

1회 AI는 무엇을 대신하고 무엇을 못 하는가

글. 정누리 (주)퍼블릭에이아이 기획교육부 이사

'방송인을 위한 AI 기술 교양' 연재 목차

PART 1. AI 기술의 기본 구조와 작동 원리

- 1회. AI는 무엇을 대신하고 무엇을 못 하는가
- 2회. 컴퓨터는 왜 '의미'를 모르는가
- 3회. AI는 어떻게 학습하는가
- 4회. 모델은 어떻게 예측하고, 왜 구조는 반복되는가
- 5회. AI 성능은 무엇으로 결정되고, 왜 틀리는가

PART 2. 방송 업무를 구성하는 핵심 AI 기술

- 6회. 듣는 일: 음성 인식 AI의 구조
- 7회. 보는 일: 영상 인식 AI의 구조
- 8회. 고르는 일: 편집과 추천 AI의 구조
- 9회. 요약하는 일: 텍스트·영상 요약 기술
- 10회. 반복하는 일: 방송 워크플로우 자동화 AI

일상과 업무로 들어온 AI

2026년 어느 방송국 제작팀. PD가 ChatGPT에 프로그램에 대한 간략한 콘셉트를 정리하여 제시하면서 “우리 프로그램 오프닝 멘트 초안 좀 써줄래?”라고 요청한다. 몇 초 만에 세 가지 버전의 대본이 생성된다. 톤앤매너를 조정해달라고 하자 즉시 수정안이 나온다. 이번에는 Gemini에 동일한 요청을 해본다. 예전 같으면 작가가 여러 시간 고민하며 초안을 작성했겠지만, 이제는 생성형 AI가 그 시작점을 제공한다. 유료 버전 구독료를 지불하며 사용한 지 이미 오래다. 편집팀에서는 ChatGPT에 “3시간 인터뷰 내용을 5분 하이라이트로 편집하려는데, 어떤 구간을 중심으로 구성하면 좋을까?”라고 물어보고 구성 아이디어를 얻는다. 음악팀은 AI 작곡 도구로 배경음악 초안을 만들어 작곡가에게 전달한다.



그런데 이 생성형 AI는 정말 창작하는 것일까? PD의 의도를 이해하고 프로그램 콘셉트에 맞는 멘트를 쓴 것일까?

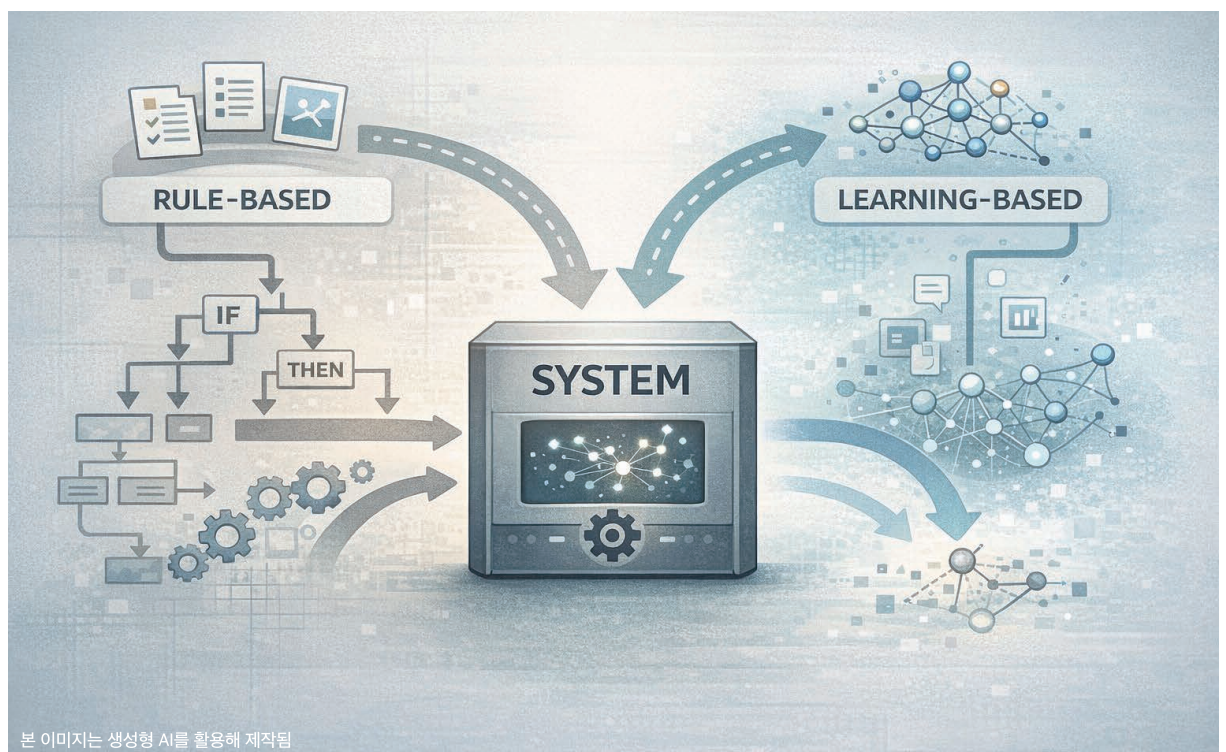
아니다. 생성형 AI는 학습한 수십억 개의 텍스트에서 '오프닝 멘트'라는 맥락에서 자주 등장하는 단어 패턴, 문장 구조, 수사법을 조합해 통계적으로 그럴듯한 문장을 만들어낸 것이다. 의미를 이해한 게 아니라, 패턴을 계산한 것이다. 하지만 그 결과물은 실무에서 충분히 유용하다. 이것이 생성형 AI 시대의 핵심 변화다.

