

네이버 AI 윤리 준칙의 방향성

2021. 2. 17.

NAVER 송대섭 이사

Agenda Research (정책연구실)

AI 윤리 준칙에 대한 고민

내용

AI 윤리 준칙에 어떤 내용을 담을 것인가?

방법

AI 윤리 준칙을 어떻게 만들 것인가?

실천

AI 윤리 준칙을 어떻게 실천할 것인가?

AI 윤리 준칙의 방향성

내용

AI 윤리 준칙에 AI에 대한 **사회의 요구**와 **네이버의 기업철학**을 반영

방법

AI 윤리 준칙을 **학과와 함께** 만들고, 이후에도 **학과와 지속적으로 협업**

실천

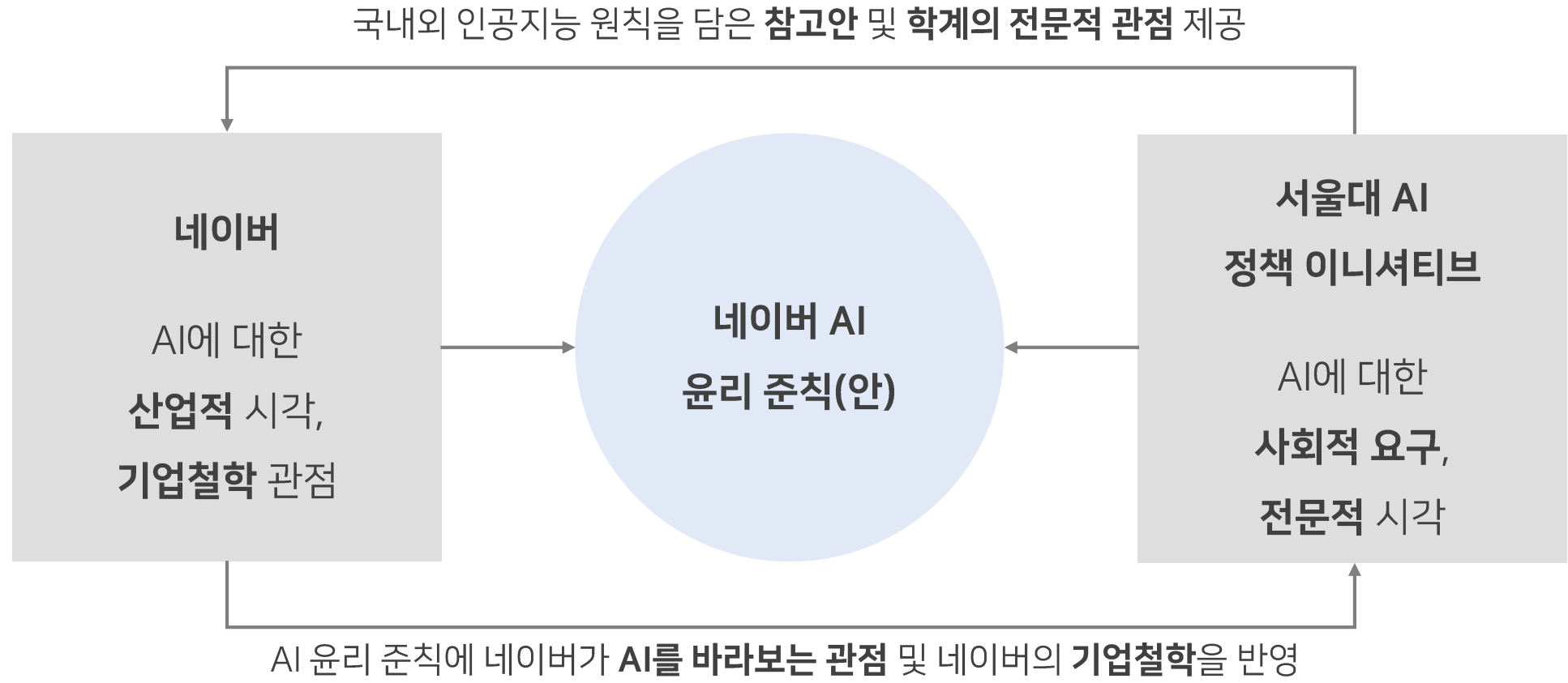
AI 윤리 준칙 적용 **사례를 통한 원칙의 구체화** 및 **지속적 개선**

