

Machine Learning Engineer

여행 가이드

kakao

카카오 추천팀
김성진 (nick.kim)

발표자 소개

Machine Learning Engineer (개인화 추천 시스템 연구/개발)

KCC 2022



kakao

가장 많이 받았던 질문들

1. ML Engineer 로 성장하려면 어떤 준비를 해야 할까요?
2. 대학원 진학을 해야 할까요?
3. AI/ML은 어떤 분야가 전망이 좋나요? NLP? Vision?

ML Engineer 로 사회에 첫걸음을 댄다면,



ML 시스템

이미 구축된 ML 시스템을 개선하거나 운영하는 일부터 시작하게 될 가능성이 높습니다.

‘구현’, ‘엔지니어링’ 실력이 기본입니다.

ML 시스템을 개선하고 운영하는 일을 하다보면

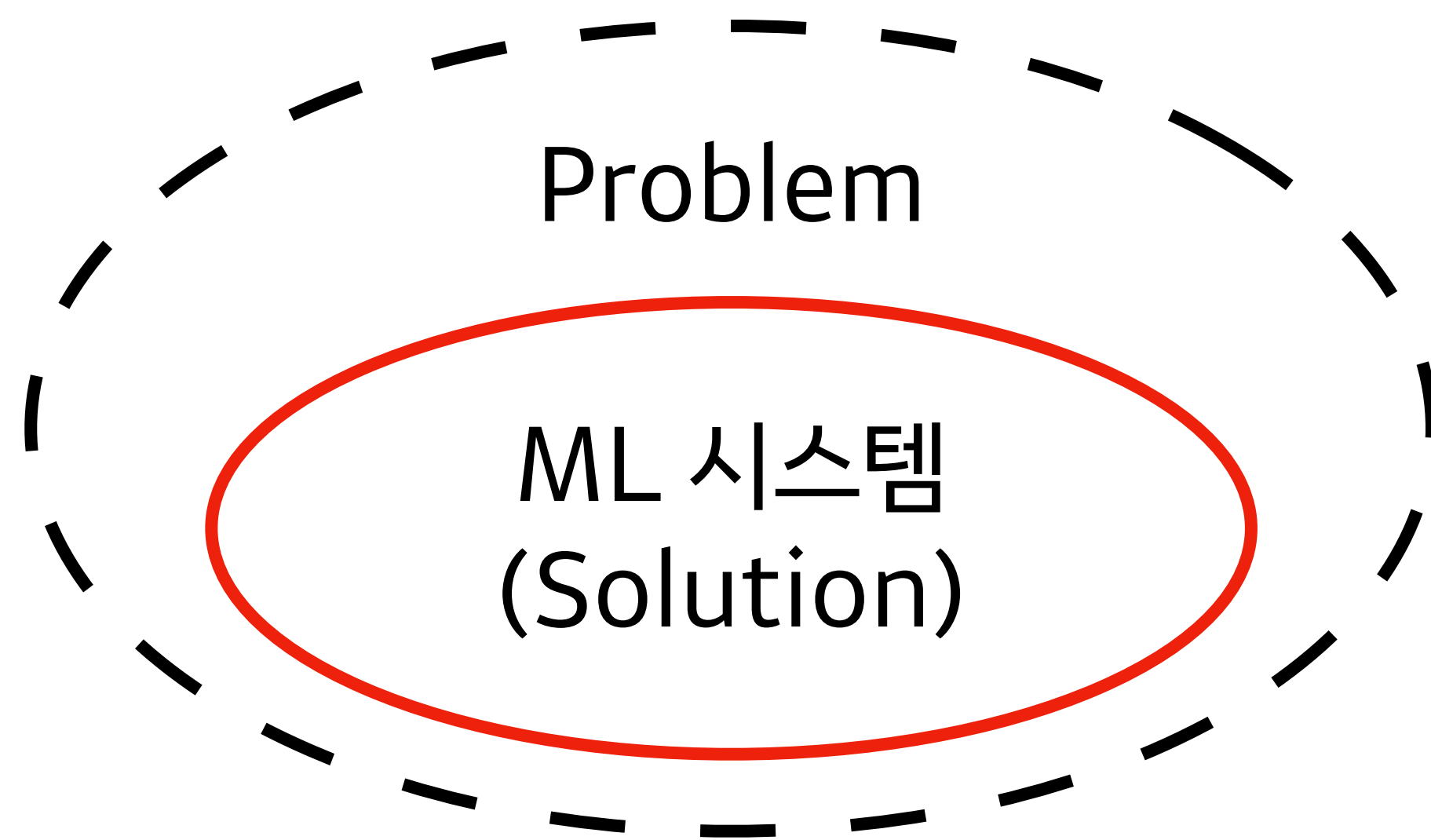


ML 시스템

자연스럽게 이런 질문이 생깁니다.