2016년 알파고가 이세돌 9단을 이겼을 때, 충격은 컸지만, 그것이 내 일상을 어떻게 바꿀지는 피부로 와닿지 않았다. 이후 다양한 전자 제품에 유행처럼 'AI'라는 키워드가 붙었지만 '이게 시인가?'라는 생각도 해봤을 것이다. 그런데 불과 몇 년 만에 ChatGPT로 대표되는 생성형 AI는 이제 우리 손안에 들어와 일상적으로 유용하게 사용하는 기술이 되었다.

그렇다면 'AI(Artificial Intelligence, 인공지능)'는 정확히 무엇인가? AI의 창시자 중 한 명인 존 매카시(John McCarthy)는 '지능적인 기계를 만드는 과학과 공학'으로 AI를 정의했다. 인간의 지각, 추론, 학습, 언어 이해 등 '지적 능력'을 컴퓨터 프로그램을 통해 구현하려는 학문과 기술을 인공지능이라고 부른다. 하지만 '인공지능'이라는 이름은 오해를 불러일으킨다. AI는 실제로 '지능을 가진 존재'가 아니다. 오히려 인간의 복잡한 판단과 창작 과정을 어떻게 '자동화'할 것인가를 고민하여 만드는 기술이다. 즉, AI의 본질은 인간의 일을 기능 단위로 분해하고, 각 기능을 데이터 처리 방식으로, 체계적으로 구현하여 '자동화하는 시스템'이라는 데 있다. 방송 현장에서 AI가 무엇을 대신할 수 있고, 무엇을 대신할 수 없는지 이해하려면, 이 본질부터 명확히 인식해야 한다.

AI를 구현하는 두 가지 방식: 규칙 기반과 학습 기반

AI를 이해하려면 AI가 어떻게 만들어지는지 알아야 한다. AI를 구현하는 방법은 크게 두 가지로 나뉜다. 하나는 '규칙 기반 인공지능'이고, 다른 하나는 '학습 기반 인공지능'이다. '인공지능'이라는 이름 때문에 사람처럼 생각하는 존재로 오해하기 쉽지만, 실제로는 '자동화 시스템'의 일종이다. 이 점을 명확히 하기 위해 잠시 '인공지능'을 '인공지능 시스템'이라고 불러보자.



