

Role of Policy

개별 설계자가 제품 설계를 변경하려고 할 수도 있지만, 이익이 되지않으면, 우리는 실질적인 변화를 보지 못할 것이다.

The Effectiveness of Regulation

기업에 재정적으로 큰 부담을 주는 규제와 작은 규제의 효과 (facebook)

- 미얀마 로힝야족에 대한 혐오 발언 유포와 폭력을 부추기는 데 face북이 사용되고 있다는 지역활동가들의 경고 → 의회 청문회 → 2013년부터, 2017년 8월 이후 북부 라흐마인주에서 최소 288개 마을이 불에 타 파괴되는 등 대량학살이 시작된 지 몇 년이 지난 2018년에 10여명을 고용하여 문제를 해결하겠다고 함.
- 혐오표현에 반대하는 독일의 새 법 → 비싼 벌금(5000만 유로)을 피하려 페이스북이 독일에서 1200명을 신속하게 고용함.

전세계적인 합의가 필요한 환경규제의 효과

제1세계 선진국의 환경규제로 관할권 경계를 넘어 조정하고 때로는 고도의 기술적 규제가 필요했다. 어떤 경우에는 오존층을 고갈시킨 상업용 냉장고에 대한 금지와 같이, 그러한 규제는 전 세계적인 합의를 필요로 했다.

프라이버시법도 시각의 변화가 필요한 시점에 와 있다.

Rights and Policy

- 깨끗한 공기, 식수
- 기술을 오용함으로 인한 의도치 않은 결과 → 많은 해악 발생 → 오염된 정보 환경이나 악화된 주변 사생활

 **조정된 규제 조치가 필요한 공공재**

- 편향된 알고리즘이 흑인 피고인들에게 더 긴 징역형을 선고할 것을 권고
- 특정 직업 광고가 젊은이들에게만 노출
- 경찰이 시위자들을 식별하기 위해 안면 인식 사용

 **법 으로 해결**

 **개별 소프트웨어 개발자와 데이터 과학자가 실제로 윤리적 결정을 내릴 수 있는 장비를 갖추는 것이 중요하다**

개별 소프트웨어 개발자와 데이터 과학자가 실제로 윤리적 결정을 내릴 수 있는 장치의 예 :

- 자동차 안전성을 높이기 위한 움직임이 한 예
- Timnit Gebru 외 "데이터셋용 데이터시트 (<https://arxiv.org/abs/1803.09010>)
 - 머신러닝 커뮤니티는 현재 데이터셋을 문서화하는 표준화된 프로세스가 없으며, 이는 높은 인지도에서 심각한 결과를 초래할 수 있다. 이러한 격차를 해소하기 위해 데이터셋용 데이터시트를 제안
 - 모든 전자부품의 작동 특성, 시험 결과, 권장 용도, 기타 정보의 데이터 세트에 동기 부여를 문서화하는 데이터 시트 첨부
 - 구성, 수집 프로세스, 권장 용도 등 데이터셋을 위한 데이터시트는 데이터셋 작성자와 데이터셋 소비자 간의 커뮤니케이션을 원활하게 하고, 기계 학습 커뮤니티가 투명성과 책임성을 우선시하도록 장려
- 99% Invisible (<https://99percentinvisible.org/episode/nut-behind-wheel/>)
 - 안전을 고려하지 않은 초기 차량의 문제점 : 안전 벨트의 부재, 충돌 시 사람들의 두개골에 들어갈 수 있는 계기판의 금속 손잡이, 위험한 방법으로 부서진 일반 판유리창, 그리고 운전자를 찌그러뜨리는 접을 수 없는 핸들기둥
 - 이 문제점들에 대한 자동차 회사들의 토론 거부감. 자동차 사용자들의 잘못된 믿음.
 - 2011년에야 평균적인 남성의 신체가 아닌 평균적인 여성을 나타내는 충돌 테스트 인체모형을 사용. 그전에는 남성 모형만 사용. 여성들은 남성들에 비해 동일한 충격의 자동차 충돌로 인해 부상을 입을 가능성이 40% 더 높음.
 - 편견, 정책, 기술이 중대한 결과를 초래하는 방식을 생생하게 보여주는 사례

Conclusion

나는 내 역할의 단점이 무엇인지 최대한 명확히 하려고 노력하고, 단점이 해결될 수 있도록 정말 정확하게 진단하는 것이라고 본다. 그것은 힘든 작업이고, 훨씬 더 많은 사람들이 그것을 해야 한다.