서울대 AI 정책 이니셔티브(SAPI)*와 협업하여 초안 작성

- AI 윤리 준칙(안)은 2020년 2월 SAPI가 제안하였던 AI 원칙 참고안을 바탕으로, 네이버가 AI를 바라보는 관점을 담아 실천 가능한 방향으로 작성
- Agenda Research(정책연구실)는 SAPI와 지난 3년 동안 협업하면서 AI가 사회에 미치는 영향에 대해 살펴보고, 학계의 의견을 지속적으로 수렴해 왔음
 - 추상적인 논의보다는 실제 AI 개발 현실에서 사전적으로 고민해 볼 필요가 있는 AI의 효용과 사회적 의미에 대해 객관적인 시각으로 접근할 필요성이 있다고 판단
 - 다양한 사회의 목소리를 반영하기 위하여 인공지능 정책에 관한 학계 전문가분들께 자문 요청

*서울대 인공지능 정책 이니셔티브(SNU AI Policy Initiative): 법학, 공학, 경제학 등 여러 분야의 협력을 통해 인공지능과 관련한 융합 연구를 하는 프로그램으로 2017년에 시작됨. 서울대 법학전문대학원의 고학수(현 인공지능법학회 회장), 임용 교수님이 공동디렉터를 맡고 있으며, 매년 인공지능 정책에 관한 국제학술행사를 주최하고 있음

AI의 산업적 의미 뿐만 아니라 사회적 의미를 고민



초안 작성 이후 사내 검토 및 의견 수렴 절차 진행

SAPI와
초안 작성

사내
개발리더
검토

SAPI와
사내 간담회
진행

SAPI와
공동 웨비나
진행

AI 윤리 준칙의 구조

전문

조항 (5개항)

1. 사람을 위한 AI 개발
2. 다양성의 존중
3. 합리적인 설명과 편리성의 조화
4. 안전을 고려한 서비스 설계
5. 프라이버시 보호와 정보 보안

“**AI와 로봇 등과 같은 첨단 기술**을 누구나 쉽고 빠르게 활용할 수 있도록
사용자들의 손에 잡히는 **일상의 도구로 바꿔내는 것**이 네이버 사업의 본질입니다.”

- Connect 2020

“취향저격 콘텐츠, 숨은 보석 같은 가게, 1인 크리에이터 등 **작은 개인이 한계를
넘어 다양성의 가치를 인정**받고, 개개인의 성공들이 모여 우리 경제에 분수효과를
가져오기를 기대합니다.”

- 프로젝트 꽃

AI 윤리 준칙 전문

일상의 도구

네이버는 첨단 AI 기술을 누구나 쉽고 편리하게 활용할 수 있는 **일상의 도구**로 만들겠습니다.

연결, 도전, 다양성

사용자에게 새로운 **연결**의 경험을 선보이는 **도전**을 멈추지 **않음**으로써 **다양한** 기회와 가능성을 열어 나가겠습니다.

모든 구성원

이를 위해 네이버의 **모든 구성원**은 AI 개발과 이용에 있어 아래와 같은 윤리 원칙을 준수하겠습니다.

AI 윤리 준칙 1. 사람을 위한 AI 개발

인간 중심의 가치

편리한 일상의 도구

네이버가 개발하고 이용하는 AI는 사람을 위한 일상의 도구입니다. 네이버는 AI의 개발과 이용에 있어 **인간 중심의 가치를 최우선**으로 삼겠습니다.

네이버는 **사용자의 일상에 편리함**을 더하기 위해 기술을 개발해왔고, AI 역시 **일상의 도구로 활용**될 수 있도록 발전시켜 나가고 있습니다. 네이버는 AI가 우리의 삶을 편리하게 만들어줄 수 있는 기술이지만, 세상의 다른 모든 것처럼 완벽할 수 없다는 점을 인식하고 있습니다. 네이버는 **AI가 사람을 위한 일상의 도구**가 될 수 있도록, 지속적으로 살펴보며 개선해 나가겠습니다.

