 기업 내 모든 환경을 스캔해 데이터 속성과 위치 파악(DataDNA 기술 활용) ② 데이터와 사용자 간 상호작용 추적(누가 데이터를 열람하고 있으며, 그 권한이 적절한지 판단) ③ AI 기반으로 실시간 유출 차단 ④ 기업 내부 AI 인프라 및 사용 환경 보호(기밀 데이터의 프롬프트 포함 여부 등을 가이드)

아직 구체적 매출을 공개하고 있지는 않으나, 사이에라는 지난 18개월 동안 10개국으로 해외 진출을 확대하고, 포춘 500대 기업 고객을 3배 이상 늘렸다. 차세대 보안 스타트업으로 부상하며 작년 연말에는 블랙스톤(Blackstone)으로부터 4억 달러 투자 유치에 성공해 기업가치가 90억 달러에 이르기도 했다. 지난 6월 60억 달러로 평가받은 지 불과 6개월 만에 50%가 증가한 것이다.

문제 정의로 시작해 도메인 전문성과 독자 기술로 완성

앞서 언급된 스타트업들은 각각 다른 분야에서 두각을 나타내고 있다. 그럼에도 공통점을 발견할 수 있었는데, **첫 번째는 모두 해결하고자 하는 문제가 명확했다는 점이다.** 단순히 어떤 문제가 어려움을 겪고 있다는 수준을 넘어, 인력과 시간이 비효율적으로 투입되는 영역을 정량적으로 명확하게 파악하고 이를 AI로 해결하고자 했다. 이러한 문제의식은 시장에서 높은 비용에 거래되는 전문직 영역으로 연결되기도 했다. 또한, 파일럿 과정에서도 해당 서비스를 사용했을 경우 절감되는 시간과 비용을 산출해 고객과 커뮤니케이션시 설득력을 높였다.

두 번째로, 이들은 AI 전문가이기에 앞서 각자가 해결하고자 하는 영역의 전문가였다. 빅테크 대비 경쟁력을 확보할 수 있었던 비결도 여기에 있다고 할 수 있다. 자체 모델이나 시스템을 만드는 과정에서 사용자들이 무엇을 필요로 하고, 그러기 위해서는 어떤 데이터가 갖춰져야 하는지 명확히 알고 있었던 것이다. 이들은 자신의 강점을 객관적으로 인지하고 있기에 타 산업으로의 확장보다는 해당 영역에서 서비스를 확장해 나가는 방식으로 포지셔닝을 공고히 하고 있다.

이처럼 투자자들이 꾸준히 찾는 미국 AI 스타트업들의 특징에는 기존의 성공하는 스타트업 공식과 일치하는 것도, 새로운 것도 존재한다. 스타트업의 핵심인 문제 정의는 여전히 유효함을 확인할 수 있었다. 여기에 AI 스타트업은 '도메인 전문성과 독자 기술력을 결합하는 역량'도 확보할 때, 비로소 고객과 시장에게 매력적인 플레이어가 될 수 있을 것으로 보인다.

주석

1. 2024년 AI 기업에 대한 투자는 전년 대비 약 25% 증가한 2,520억 달러를 기록했다. : Global corporate investment in AI by investment activity (Stanford HAI, 2025)

2. 전년 대비 45% 증가한 1,510억 달러에 이르렀으며, 건당 평균 투자액은 약 44% 증가한 4,540만 달러를 기록했다. : Global private investment in AI & Average size of AI private investment events (Stanford HAI, 2025)

3. 2024년 미국에서는 약 1,090억 달러가 비상장 AI 기업에 투자되며, 2위인 중국(90억 달러)과 3위인 영국(44억 달러)을 크게 앞섰다. : Global private investment in AI by geographic area (Stanford HAI, 2025)

4. 라운드 당 1억 달러 이상의 높은 투자금을 유치한 미국 스타트업이 2024년 73개에서 2025년 123개로 약 두 배 가량 증가했기 때문이다. : Seed ~ Series G 대상 최소 한 번 이상 1억 달러 이상 투자 유치한 스타트업 (Crunchbase Database)

5. 생성형 AI의 원년이라 할 수 있는 2022년부터 지난 4년간 수차례에 걸쳐 대규모 투자를 유치한 미국 AI 스타트업은 총 열한 개였으며, 그 중 작년에만 두 번 이상 대규모 투자 유치에 성공한 곳은 네 군데였다. : 설립 연도 순으로 Abridge, Cyera, Harvey, Cognition 해당. 라운드 당 10억 달러 이상 유치한 Anthropic, Thinking Machine Lab 등의 아웃라이어는 분석의 일관성을 위해 제외 (Crunchbase Database)

6. 병원에서 널리 사용하는 표준 문서 양식 : 의사가 환자 진료 후 작성하는 구조화된 진료 기록은 주로 SOAP 형식을 따르며, 이는 S(subjective, 환자가 직접 말한 증상) – O(Objective, 의사가 관찰한 사실) – A(Assessment, 의학적 판단) – P(Plan, 향후 치료 계획)로 구성

7. 질환 심각도 코드 : 의학계에서는 HCC(Hierarchical Condition Category) 코드라고 일컬음. 환자가 가진 만성질환의 심각도를 숫자로 표현 (e.g. HCC 38은 합병증이 없는 당뇨, HCC 37은 만성 합병증이 있는 당뇨)

8. 법률 사고 체계를 갖춘 맞춤학습 모델 : 법률 맞춤학습 모델(Custom-trained model)은 어떤 문서를 요약해달라는 명령을 넣으면 단순히 텍스트를 요약하는 것이 아니라, 판례를 찾고, 조항을 참고하고, 논리를 구성하고, 근거를 붙이는 법률 추론 구조를 따라 요약을 수행한다.

9. 이스라엘 사이버 보안 부대 Unit 8200 : 이스라엘 군사 정보기관인 아만(Aman)의 한 부서로, 통신·전자 신호 수집에 능숙하며 상대국 침투를 위한 사이버 보안 시스템 개발 등의 업무 수행. 8200 부대 출신이 창업한 스타트업 중 5개 이상이 미국 주식시장에 상장했으며, 이외에도 여러 곳이 실리콘밸리에서 두각을 나타내고 있음.