- 전 프로퍼블리카 선임기자 줄리아 앙윈(Julia Angwin)은 2019년 인터뷰 -

논의된 많은 문제들(모델이 훈련된 것과 극적으로 다른 데이터 분포, 일단 전개된 모델에 대한 영향 피드백 루프 등)도, 수 익성 있는 거래 모델을 구축하는 데 있어 핵심 쟁점이다.

- 프레드 먼로는 헤지펀드 거래인 -

윤리에 대해 신중하게 생각하는 것도, 데이터 제품을 더 일반적으로 성공시키도록 신중하게 생각하는 데 도움이 될 수 있다.



[\[뉴스IT\(잇\)쥬\] '통신비 2만원'이 거기서 왜 나와... "카카오 들어와" 시끌](#)

뉴스1 | 11시간 전 | 네이버뉴스 | [🔗](#)

◇ "카카오 들어와" **윤영찬** 의원 문자에 불붙은 '포털뉴스 편집' 논란 **윤영찬** 의원이 국내 양대 포털사이트(네이버·**다음**) 중 하나인 **다음**(Daum)에 대한 뉴스 편집 개입을 시사한 발언으로 논란이 되면서 현 포털사이트의...



[네이버·카카오, 또 '뉴스 편집' 논란...AI 신뢰도 확보·외압 차단 '숙제'](#)

매일일보 | 11시간 전 | [🔗](#)

윤영찬 더불어민주당 의원이 지난 8일 오후 국회 본회의장에서 **카카오**톡 뉴스에 실린 국민 의힘 주호영... **다음**을 창업한 **이재웅** 전 쏘카 대표는 "국회의원이 마음에 안 드는 뉴스가 메인에 올라왔다고 바로 포털 담당자를..."



[\[서정욱의 전복후계\] 언론 장악·여론 조작, 빙산의 일각 윤영찬 의혹 전모 밝...](#)

데일리안 PICK | 15시간 전 | 네이버뉴스 | [🔗](#)

이 점에서 민주당 **윤영찬** 의원이 '**다음**'의 뉴스 배치에 불만을 품고 "카카오에 강력 항의해 주세요.... 또한 **다음** 창업자인 **이재웅** 전 쏘카 대표도 "시는 가치중립적으로 판단하지 않는다. 설계한 대로 판단할 가능성이..."



[\[지난주 산업계 주요이슈\]던파, 네오플 직원 직권남용 적발... 삼성전자, 3분기...](#)

청년일보 | 21시간 전 | [🔗](#)

◇ "AI 설계자 의도 반영"...**이재웅** "윤의원 항의, 카카오 태도 모두 잘못" 포털 사이트 '**다음**'을 창업한 **이재웅** 전 쏘카 대표가 '포털 압박 문자'를 보낸 **윤영찬** 더불어민주당 의원과 뉴스편집 인공지능(AI)에 책임을...



[네이버 출신 윤영찬은 왜 카카오를 부르려 했나... "포털 임직원도 김영란법 ...](#)

법률방송뉴스 | 3일 전 | [🔗](#)

일단 **윤영찬** 의원이 몸담았던 네이버와 "들어오라고 하세요"의 대상으로 지목된 **다음카카오** 모두 당혹스러워... 하지만 이와 관련 **다음**을 창업한 **이재웅** 전 쏘카 대표는 그제 자신의 페이스북에 "포털의 '시'가 했으니까..."



[\[뉴스하이킥\] 민주당, 언제까지 자유한국당 시절 얘기할 건가.. 누가 더 나쁜...](#)

MBC | 3일 전 | 네이버뉴스 | [🔗](#)

지난 8일 국회에서 민주당 **윤영찬** 의원이 보내 던 휴대전화 메시지 '**카카오** 들어오라 하세요'라는 내용이 논란되고 있습니다. 국민의힘 원내대표 연설이 포털사이트 **다음**에 메인화면에 뜨자 이런 반응을 보인 건데요....