먼저, **규칙 기반 인공지능 시스템**은 명확한 조건과 규칙을 미리 정의해두고, 컴퓨터가 그 규칙을 기계적으로 따라가는 방식이다. 'if-Then(만약 ~라면, ~하라)' 식의 논리로 직접 코딩해서 주입하는 방식으로, 전문가 시스템(Expert System)이 대표적 예시이다. 일반적인 소프트웨어가 단순히 '기능'을 수행한다면, 규칙 기반 인공지능 시스템은 '인간의 전문적인 추론 과정'을 흉내 내는 데 초점을 맞춘다.

방송계에서도 전문가 시스템은 매우 활발하게 사용되어 왔으며, 특히 사람이 일일이 챙기기 힘든 복잡한 규칙과 시간 싸움이 관건인 분야에서 큰 힘을 발휘했다. 과거에는 베테랑 편성 PD들이 머릿속으로 규칙들을 다 계산하여 편성표를 짰지만, 90년대부터는 이를 전문가 시스템으로 구현하여 자동으로 최적의 편성표를 짜고 광고를 배치했다. 현재의 BMS(Broadcast Management System)의 뿌리가 바로 이 규칙 기반 시스템이다.

규칙 기반 시스템의 장점은 명확하다. 예측 가능하고 정확하다. 왜 그렇게 작동했는지 설명할 수 있고, 오류가 발생하면 어느 규칙이 문제인지 추적하기 쉽다. 하지만 한계도 분명하다. 미리 정의되지 않은 상황에는 대응할 수 없다. 예외 상황이 생길 때마다 새로운 규칙을 추가해야 한다. 규칙이 수백, 수천 개로 늘어나면 관리가 어려워지고, 규칙 간 충돌도 발생한다.

다른 인공지능 시스템 구현 방식은 '학습'이다. 최근 인공지능으로 불리는 대부분의 기술은 **학습 기반 인공지능 시스템**을 뜻한다. 규칙을 직접 설정하는 대신, 데이터에서 패턴을 학습하도록 만드는 방식이다. 이미지에서 얼굴을 인식하여 화장을 한 듯 보정하거나, 흑백사진을 컬러 사진으로 변환하는 작업은 규칙으로 정의하기 거의 불가능하지만 학습을 통해 처리할 수 있다. 대규모 데이터를 학습한 인공지능 시스템은 새로운 데이터에서도 비슷한 패턴을 인식하고 처리할 수 있다.

학습 기반 인공지능의 강점은 유연성이다. 명시적으로 규칙을 정의하지 않아도 복잡한 패턴을 인식하여 처리한다. 하지만 단점도 있다. 왜 그런 결과가 나왔는지 설명하기 어렵다. 학습 데이터에 없던 상황에서는 엉뚱한 결과를 낼 수 있다. 그리고 무엇보다, 데이터에서 학습한 '통계적 패턴'을 따를 뿐, 의미를 이해하지 않는다.

시가 잘하는 일: 분류와 예측, 패턴 인식과 조합, 그리고 반복

시는 특정 종류의 작업에서 인간을 압도한다. **첫째, 분류와 예측 작업이다.** 시는 주어진 데이터를 미리 정의된 범주로 구분하는 일을 잘한다. 음성을 '남성', '여성', '아동'으로 구분하는 작업이나 이메일을 '스팸'인지 아닌지 구분하는 일을 쉽게 해낸다. 최근에는 감정 분류도 가능해졌다. 영상 속 출연자의 표정을 분석해 '기쁨', '슬픔', '놀람', '분노'로 자동 태깅하거나, 댓글을 '긍정', '부정', '중립'으로 분류하는 시가 실무에 도입되고 있다. 한편, 주어진 데이터를 범주가 아닌 연속적인 수치나 확률을 추정하는 예측 작업도 잘한다. 접속 빈도 하락, 검색 패턴 변화, 관심 콘텐츠 소진 등을 파악해 '이 고객이 한 달 안에 구독을 해지할 확률'을 계산한다. 그 결과는 해지 확률이 높게 점쳐지는 고객에게만 특별 할인 쿠폰을 보내거나, 그가 좋아할 만한 새로운 오리지널 콘텐츠를 푸시 알림으로 보내는 방식으로 활용된다.