“이 ML 시스템을 ‘잘’ 만들었다는 기준이 뭘까?”

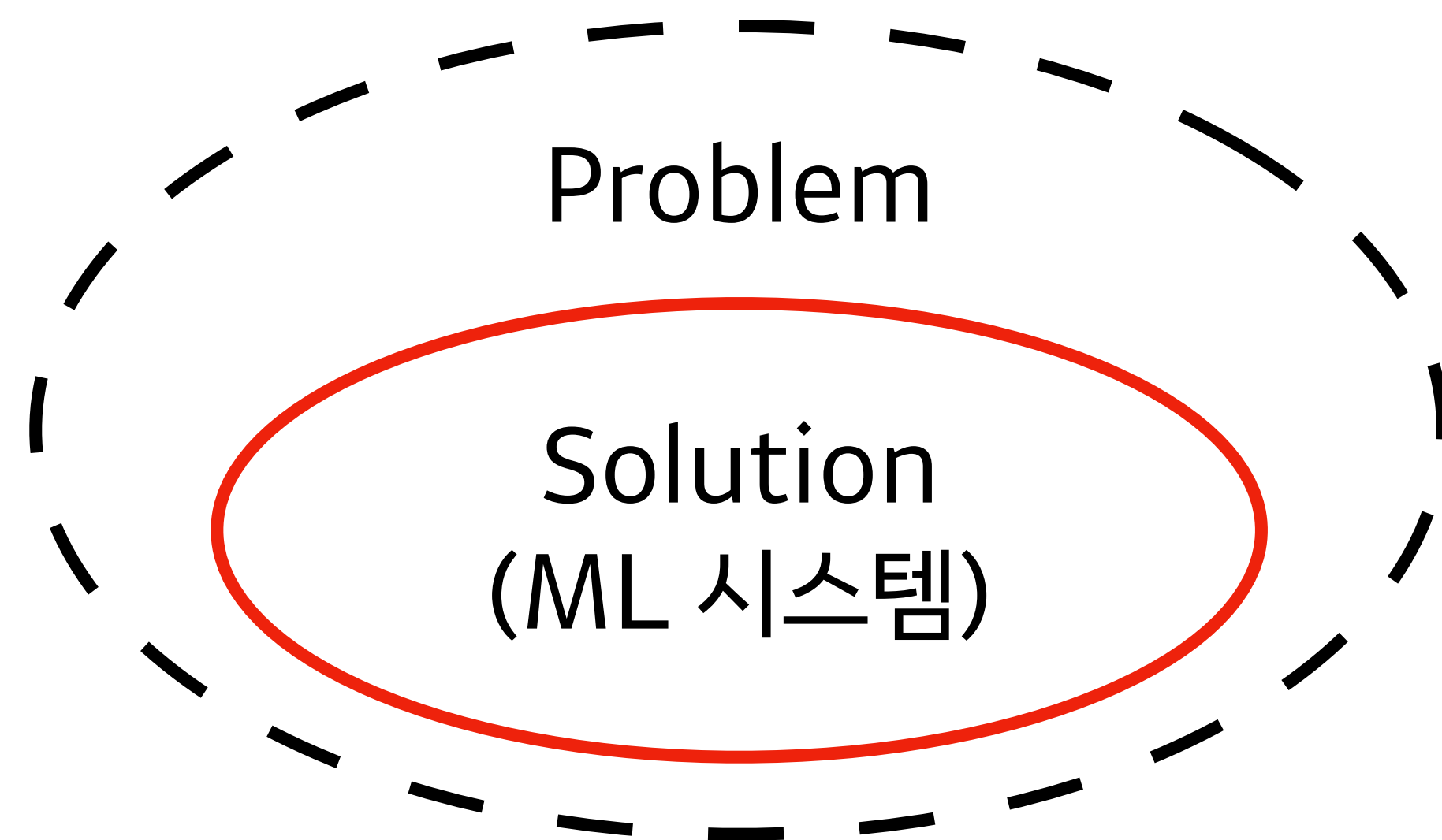
‘잘’ 만들었다는 기준이 뭘까? ..에 대해 생각하다보면,



해당 ML 시스템이 풀어야 하는 문제가 보이기 시작합니다.

내가 개선하고 있던 ML 시스템은, 어떠한 문제를 해결하기 위한 하나의 ‘**솔루션**’ 이라는걸 알게 됩니다.

내가 만들고 있는 ML 시스템을 하나의 솔루션으로 바라보면,



이런 질문이 자연스럽게 따라옵니다.