Questionnaire

1. 윤리는 "정답"의 목록을 제공하는가?
2. 다른 배경을 가진 사람들과 함께 일하는 것이 윤리적인 질문을 고려할 때 어떻게 도움이 될 수 있을까?
3. 나치 독일에서 IBM의 역할은 무엇이었는가? 왜 그 회사는 그대로 참여했는가? 노동자들은 왜 참여했는가?
4. 폭스바겐 디젤 스캔들로 수감된 첫 번째 인물의 역할은 무엇이었을까.
5. 캘리포니아 주 사법경찰 관리들이 관리하는 폭력조직 혐의자들의 데이터베이스에는 어떤 문제가 있었는가?
6. 유튜브의 추천 알고리즘을 구글의 직원이 이 기능을 프로그래밍하지 않았음에도 불구하고, 왜 옷을 부분적으로 입은 아이들의 동영상
을 소아성애자들에게 추천했을까?
7. 측정기준의 중심성에 대한 문제점은 무엇인가?
8. 왜 아마존닷컴은 기술 미팅을 위한 추천 시스템에 성별을 포함하지 않았는가?
9. 수레쉬와 구트태그에 따르면 기계학습에서 편향의 6가지 유형은 무엇인가?
10. 미국의 역사적 인종 편향의 두 가지 예를 들어보자.
11. ImageNet의 대부분의 이미지 원본 위치는 어디인가?
12. "기계학습이 도덕적 해이와 오류를 자동화하는가"(<https://scholar.harvard.edu/files/sendhil/files/aer.p20171084.pdf>)라는 논문에서 축농증은 왜 뇌졸중을
예측하는 것으로 밝혀졌는가?
13. 표현편향이란 무엇인가?
14. 의사결정에 사용하는 기계와 사람은 어떻게 다른가?
15. 위조지폐는 "가짜 뉴스"와 같은가?
16. 왜 자동 생성 텍스트를 통한 비양식이 특히 중요한 문제인가?
17. 마크쿨라 센터에서 설명한 5가지 윤리적 렌즈는 무엇인가?
18. 데이터 윤리 문제를 해결하기 위한 적절한 도구는 어디인가?

Further Research:

1. "알고리즘이 당신의 건강을 해칠 때 일어나는 일"(<https://www.theverge.com/2018/3/21/17144260/healthcare-medicaid-algorithm-arkansas-cerebral-palsy>) 기사를 읽으세요. 앞으로 어떻게 이런 문제들을 피할 수 있을까?
2. 유튜브의 추천 시스템과 사회적 영향에 대해 자세히 알아보기 위한 연구. 권장 시스템에는 항상 부정적인 결과가 있는 피드백 루프가 있어야 한다고 생각하십니까? 구글이 그들을 피하기 위해 취할 수 있는 접근법은 무엇인가? 정부는?
3. "온라인 광고 전달 시 차별"(<https://arxiv.org/abs/1301.6822>) 문서를 읽어보십시오. 당신은 구글이 닥터에게 일어난 일에 대해 책임을 져야 한다고 생각하는가? 스위니? 적절한 대응책은 무엇인가?
4. 학제 전반에 걸친 팀이 어떻게 부정적인 결과를 피하도록 도울 수 있을까?
5. "기계 학습이 도덕적 해이와 오류를 자동화하는가"(<https://scholar.harvard.edu/files/sendhil/files/aer.p20171084.pdf>)를 읽으십시오. 본 논문에서 확인된 문제에 대처하기 위해 어떤 조치를 취해야 한다고 생각하십니까?
6. "AI 기반 위조를 어떻게 방지할 것인가?"(<https://hbr.org/2019/03/how-will-we-prevent-ai-based-forgery>) 기사를 읽으십시오. 에지오니의 제안된 접근법이 효과가 있을 거라고 생각하십니까? 왜?
7. 이 장의 "작업 중인 프로젝트 분석" 섹션을 완료하십시오.
8. 여러분의 팀이 더 다양해질 수 있을지 생각해 보십시오. 만약 그렇다면, 어떤 접근법이 도움이 될 수 있을까?

Deep Learning in Practice: That's a Wrap

- . 이 책의 웹사이트에서 추천된 GPU 주피터 서버 중 하나에 연결하십시오.
- . 첫 번째 노트북 직접 실행하십시오.
- . 첫 번째 노트북에서 찾은 이미지를 업로드한 다음, 다른 종류의 몇 가지 이미지를 사용해 보십시오.
- . 생각해 낸 이미지 검색 쿼리에 따라 사용자 자신의 데이터 집합을 수집하는 두 번째 노트북을 실행하십시오.
- . 어떤 종류의 데이터를 사용할 수 있는지, 어떤 종류의 문제가 발생할 수 있는지, 실제로 이러한 문제를 어떻게 완화할 수 있는지 등 딥러닝을 사용하여 프로젝트를 수행하는 방법에 대해 생각해 보십시오.

다음 세션 : 딥러닝이 어떻게 그리고 왜 효과가 있는지에 대해..