둘째, 패턴 인식과 조합이다. 시는 복잡한 데이터에서 인간이 놓치기 쉬운 규칙성을 찾아낸다. 예를 들어 '고양이'라는 개념을 이해하는 게 아니라, 고양이 이미지들의 픽셀 배치 패턴을 인식한다. 수백만 장의 사진을 통해 '뽀족한 귀', '가느다란 눈동자', '수염의 위치' 등이 어떤 수학적 패턴으로 배치되는지 학습한다. 또한 '사이버펑크' 스타일의 색감(보라, 네온블루) 패턴도 따로 인식해둔다. '사이버펑크 스타일의 고양이'라는 요청을 받으면, 고양이의 형태 패턴 위에 사이버펑크의 색감과 질감 패턴을 수학적으로 겹쳐서(diffusion) 이전에 없던 이미지를 생성한다. 데이터를 학습하여 패턴을 인식하고 조합해 '새로운 형태'를 만들어내는 능력, 이것이 생성형 시가 동작하는 방식이다.



본 이미지는 생성형 AI를 활용해 제작됨 (각주: "사이버펑크 고양이 일러스트 그려줘"라는 명령에 고양이가 담배를 피우고 있고, 배경에 일본어가 쓰여진 간판이 있어 몇 번의 편집 과정을 거쳤다.)

마지막은 반복이다. 분류와 예측, 패턴 인식과 조합 그 자체만으로 AI가 인간을 압도한다고 볼 수는 없을 것이다. 중요한 차이 점은 AI는 사람이 아닌 컴퓨터 시스템이기 때문에 이러한 작업을 쉼 없이 '반복'할 수 있다는 점이다. 예를 들어, 방송 편집자가 3시간 분량의 인터뷰 내용을 자막으로 작성하는 작업을 직접 하면 수 시간이 걸리지만, AI는 몇 분 안에 음성 파일을 텍스트 파일로 변환 처리한다. 더 중요한 것은, 이러한 반복 작업에서 사람은 집중력을 잃거나 실수를 범하기 쉽지만, AI는 처음부터 끝까지 동일한 성능을 유지한다는 점이다. 이것이 AI가 '반복적 노동'을 근본적으로 바꾸는 이유다. 그렇다면 AI는 우리가 하는 일의 모든 것을 바꿀까? 그렇지 않다. 여전히 넘지 못하는 근본적인 한계가 있다.

AI가 여전히 못하는 일: 진정한 맥락 이해, 독창적 창조, 윤리적 책임

AI가 아무리 발전해도, 여전히 넘지 못하는 벽이 있다. **첫째, 진정한 맥락 이해의 한계다.** 생성형 AI는 문장 간 관계를 계산할 수 있지만, 그것이 '이해'는 아니다. 예를 들어 방송 편집자가 "이 장면은 감동적이니까 음악을 깔아야 해"라고 판단할 때, 그 판단에는 영상의 전체 흐름, 출연자의 표정과 대사, 프로그램의 톤, 시청자가 느낄 감정에 대한 예측이 모두 포함되어 있다. 하지만 생성형 AI에 '감동적'이라는 개념은 학습 데이터에서 추출한 통계적 패턴일 뿐이다. '슬픈 표정 + 느린 음악 + 클로즈업 = 감동 장면'이라는 상관관계를 학습했을 뿐, 왜 이 장면이 감동적인지, 이 맥락에서 음악이 적절한지 '이해'하지 않는다. ChatGPT가 긴 대화를 기억하고 반응하는 것 같이 보여도, 그것은 앞서 주고받은 단어 시퀀스의 확률을 계산한 것일 뿐이다.