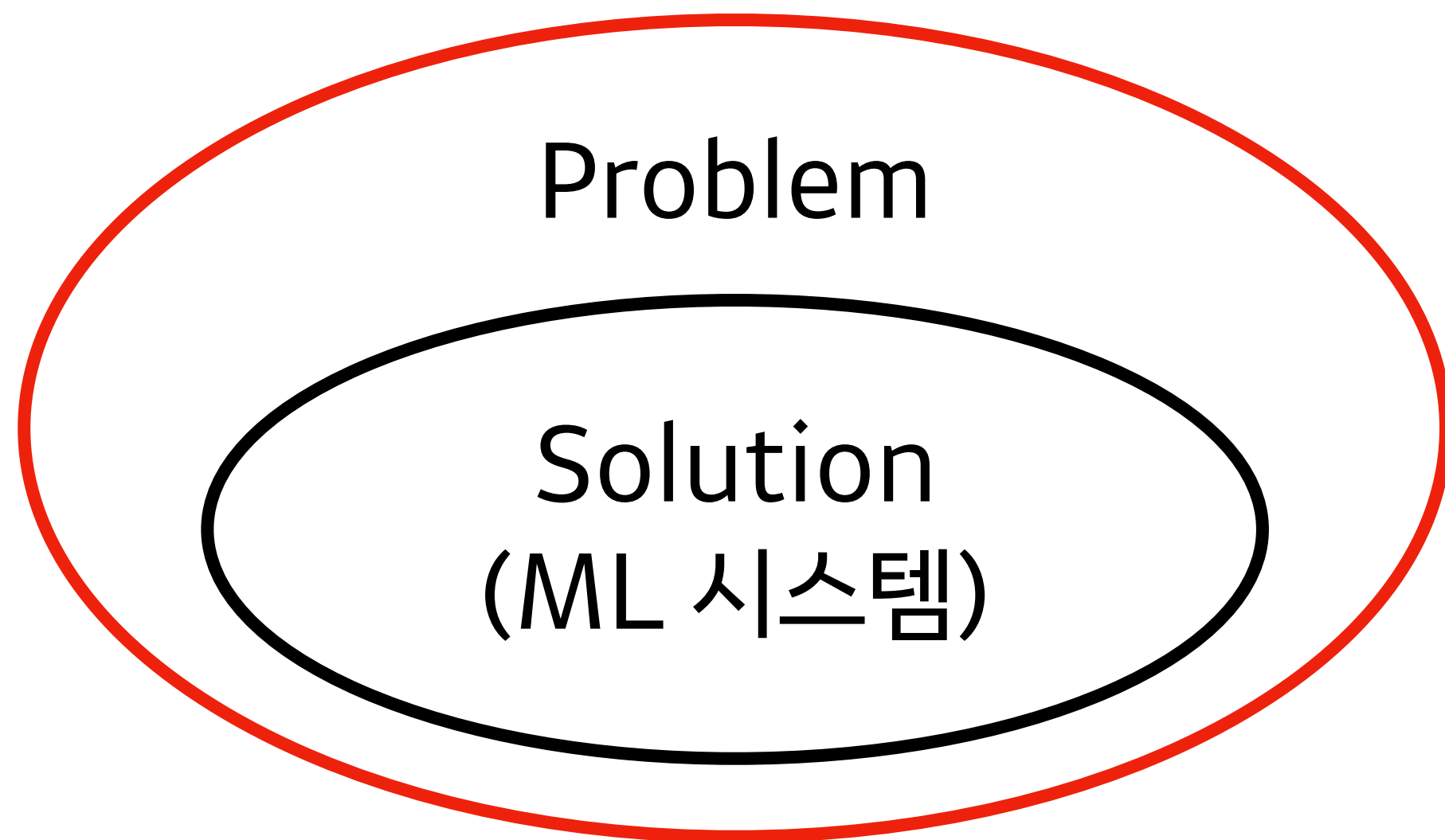
“왜 이 솔루션을 선택했을까?”

“이 솔루션이 최선일까?”

“이 솔루션이 최선일까?”

- 최고의 솔루션을 만들어가는 과정에서 끊임 없이 해야하는 질문
- 자신이 생각한 솔루션이 왜 잘 동작할거라 생각하는지 설명할 수 있어야 함.
 - (= 자신이 생각한 솔루션의 ‘승부수’가 무엇인지 설명할 수 있어야 함.)
 - 참고 영상: 실험을 잘 한다는 것은 무엇일까? (ifkakao 2022) [youtube](#)

솔루션 고도화를 하다보면,



‘문제 정의’에 자연스럽게 관심이 생깁니다.

이 단계에서 드는 질문은,

“이렇게 문제 정의하는게 최선일까?”

