DAY 55 · 5부 안을 들여다보기

# 내 폰은 AI를 혼자 돌릴 수 있을까? — 내 폰 안의 AI

하루 5분 AI 수업 · 2026년 10월 31일

폰은 전부 보내지도, 전부 품지도 않아. 할 수 있는 조각만 붙잡아 두는 거야.

# 한 줄로 말하면

1. 일부는 혼자 돌려. 요즘 폰 안에는 **숫자 30억 개쯤 되는 작은 모델**이 들어 있어서, 인터넷이 끊겨도 요약하고 다듬고 알림을 정리해.

2. 대신 큰 일은 버거워. 긴 글이나 어려운 질문은 **어제 본 그 건물**까지 보내서 답을 받아 와. 폰 안 모델과 서버 모델은 크기가 수십에서 수백 배 차이 나거든.

3. 그래서 요즘 폰은 **조각을 나눠** 써. 할 수 있는 건 안에서 끝내 — 빠르고, 전기도 덜 들고, 내 글이 밖으로 안 나가니까.

# 폰 안 조각 vs 서버 조각

### 폰 안에서 도는 조각

· 숫자 약 30억 개짜리 작은 모델
· 인터넷 없이도 바로 동작
· 내 글이 폰 밖으로 안 나감

### 서버에서 도는 조각

· 숫자 수천억 개짜리 큰 모델
· 연결돼 있어야 동작
· 어려운 일을 훨씬 잘함

폰이 먼저 해 보고 버거운 조각만 골라 내보내. 요즘 폰은 왼쪽과 오른쪽을 섞어 써.

**핵심: 전부 보내지도, 전부 품지도 않는다 — 조각으로 나눈다**

실험 · 쪼개보기

# 이 일은 폰 안에서 끝날까, 밖으로 나갈까

## 인터넷을 끊으면

**비행기 모드에서 되는 일**

- 사진 속 글자 읽기
- 알림 한 줄 요약
- 맞춤법 다듬기
- 짧은 답장 제안
- 긴 글 써 주기
- 웹에서 찾아 답하기
- 어려운 수학 풀이

초록 넷은 폰 안 작은 모델이 혼자 해. 비행기 모드에서도 돌고, 무엇보다 내 글이 폰 밖으로 안 나가.

## 인터넷을 켜면

**버거운 일만 건물까지**

- 사진 속 글자 읽기
- 알림 한 줄 요약
- 맞춤법 다듬기
- 짧은 답장 제안
- 긴 글 써 주기
- 웹에서 찾아 답하기
- 어려운 수학 풀이

주황 셋은 폰을 떠나 데이터센터까지 다녀와. 훨씬 잘하는 대신 연결과 전기와 내 글을 내주는 거야.

## 사진 한 장 정리

**한 가지 일도 조각으로 갈린다**

- 얼굴 찾기
- 글자 뽑기
- 흐린 곳 다듬기
- 무슨 장면인지 설명
- 비슷한 사진 찾기

요청 하나도 통째로 가지 않아. 폰이 먼저 해 보고 버거운 조각만 골라 내보내. 이게 요즘 폰이 실제로 하는 방식이야.

버튼을 눌러 상황을 바꿔 봐. 초록은 폰 안에서 끝나는 조각, 주황은 폰을 떠나는 조각, 흐린 건 아예 안 되는 조각이야.

웹 수업에서 직접 해 볼 수 있어요 · next-token.net/learn/day-55/

확인해보기

# 세 문제만 풀어 보자

**Q1** 폰 안에서 도는 AI 모델의 특징은?

① 서버 모델보다 훨씬 크다
② 사진은 전혀 못 다룬다
③ 전기를 더 많이 쓴다
④ 작고, 인터넷 없이도 돌아간다

**Q2** 어려운 질문을 서버로 보내면 생기는 일은?

① 더 잘하는 대신 내 글이 폰 밖으로 나간다
② 폰 배터리가 늘어난다
③ 인터넷이 없어도 된다
④ 답이 반드시 틀린다

**Q3** 요즘 폰이 실제로 하는 방식은?

① 전부 폰 안에서
② 조각을 나눠서
③ 전부 서버로
④ 아무것도 안 함

# 왜 그런지까지 알고 가자

Q1

④

2024년에 공개된 폰 안 모델은 숫자가 약 30억 개였어. 서버 모델보다 수십에서 수백 배 작은 대신, 비행기 모드에서도 돌아가.

Q2

①

큰 모델은 훨씬 잘해. 대신 내 글이 어제 본 그 건물까지 다녀오지. 얻는 것과 내주는 것이 같이 있는 거야.

Q3

②

실험 셋째 칸에서 봤지. 사진 한 장 정리도 폰이 할 조각과 서버가 할 조각으로 갈려. 폰이 먼저 해 보고 버거운 조각만 내보내.

# 폰은 전부 보내지도, 전부 품지도 않아.
# 할 수 있는 조각만 붙잡아 두는 거야.



내일은 AGI 논쟁 — 사람만큼 하는 AI가 정말 올까. 어른들이 왜 이걸로 다투는지 보자.