DAY 02 · 1부 AI가 뭔지

AI가 말하는 '한 조각', 토큰이 뭘까?

하루 5분 AI 수업 · 2026년 9월 8일

문장은 결국, 토큰이라는 조각들의 줄이다

오늘의 답

한 줄로 말하면

1. 어제 배운 다음 조각을, AI 세계에서는 **토큰**이라고 불러.

2. 토큰은 **글자도 단어도 아니고**, AI가 계산하기 편하게 자른 조각이야 — 크기는 그때그때 달라.

3. 그래서 같은 문장도 **토큰 개수가 다르게** 세어지고, 그 개수만큼 시간과 비용이 들어.

그림 한 장

문장 → 어절 → 토큰

예시 학교 끝나고 친구랑 떡볶이를 먹으러 갔어

↓ 어절로 나누면 (사람이 읽는 단위)

학교	끝나고	친구랑	떡볶이를	먹으러	갔어

↓ 토큰으로 한 번 더 쪼개면 (AI가 세는 단위)

학교	끝나	고	친구	랑	떡볶이	를	먹으	러	갔	어

어절 6개 → 토큰 11개. 사람이 세는 단위와 AI가 세는 단위는 다르다.

같은 문장, 세 가지 방식으로 잘라보면

7개

글자 하나하나

오 | 늘 | 급 | 식 | 뭐 | 야 | ?

가장 잘게 쪼갠 것. 조각 수는 많지만 의미 단위는 아니다.

3개

띄어쓰기 기준(어절)

오늘 | 급식 | 뭐야?

사람이 읽긴 편하지만, 처음 보는 단어는 AI가 통째로 낯설어한다.

5개

토큰처럼(모의)

오늘 | 급 | 식 | 뭐 | 야?

글자보다 크고 어절보다 작을 수 있다 — 자주 쓰는 덩어리는 통째로, 낯선 덩어리는 잘게.

이렇게 조각 내는 규칙에도 이름이 있어 — 토크나이저(tokenizer).

※ 위 쪼개기는 설명을 위해 단순화한 모형이다.

확인해보기

세 문제만 풀어 보자

Q1 **“토큰”이 뭘까?**

① 항상 한 글자야
② 항상 한 단어야
③ AI가 계산하기 편하게 자른 조각이야(크기는 그때그때 달라)
④ 사람이 미리 정해 둔 단어장 전체야

Q2 **토큰 수가 많아질수록 AI에게 생기는 일은?**

① 더 빨라진다
② 계산할 것도, 걸리는 시간도 늘어난다
③ 아무 영향 없다
④ 무조건 더 똑똑해진다

Q3 **(O / X)** 같은 문장이라도 AI마다 토큰으로 쪼개는 방식이 다를 수 있다.

왜 그런지까지 알고 가자

Q1

③

토큰은 AI가 계산하기 편하게 자른 조각이야. 글자 하나일 때도, 단어 전체일 때도 있어 — 크기가 정해져 있지 않아.

Q2

②

토큰이 많을수록 AI가 계산할 것도, 걸리는 시간(과 비용)도 늘어나. 같은 내용도 짧게 쓰면 AI가 더 빨리 처리해.

Q3

O

맞아. 어떤 AI는 글자 단위로, 어떤 AI는 자주 쓰는 덩어리 단위로 쪼개 — 그래서 같은 문장도 AI마다 토큰 개수가 달라질 수 있어.

AI에게 문장은 글자도 단어도 아니라,
토큰이라는 조각들의 줄이다.



내일은 확률과 온도 — 어제 살짝 나온 '모험심 손잡이'를 더 깊이 파헤쳐 볼 거야.