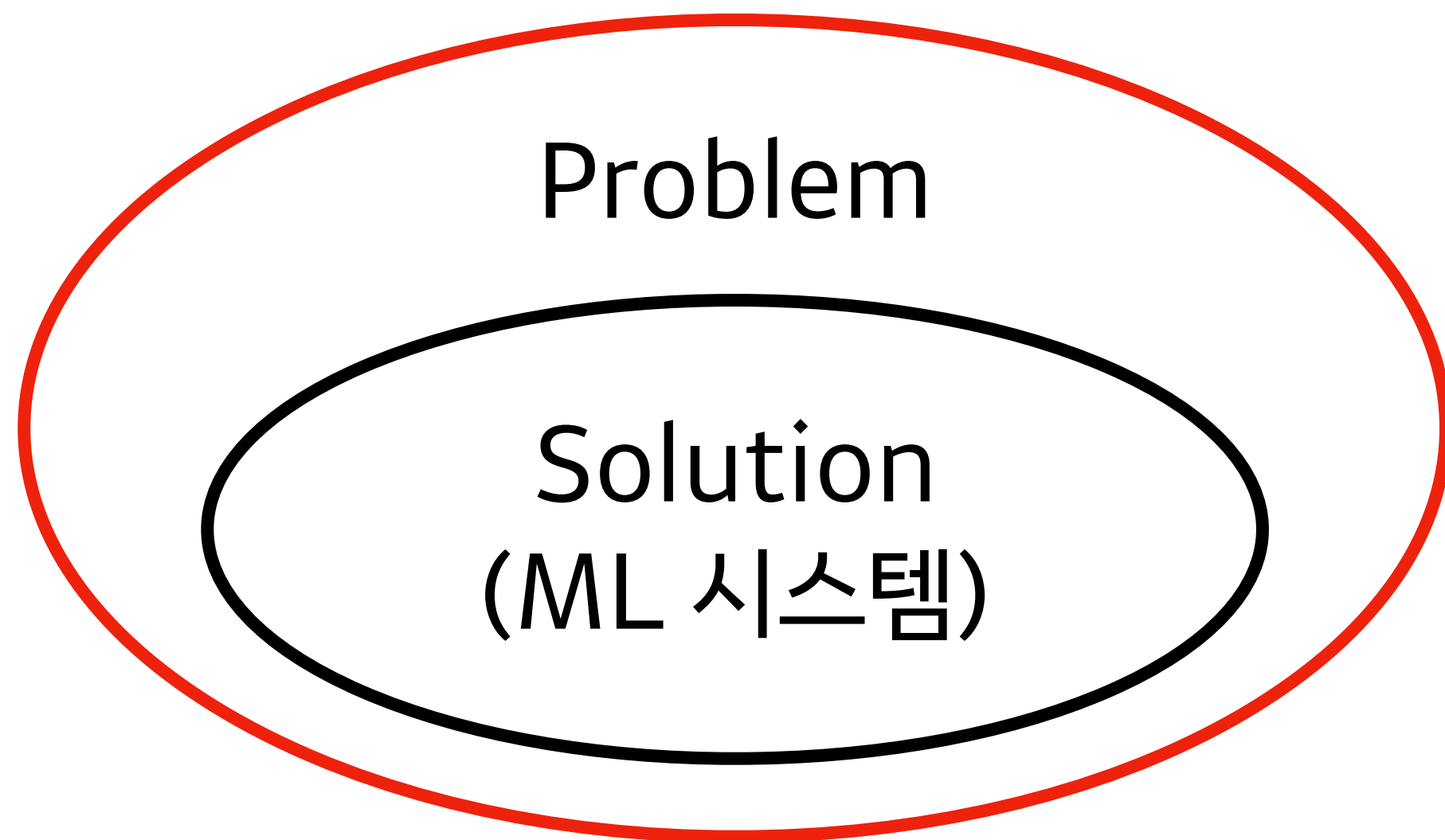
문제 정의

- 현실 세계의 문제를 잘 정의해낸다는건 쉽지 않습니다.
- 솔루션이 최적화할 수 있는 수준으로 문제를 축소해서 정의할 필요가 있습니다.
- 여기에서 실제 문제와 솔루션이 풀이하는 문제의 차이가 생깁니다.
- (이 주제에 더 관심이 있으신 분들은 'Proxy metric' 으로 검색해보시기 바랍니다.)

