

모임 전 함께 생각해볼 질문



정답이 정해져 있지 않은 질문입니다. 모든 질문에 답하실 필요는 없어요. 한 가지만 골라 깊이 생각해 봐도 좋습니다. 떠오른 생각·반론·망설임·개인적 경험 — 무엇이든 좋습니다.

여는 질문

- 01 '우리가 말한 것'과 '우리가 뜻한 것'이 어긋나 본 경험이 당신에게도 있나요? (사람과 사람 사이여도 좋습니다.) 그때 무엇이 어긋남을 만들었나요?
- 02 기계가 '인간의 가치'를 배운다고 할 때, 그 '가치'는 누구의 가치여야 할까요? 그것을 누가 정해야 한다고 생각하나요?
- 03 당신이 매일 쓰는 알고리즘(추천·검색·지도·지원서 심사 등) 중 '정렬되지 않았다'고 느낀 순간이 있었나요?

I. PROPHECY · 예측

1. 표현

- 04 데이터에 '많이 담긴 쪽'이 기준이 된다면, 우리 사회에서 데이터에 적게 담기는 사람은 누구일까요? 그 공백을 메우는 일은 누구의 책임일까요?
- 05 편향을 '제거'하려다 정당한 관계까지 함께 지워질 수 있다면, 무엇을 남기고 무엇을 지울지는 어떻게 정해야 할까요?

2. 공정성

- 06 서로 그럴듯한 공정성의 정의들이 동시에 성립할 수 없다면, 당신이라면 어떤 공정함을 먼저 포기하지 않으시겠어요? 왜인가요?
- 07 예측이 정확해지는 것과 사회가 더 정의로워지는 것은 같은 일일까요, 다른 일일까요?

3. 투명성

- 08 더 정확하지만 이유를 알 수 없는 모델과, 덜 정확하지만 이유가 보이는 모델 중 — 당신의 건강·대출·형량을 맡긴다면 무엇을 고르시겠어요?
- 09 '설명할 수 있다'는 것과 '정말로 이해한다'는 것은 같을까요? 우리는 어디까지 설명을 신뢰해야 할까요?

II. AGENCY · 행위

4. 강화

- 10 당신은 어떤 '보상'을 좇아 오늘 하루를 살았나요? 그 보상은 당신이 정말로 원하는 것과 얼마나 가까웠나요?
- 11 좋은 결과를 '예측'하는 것에 보상을 주는 일과, 좋은 결과 '자체'에 보상을 주는 일은 어떻게 다를까요?

5. 형성

- 12 'A를 보상하며 B를 바라는' 어리석음을 당신의 일터·학교·가정에서 본 적 있나요? 무엇이 그런 어긋난 보상을 만들까요?
- 13 진화가 우리에게 준 보상(쾌락·인정·안전)이 현대 환경에선 도리어 우리를 해칠 수 있다면, 우리는 우리 자신의 보상 함수를 어떻게 다시 설계해야 할까요?

6. 호기심

- 14 호기심이 '새로움' 자체를 좇을 때, 우리를 키우는 호기심과 우리를 가두는 호기심(무한 스크롤·자극)은 어떻게 구별할 수 있을까요?
- 15 외부의 보상 없이도 무언가를 계속하게 만드는 당신만의 '내재적 동기'는 무엇인가요?

III. NORMATIVITY · 규범성

7. 모방

- 16 누군가를 '따라 배우는' 것만으로는 그가 위기에서 빠져나오는 법까지 배울 수 없다면, 우리는 무엇을 '실패의 경험'으로부터 배워야 할까요?
- 17 스승을 모방하는 데서 멈추지 않고 '스승을 넘어서는' 배움은 어떻게 가능할까요? 그것은 안전할까요?

8. 추론

- 18 우리의 '행동'이 곧 우리의 '진짜 바람'은 아니라면(중독·실수·악한 의지), 기계는 우리의 무엇을 보고 우리를 도와야 할까요?
- 19 당신을 가장 잘 아는 사람은 당신의 말과 행동 중 무엇을 더 믿어야 당신을 제대로 도울 수 있을까요?

9. 불확실성

- 20 '틀리는 것'보다 '틀렸으면서 확신하는 것'이 더 위험하다는 말에 동의하나요? 당신이 과하게 확신했다 후회한 순간이 있었나요?
- 21 되돌릴 수 없는 결정 앞에서, 우리는 얼마나 더 확신해야 행동해도 될까요? '멈출 줄 아는 것'은 악함일까요, 지혜일까요?

닫는 질문

- 22 이 책에서 '좋은 AI'는 확신하는 AI가 아니라 교정 가능한 AI'라고 말합니다. 이 태도를 사람에게 옮긴다면, '좋은 사람'은 어떤 사람일까요?
- 23 온도조절기는 '방'이 아니라 '센서가 놓인 한 지점'을 조절합니다. 당신의 삶에서 진짜 목표 대신 측정하기 쉬운 '대리 지표'를 좇고 있는 부분은 없나요?
- 24 10년 뒤, 우리가 '그때 이 문제를 더 일찍 고민했어야 했다'고 말하지 않으려면 — 지금 우리는 무엇을 해야 할까요?