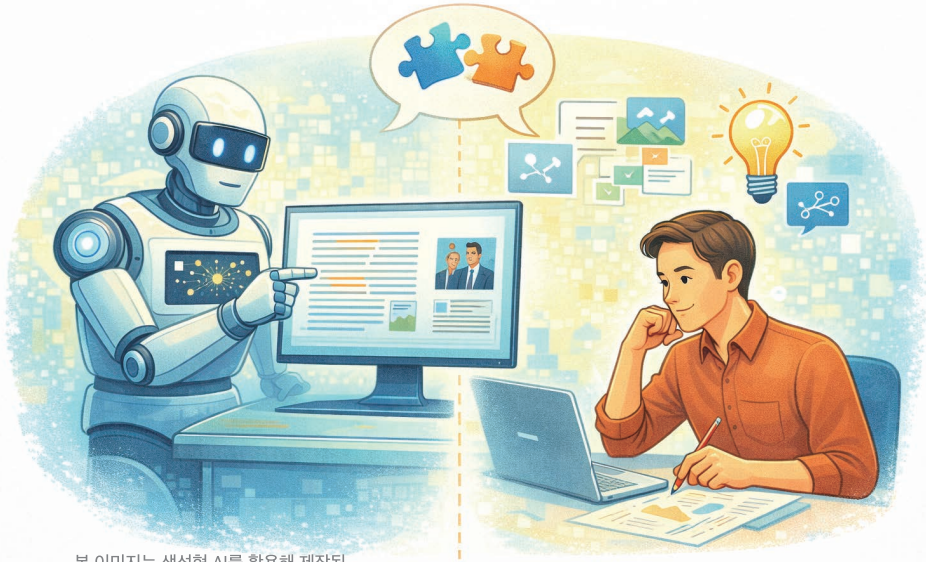
둘째, 무(無)에서의 창조는 어렵다. AI는 수십억 개의 데이터를 섞어 가장 그럴듯한 다음 조각을 찾는 데 능숙하지만, 기존에 없던 완전히 새로운 장르를 만들거나 철학적 패러다임을 전환하는 일은 못 한다. 완전히 새로운 편집 스타일을 '발명'하거나, 기존 관습을 깨는 실험적 시도를 '결단'하는 것은 인간 창작자의 영역이다. 아무리 수십억 개의 데이터를 섞더라도, 그 결과는 항상 기존 패턴의 '통계적 가능성'에 머문다. 반면 인간의 창조는 데이터에 없던 질문에서 시작된다. "왜 지금까지 이렇게만 들어왔을까?"라는 의문과 "이건 이 방식이 아니라 전혀 다르게 해봐야 한다"는 결단이 새로운 장르를 만든다. 이것은 단순히 패턴을 잘 조합하는 능력으로는 구현될 수 없는, 의도적 선택의 영역이다.

또 하나 중요한 한계가 있다. AI는 책임을 지지 않는다. 편집자가 특정 장면을 삭제하거나, 자막을 수정하거나, 음악을 선택하는 결정에는 윤리적·법적 책임이 따른다. 명예훼손, 저작권 침해, 편파 편집 같은 문제가 발생하면 편집자와 방송국이 책임을 진다. 하지만 AI는 책임을 지지 않는다. AI 도구가 아무리 정교해져도, AI가 만든 결과물에 문제가 발생하면, 결국 그 결과를 검토하고 최종 결정을 내린 사람이 책임을 진다. 이것이 AI를 '보조 도구'로 봐야 하는 이유다.

'지능'이 아니라 '처리 능력'으로서의 AI

이처럼 AI는 '지능'을 가지고 있지 않다. 하지만 여전히 생성형 AI가 유창한 문장을 쓰고, 그럴듯한 이미지를 만들어내는 모습을 보고 있다면 이 녀석이 진정 '지능'을 가진 존재처럼 느껴질 것이다. 그럴 때, AI는 데이터를 빠르게 처리하고 패턴을 조합하는 능력을 가진 '자동화 시스템'이라는 점을 다시 기억해보자. 그렇지 않고 AI를 '지능을 가진 인간과 유사한 존재'로 이해하면 오해가 생긴다. "AI가 이렇게 판단했으니 맞을 거야"라고 맹신하거나, "AI가 이걸 못하네? 쓸모없네"라고 실망할 수 있다. 때로는 친구보다 대화가 더 잘 통하니 사람보다 낫다는 생각도 할 수 있다. 이 모든 것은 AI의 본질을 오해한 결과다. AI는 특정 작업을 자동화하는 도구일 뿐이다. 그 작업이 무엇인지, 어떤 조건에서 잘 작동하는지, 어떤 한계가 있는지를 정확히 이해해야 제대로 활용할 수 있다.