문제 정의 - 예시

- 풀고자하는 문제: 콘텐츠 서비스에서 개인화 추천 기술을 사용해서 이용자가 늘어났으면 좋겠다.
- 문제 정의 예시:
 - a. 클릭수를 최대화하도록 추천 로직을 최적화한다.
 - b. 이용자 잔존률을 최대화하도록 추천 로직을 최적화한다.
- 참고자료: Rule of ML([google](#)) link

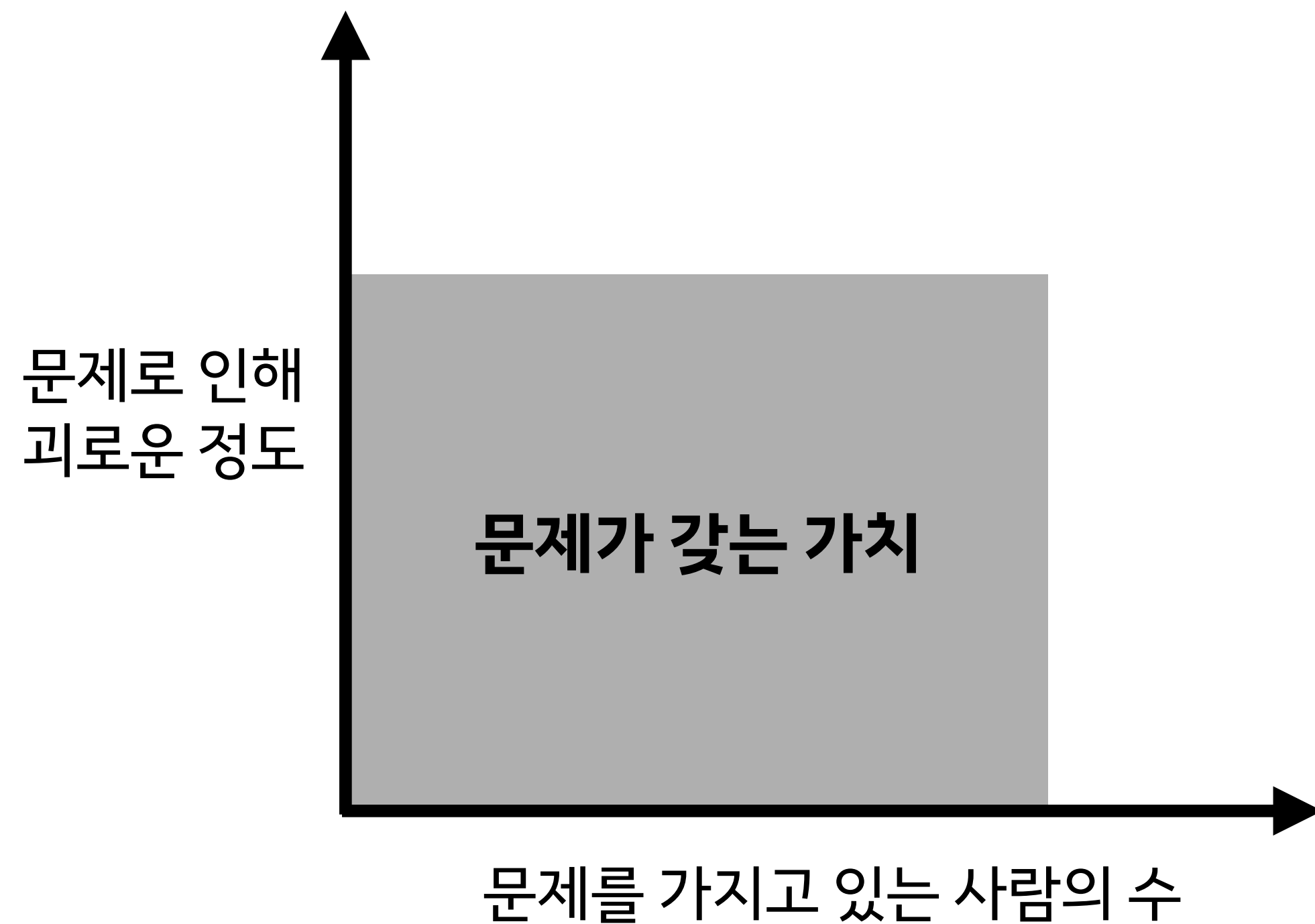
문제 정의를 열심히 하다보면,



자연스럽게 이런 질문도 하게 됩니다.