AI 윤리 준칙 2. 다양성의 존중

다양한 가치의 공존

부당한 차별 방지

네이버는 **다양성의 가치**를 고려하여 AI가 **사용자를 포함한 모든 사람에게 부당한 차별을 하지 않도록** 개발하고 이용하겠습니다.

네이버는 **다양성을 통해 연결이 더 큰 의미**를 가질 수 있도록 기술과 서비스를 구현해 왔습니다. 그 과정에서 사용자에게 다채로운 기회와 가능성을 열어왔고, **합리적 기준 없는 부당한 차별이 발생하지 않도록** 노력해 왔습니다. 네이버는 AI 서비스에서도 **부당한 차별을 방지하고 다양한 가치가 공존**하는 경험과 기회를 제공해 나가겠습니다.

AI 윤리 준칙 3. 합리적인 설명과 편리성의 조화

합리적 설명에 대한 책무

네이버는 누구나 편리하게 AI를 활용하도록 도우면서, 일상에서 AI의 관여가 있는 경우 사용자에게 그에 대한 합리적인 설명을 하기 위한 책무를 다하겠습니다.

네이버는 AI에 관한 합리적인 설명의 방식과 수준이 다양할 수 있다는 점을 고려해, 이를 구체적으로 실현하기 위하여 노력하겠습니다.

사용자의 눈높이에 맞춘 설명

네이버의 AI는 기술을 위한 기술이 아니며, **기술적 지식이 없이도 누구나 손쉽게 활용할 수 있는** 도구가 될 것입니다. 네이버는 서비스의 편리함을 추구하면서, 사용자의 요구가 있거나 필요한 경우에는 AI 서비스에 대해 쉽게 이해할 수 있도록 **사용자의 눈높이에 맞춰 설명**하겠습니다.

AI 윤리 준칙 4. 안전을 고려한 서비스 설계

사람의 생명과 신체 안전

네이버는 안전에 유의하여, 서비스의 전 과정에서 사람에게 유해한 영향을 미치지 않는 AI 서비스를 설계하겠습니다.

사람을 위한 일상의 도구인 AI가 **사람의 생명과 신체를 위협하는 상황이 발생하지 않도록**, 네이버는 전 과정에서 안전을 고려해 서비스를 설계하고, 테스트를 거치며, 배포 이후에도 **안전성**에 대해 지속적으로 살펴보겠습니다.

AI 윤리 준칙 5. 프라이버시 보호와 정보 보안

프라이버시 바이 디자인(PbD)

네이버는 AI를 개발하고 이용하는 과정에서 개인정보 보호에 대한 법적 책임과 의무를 넘어 **사용자의 프라이버시가 보호될 수 있도록** 노력하겠습니다. 또한 개발 단계를 포함해 **AI 서비스의 전 과정에서 정보 보안을 고려한 설계**를 적용하겠습니다.

네이버는 개인정보 활용에 있어 법적 책임과 의무를 다하는 것을 넘어 **개인의 프라이버시도 적극적으로 보호**하고 있습니다. 또한 사용자가 서비스를 활용하면서 정보 보안을 우려하게 되는 상황을 원천적으로 차단할 수 있도록, **서비스 전 과정에서 정보 보안을 고려한 설계를 적용**하고 있습니다. AI 서비스에 있어서도 마찬가지로, 사용자가 프라이버시와 정보 보안을 걱정하지 않고 AI 서비스를 자유롭게 활용해 삶에 편리함을 더할 수 있도록 노력하겠습니다.

네이버 AI 윤리 준칙

NAVER Artificial Intelligence Ethics Principles

네이버는 첨단 AI 기술을 누구나 쉽고 편리하게 활용할 수 있는 일상의 도구로 만들겠습니다. 사용자에게 새로운 연결의 경험을 선보이는 도전을 멈추지 않음으로써 다양한 기회와 가능성을 열어 나가겠습니다. 이를 위해 네이버의 모든 구성원은 AI 개발과 이용에 있어 아래와 같은 윤리 원칙을 준수하겠습니다.

1. 사람을 위한 AI 개발

네이버가 개발하고 이용하는 AI는 사람을 위한 일상의 도구입니다. 네이버는 AI의 개발과 이용에 있어 인간 중심의 가치를 최우선으로 삼겠습니다.