특히 "AI가 알아서 해줄 거야"라는 기대는 위험하다. AI는 인간이 정의한 작업을 자동화할 뿐, 스스로 판단하고 결정하지 않는다. 편집자의 역할은 사라지지 않는다. 오히려 역할이 바뀐다. 반복 작업은 AI에게 맡기고, 편집자는 맥락을 이해하고 의도를 판단하고 창의적으로 구성하는 데 집중하게 된다. AI는 편집자를 대체하는 것이 아니라, 편집자가 본질적인 창작 작업에 더 많은 시간을 쓸 수 있도록 돕는 것이다.



본 이미지는 생성형 AI를 활용해 제작됨

AI를 제대로 이해하고 활용하기 위해 AI 기술의 작동 원리를 이해해야 한다. AI를 블랙박스로 보지 말고, 어떤 데이터를 학습했는지, 어떤 방식으로 패턴을 인식하는지, 어떤 조건에서 성능이 떨어지는지 알아야 한다. 그래야 AI 도구를 제대로 활용하고, 문제가 생겼을 때 원인을 파악하고, 더 나은 방식으로 개선할 수 있다. 이 연재의 목적도 바로 여기에 있다. AI를 막연한 대상으로 보지 않고, 구조와 원리를 이해해서 실무에 적용할 수 있는 기초를 만드는 것이다.

AI 구조 이해가 실무에서 중요한 이유

그런데 문득 이런 의문이 들 수 있다. 'AI 작동 구조와 원리를 이해하는 게 정말 실무에 도움이 될까?' 기술 세부사항을 몰라도

도구는 쓸 수 있지 않나? 맞다. 스마트폰 내부 회로를 몰라도 전화를 걸 수 있다. 하지만 스마트폰의 유용한 기능과 다양한 애플리케이션을 알고 있으면 더 똑똑하게 사용할 수 있다. AI도 마찬가지이다. AI의 구조와 원리를 이해하면 AI를 똑똑하게 사용하는 방법을 알 수 있게 된다. 즉, 문제가 생겼을 때 대응 방식이 바뀌고, 도구 선택 기준이 명확해지며, 성능을 끌어올리는 방법을 알게 된다. 구체적으로 살펴보자.

첫째, 문제 진단과 개선 능력이 생긴다. ChatGPT를 활용해 기획안을 작성하던 중, 참고문헌 목록을 요청했더니 그럴듯한 논문 제목과 저자명, 출판 정보가 나왔다. 그런데 실제로 검색해보니 그런 논문은 존재하지 않았다. 이것이 바로 생성형 AI의 '환각(hallucination)' 현상이다. AI 구조를 모르는 사람은 "이 AI 거짓말하네, 못 믿겠어"라고 실망한다. 하지만 구조를 아는 사람은 다르게 이해한다. "생성형 AI는 학습한 패턴으로 그럴듯한 문장을 생성할 뿐, 사실 여부를 검증하지 않는다. 참고문헌처럼 사실 확인이 중요한 정보는 AI가 생성한 후 반드시 검증하자."라고 판단한다. AI 결과물을 맹신하지 않고, 사실 확인이 필요한 부분은 별도로 검증하는 프로세스를 만드는 것이다. 또는 명확한 참고문헌을 찾은 후에 번역과 요약 도구로 활용하여 내용을 빠르게 파악하는 방식도 고려해볼 수 있을 것이다. 문제의 원인을 알면 대응 방법도 명확해진다.