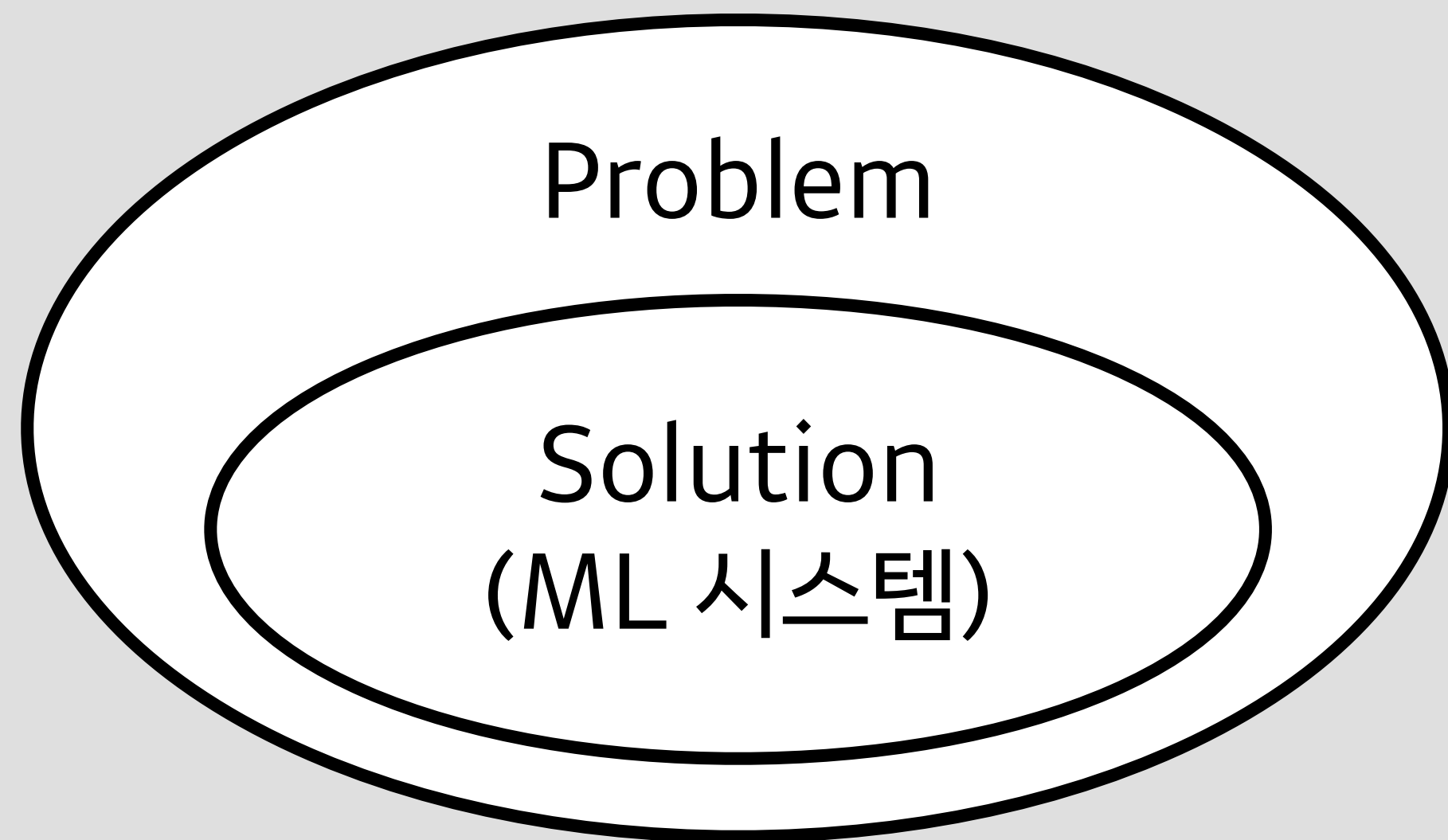
“이 문제의 가치가 어느 정도일까?”

문제의 가치



- (발표자 생각) 저는 문제의 가치가 아래 2가지에 비례한다고 생각합니다.
 - (1) 그 문제를 가지고 있는 사람의 수
 - (2) 그 문제로 인해 괴로운 정도