네이버는 사용자의 일상에 편리함을 더하기 위해 기술을 개발해왔고, AI 역시 일상의 도구로 활용될 수 있도록 발전시켜 나가고 있습니다. 네이버는 AI가 우리의 삶을 편리하게 만들어줄 수 있는 기술이지만, 세상의 다른 모든 것처럼 완벽할 수 없다는 점을 인식하고 있습니다. 네이버는 AI가 사람을 위한 일상의 도구가 될 수 있도록, 지속적으로 살펴보며 개선해 나가겠습니다.

2. 다양성의 존중

네이버는 다양성의 가치를 고려하여 AI가 사용자를 포함한 모든 사람에게 부당한 차별을 하지 않도록 개발하고 이용하겠습니다.

네이버는 다양성을 통해 연결이 더 큰 의미를 가질 수 있도록 기술과 서비스를 구현해 왔습니다. 그 과정에서 사용자에게 다채로운 기회와 가능성을 열어왔고, 합리적 기준 없는 부당한 차별이 발생하지 않도록 노력해 왔습니다. 네이버는 AI 서비스에서도 부당한 차별을 방지하고 다양한 가치가 공존하는 경험과 기회를 제공해 나가겠습니다.

3. 합리적인 설명과 편리성의 조화

네이버는 누구나 편리하게 AI를 활용하도록 도우면서, 일상에서 AI의 관여가 있는 경우 사용자에게 그에 대한 합리적인 설명을 하기 위한 책무를 다하겠습니다. 네이버는 AI에 관한 합리적인 설명의 방식과 수준이 다양할 수 있다는 점을 고려해, 이를 구체적으로 실현하기 위하여 노력하겠습니다.

네이버의 AI는 기술을 위한 기술이 아니며, 기술적 지식이 없어도 누구나 손쉽게 활용할 수 있는 도구가 될 것입니다. 네이버는 서비스의 편리함을 추구하면서, 사용자의 요구가 있거나 필요한 경우에는 AI 서비스에 대해 쉽게 이해할 수 있도록 사용자의 눈높이에 맞춰 설명하겠습니다.

4. 안전을 고려한 서비스 설계

네이버는 안전에 유의하여, 서비스의 전 과정에서 사람에게 유해한 영향을 미치지 않는 AI 서비스를 설계하겠습니다.

사람을 위한 일상의 도구인 AI가 사람의 생명과 신체를 위협하는 상황이 발생하지 않도록, 네이버는 전 과정에서 안전을 고려해 서비스를 설계하고, 테스트를 거치며, 배포 이후에도 안전성에 대해 지속적으로 살펴 보겠습니다.

5. 프라이버시 보호와 정보 보안

네이버는 AI를 개발하고 이용하는 과정에서 개인정보 보호에 대한 법적 책임과 의무를 넘어 사용자의 프라이버시가 보호될 수 있도록 노력하겠습니다. 또한 개발 단계를 포함해 AI 서비스의 전 과정에서 정보 보안을 고려한 설계를 적용하겠습니다.

네이버는 개인정보 활용에 있어 법적 책임과 의무를 다하는 것을 넘어 개인의 프라이버시도 적극적으로 보호하고 있습니다. 또한 사용자가 서비스를 활용하면서 정보 보안을 우려하게 되는 상황을 원천적으로 차단할 수 있도록, 서비스 전 과정에서 정보 보안을 고려한 설계를 적용하고 있습니다. AI 서비스에 있어서도 마찬가지로, 사용자가 프라이버시와 정보 보안을 걱정하지 않고 AI 서비스를 자유롭게 활용해 삶에 편리함을 더할 수 있도록 노력하겠습니다.