둘째, 적합한 도구를 선택할 수 있다. 새로 론칭하는 OTT 서비스에서 사용자 행동 기반 추천시스템을 구축하려고 한다. 넷플릭스처럼 시청 이력, 평점, 시청 시간대를 분석해 개인화된 콘텐츠를 추천하고 싶다. 하지만 신규 서비스라 사용자 데이터가 충분하지 않다. 이것이 추천시스템의 고질적인 문제인 '콜드 스타트(Cold Start)' 상황이다. 구조를 모르면 '유명한 협업 필터링 모델을 도입하면 되겠지'라고 생각한다. 하지만 협업 필터링 모델은 사용자 간 유사도를 계산해 추천하는 방식이라 충분한 사용자의 행동 데이터가 필요하다. 데이터가 없는 초기에는 제대로 작동하지 않는다. 구조를 알면 다르게 접근한다. 신규 가입 시 '선호 장르 3가지를 선택하세요'처럼 명시적으로 초기 프로필을 입력받는다. 이 정보로 콘텐츠 메타데이터(장르, 감독, 배우)를 기반으로 추천한다. 동시에 전체 사용자에게 인기 있는 콘텐츠를 노출해 초기 시청 데이터를 확보한다. 사용자가 콘텐츠를 시청하고 평점을 남기면서 행동 데이터가 쌓이면, 그때 협업 필터링 모델을 단계적으로 도입한다. 구조를 알면 '지금 우리에게 필요한 건 데이터 확보'라는 목표가 명확해지고, 콜드 스타트 문제를 단계적으로 해결하는 전략을 세울 수 있다.

셋째, 성능을 높여 활용할 수 있다. 방송 프로그램 제작 과정에서 실시간 자막 생성 AI를 사용하고 있다. 인터뷰나 토크쇼를 녹화할 때 음성을 텍스트로 자동 변환하면 후반 편집 시간이 크게 단축된다. 그런데 문제가 있다. 의학 다큐멘터리를 작업하는데, 전문 용어를 잘못 인식한다. 구조를 모르면 "이 AI 정확도가 너무 낮네, 쓸모없어"라고 불평한다. 하지만 구조를 알면 다르게 접근한다. 음성인식 AI는 학습한 일반 언어 데이터를 기반으로 음성을 텍스트로 변환한다. 문제는 방송 분야별로 등장하는 전문 용어, 신조어, 인명, 지명 같은 특수 어휘는 학습 데이터에 충분히 포함되지 않았다는 점이다. 이때 활용할 수 있는 기능이 '특수용어 사전' 추가다. 대부분의 음성인식 시스템은 도메인별 전문 용어를 사전에 등록하는 기능을 제공한다. 이렇게 하면 AI는 해당 발음이 들렸을 때 일반 단어보다 사전에 등록된 특수 용어를 우선해서 선택한다.

결국, AI 구조를 이해한다는 것은 AI를 '신기한 블랙박스'가 아니라 '내가 제어할 수 있는 도구'로 보게 된다는 뜻이다. 문제가 생기면 원인을 진단하고, 도구를 선택할 때는 우리 현장에 맞는지 판단하고, 성능을 끌어올릴 방법을 찾게 된다. 이것이 AI 구조 이해가 실무에서 중요한 이유다. 기술 세부사항을 모두 알아야 한다는 뜻이 아니다. 'AI가 무엇을 학습했고, 어떤 방식으로 작동하며, 어떤 조건에서 성능이 떨어지는가'를 이해하면 된다. 그 이해가 있으면 AI는 더 이상 막연한 대상이 아니다. 내가 활용하고, 개선하고, 문제를 해결할 수 있는 실무 도구가 된다.

AI는 혁명적인 기술이 맞다. 하지만 AI의 진짜 혁명은 '컴퓨터가 사람처럼 생각하게 되었다'는 데 있지 않다. 오히려 '인간의 일을 기능 단위로 쪼개서 자동화하는 시스템을 만들었다'는 점에 있다. 더 중요한 것은, 기술 발전으로 이 자동화 시스템의 비용이 급격히 낮아지면서 이제 누구나 이 혜택을 누릴 수 있게 되었다는 사실이다. 이 차이를 이해하는 것이 AI 시대를 살아가는 우리에게 가장 중요한 출발점이다. 