문제의 가치를 고민하는 것에도 익숙해지면,



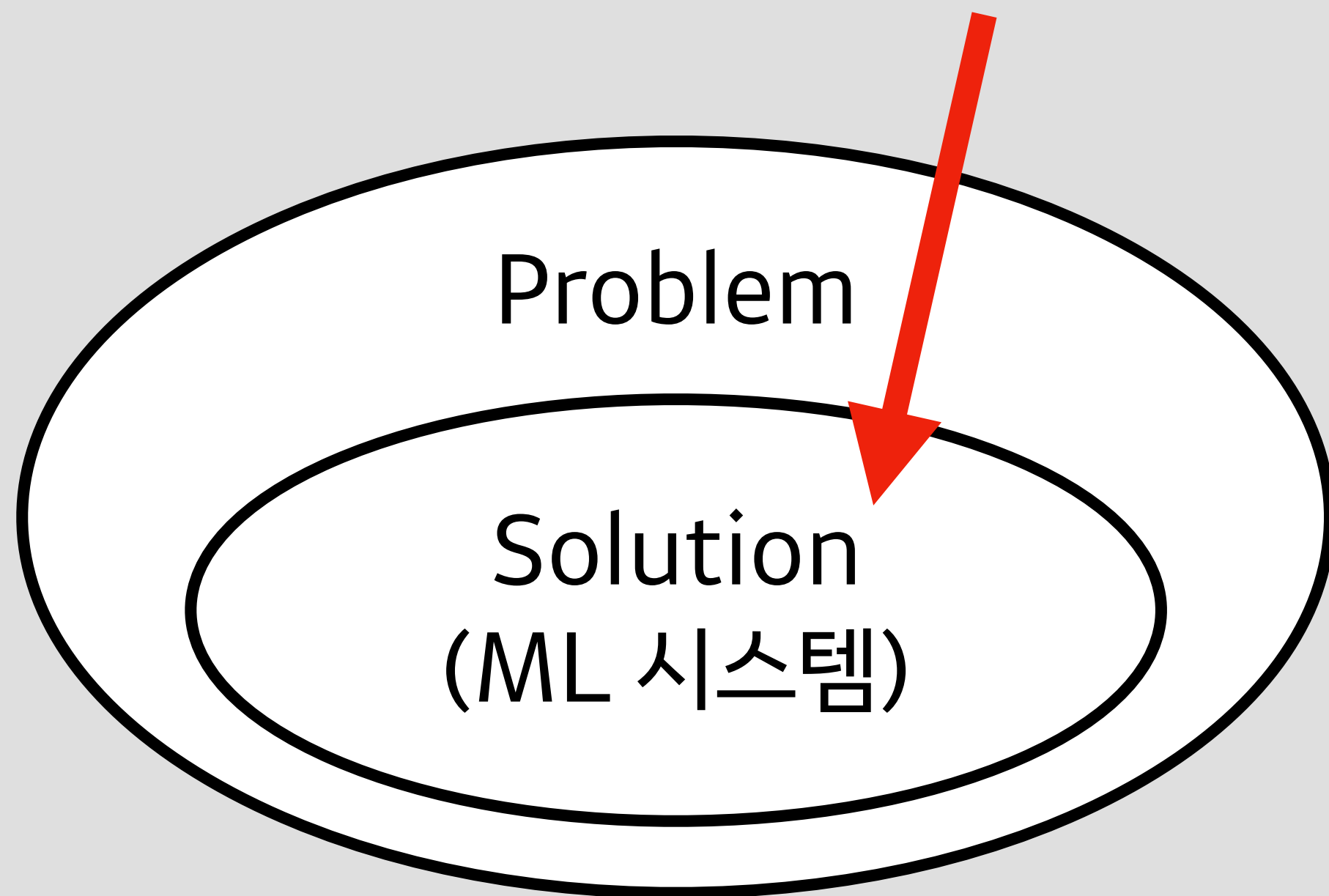
가치 있는 다른 문제들에 관심이 생깁니다.

“밖(회색 영역)에는 뭐가 있을까?”

-> 다른 재미있는 문제들이 기다리고 있습니다.

정리: 문제 해결 과정

문제 해결 과정 (바깥 -> 안)