단계적인 여러 실험들을 통해 AI 윤리 준칙의 실천 방법 모색

사안 중심의 유연한 커뮤니케이션 채널 마련

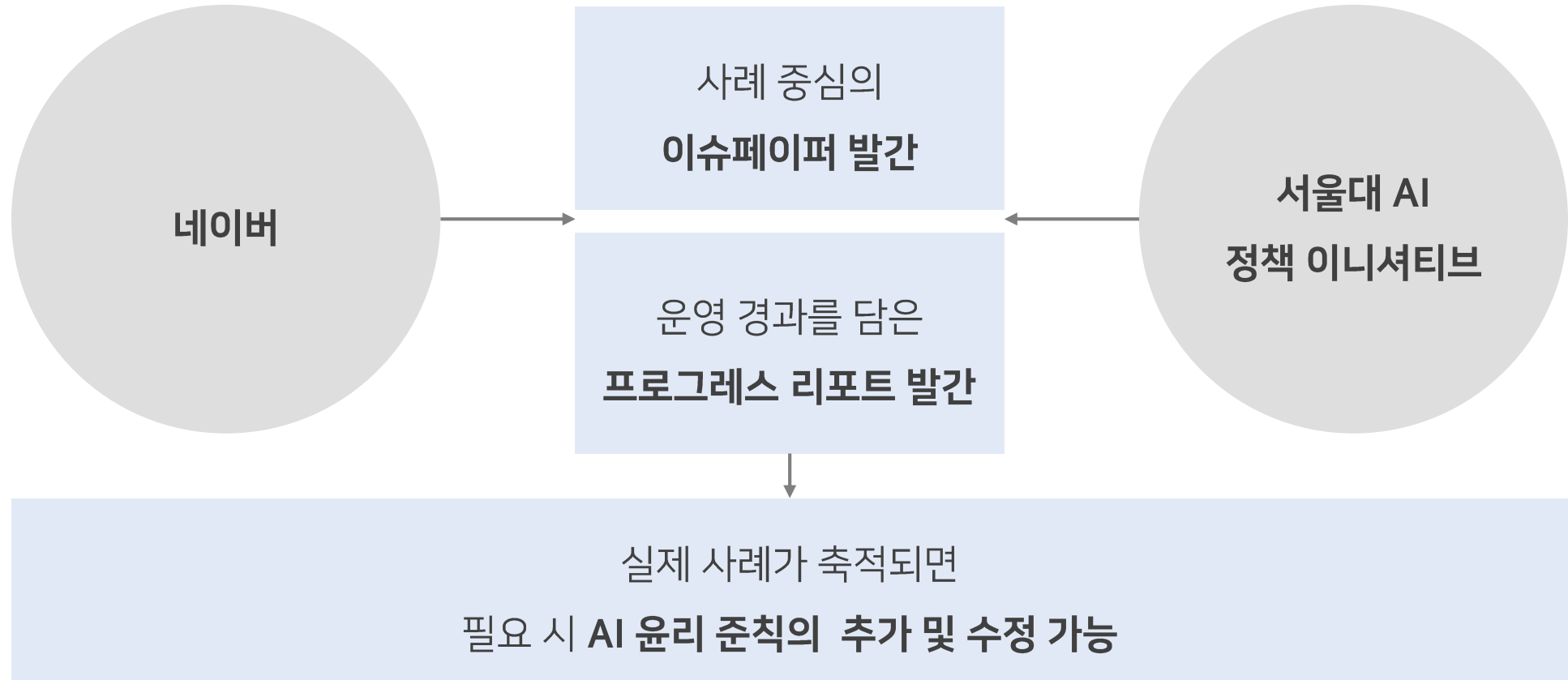
사안 중심의
커뮤니케이션
채널

프로젝트 진행 또는 서비스 개발 시
AI 윤리 준칙과 관련된 사안 감지 및 문의

사안 별 문의 사항에 대한 논의

AI 업무 관련성이 높으면서도
다양한 관점을 가진 직군이 참여하는 메일링 그룹

SAPI와 협력해 이슈페이퍼와 프로그레스 리포트 발간



AI 윤리 준칙 설계 및 실천에 대한 경험 등을 스타트업에게 공유

네이버

서울대 AI
정책 이니셔티브

스타트업
지원기관

스타트업

네이버가 AI 윤리 준칙의 설계 및 실천을 통해 얻은
경험, 고민, 노하우를
SAPI와 스타트업 지원기관을 통해
스타트업에게 공유

개별 스타트업 문화에 맞는
AI 윤리 원칙 설계,
독자적으로 판단하기 어려운 AI 이슈를
사전에 고민할 수 있는 계기 제공

감사합니다