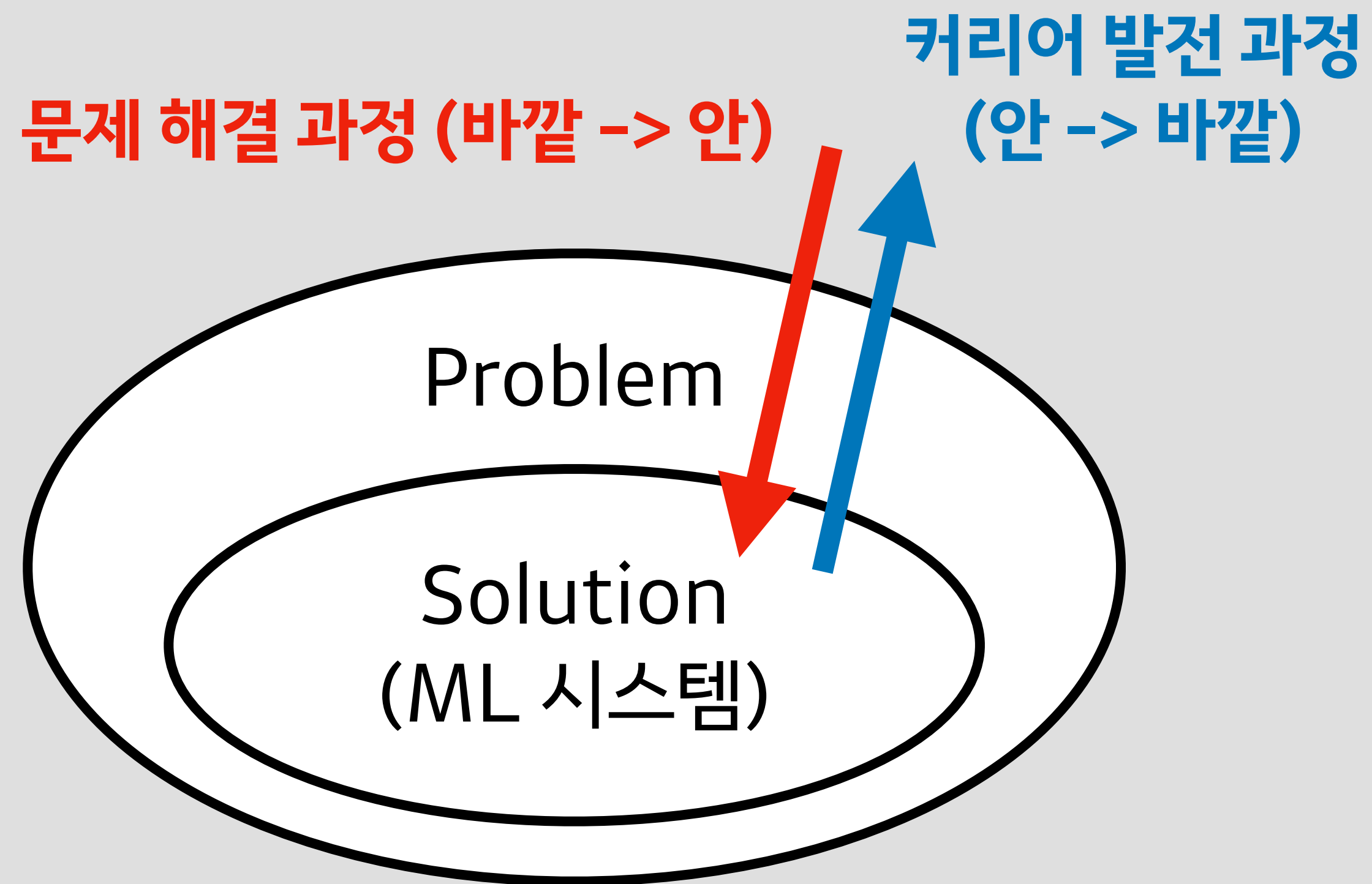


문제 해결 과정

1. 가치 있는 문제를 발견하고
2. 문제를 정의하고
3. 솔루션을 구현하고
4. 실험하고
5. 회고하고

이 과정을 반복합니다.

‘문제 해결 과정’과 ‘커리어 발전 과정’



커리어 발전 과정은 보통 솔루션에서 먼저 시작하게 됩니다.

“왜 커리어 발전 과정은 안->바깥 방향일까요?”

“왜 커리어 시작을 솔루션으로 하게 될까?”

- (발표자의 생각입니다만)
 - 첫 째, 솔루션을 결정하고 구현하는 일이 문제 정의하는 것보다 쉬운 경우가 많습니다.
 - 둘 째, 문제 해결 과정 전체의 성패에, 문제 정의가 솔루션보다 훨씬 큰 기여를 하는 경우가 많습니다.

ML Engineer 여정을 떠날 때 챙기면 좋을 준비물

1. 구현, 엔지니어링 역량
2. 자신이 선택한 솔루션이 문제 해결에 왜 도움이 되는지 설명할 수 있는 역량

Tip: 구현, 엔지니어링 역량

1. 구현 역량: 논문을 많이 구현해보시고 직접 실험해보시면 좋습니다.
2. 엔지니어링 역량
 1. 엔지니어링 역량이 중요한 이유:
 1. 결국 우리가 만드는 건 누군가가 실제로 사용할 ‘제품’이기 때문입니다.
 2. 우리가 만드는 ML 시스템이 안정적인 성능을 내는 것이 중요합니다.
 3. 현업에서는 스냅샷 데이터가 아닌 실시간 데이터를 사용합니다. 늘 균일한 품질의 데이터를 모델에 제공할 수 있는 안정적인 파이프라인이 중요합니다.
 4. 버그로 인해 실험을 다시하게 되는 경우도 많습니다.
 2. 참고자료: Rule of ML([google](#)) link

마지막으로 챙겨가야 할 것: 마음 가짐

1. 실험은 대부분 실패한다는 것.
2. Up & Down 이 늘 있다는 것.
3. 연구자로서의 마음 가짐. (참고 글)

남은 질문에 대한 이야기

1. 대학원 진학
2. 기술 전망에 대한 이야기

감사합니다.

<https://github.com/kakao/recoteam>
nick.kim@kakaocorp.com

kakao