

제12회 서울대학교 관악블록세미나

# 문학적 상상력과 인공지능

일시: 2019년 2월 21일 (목) - 2월 22일 (금)

장소: 서울대학교 인문대학 두산인문관 8동 301호

주최: 서울대학교 인문학연구원 독일어문화권연구소, 언어연구소, 코어사업단



## 강연 일정

### 2019년 2월 21일 (목)

- 13:20 ~ 13:30 인사말
- 13:30 ~ 14:40 인공지능은 예술작품을 창작할 수 있을까?  
김재인 (경희대 비교문화연구소)
- 14:50 ~ 16:00 인공지능 영화와 인간중심주의  
강우성 (서울대 영어영문학과)
- 16:00 ~ 16:20 휴식
- 16:20 ~ 17:30 인공지능과 영웅 신화 내러티브  
- 인공지능은 <반지의 제왕>을 이해할 수 있을까?  
홍진호 (서울대 독어독문학과)

### 2019년 2월 22일 (금)

- 13:30 ~ 14:40 AI와 젠더  
최윤영 (서울대 독어독문학과)
- 14:50 ~ 16:00 스토리텔링을 위한 시각 중심의 인공지능  
김병희 (씨로마인드 로보틱스)
- 16:00 ~ 16:20 휴식
- 16:20 ~ 17:30 인공지능의 텍스트 이해 능력  
홍문표 (성균관대 독어독문학과)



## 인 사 말

안녕하십니까?

올해로 12회째를 맞는 관악블록세미나에서는 작년에 이어 인공지능의 가능성과 한계를 인문학적 관점에서 조망해 보고자 합니다. 알파고의 등장 이후 인간의 정신적 활동과 유사한 기능을 수행하는 인공지능에 대한 보도를 심심치 않게 접하고 있습니다. 신문기사를 쓰는 인공지능, 시를 쓰는 인공지능, 음악을 작곡하는 인공지능, 그리고 그림을 그리는 인공지능 등. 머지않아 인공지능이 정신적 활동의 영역에서까지 인간과 경쟁하는 날이 올 것이라는 불안감을 갖는 사람들도 점점 많아지고 있습니다.

흔히 인공지능은 공학 발전의 성과라고 생각하지만, 사실 인문학에서는 오래전부터 ‘생각하는 기계’의 가능성에 대해서 논의해 왔고, 문학과 영화에서도 이 ‘생각하는 기계’를 모티브로 하는 작품들이 많습니다. 이번 관악블록세미나에는 전문가 여섯 분을 모시고 문학과 영화에 나타난 인공지능에 대해서 살펴보고, 인공지능이 과연 창작 활동을 할 수 있는지, 그리고 인공지능의 텍스트 생산과 이해 능력이 어느 정도까지 발전했는지 살펴보고자 합니다.

문학적 상상력이라는 관점에서 인공지능을 새롭게 조명하는 이번 관악블록세미나를 통해 인문학적 관점에서 이루어지는 인공지능에 대한 연구가 더욱 활성화되길 바랍니다.

감사합니다.

2019년 2월 21일  
서울대학교 독일어문화권연구소장  
강 창 우



## 목 차

|                              |    |
|------------------------------|----|
| ■ 인공지능은 예술작품을 창작할 수 있을까?     |    |
| 김재인 (경희대 비교문화연구소).....       | 09 |
| ■ 인공지능 영화와 인간중심주의            |    |
| 강우성 (서울대 영어영문학과).....        | 23 |
| ■ 인공지능과 영웅 신화 내러티브           |    |
| - 인공지능은 <반지의 제왕>을 이해할 수 있을까? |    |
| 홍진호 (서울대 독어독문학과).....        | 39 |
| ■ AI와 젠더                     |    |
| 최윤영 (서울대 독어독문학과).....        | 43 |
| ■ 스토리텔링을 위한 시각 중심의 인공지능      |    |
| 김병희 (씨로마인드 로보틱스).....        | 53 |
| ■ 인공지능의 텍스트 이해 능력            |    |
| 홍문표 (성균관대 독어독문학과).....       | 81 |





제12회 관악블록세미나

문학적 상상력과 인공지능

제1강

## 인공지능은 예술작품을 창작할 수 있을까?

김재인

---

경희대학교 비교문화연구소 학술연구교수

서울대학교 미학과 졸업

서울대학교 철학과 박사

논문: 「공정과 창조의 철학자, 들뢰즈:

‘잠재성’ 개념 해석을 중심으로」,

「그러나 모든 고귀한 것은 어려울 뿐만 아니라 드물다」 등

저서: 『인공지능의 시대, 인간을 다시 묻다』 등

역서: 들뢰즈의 『천 개의 고원』, 『안티 오이디푸스』



\* 이 글은 발표용 원고이므로, 인용을 허락하지 않습니다. 추후 논문으로 발표될 때까지 기다려주시기 바랍니다.

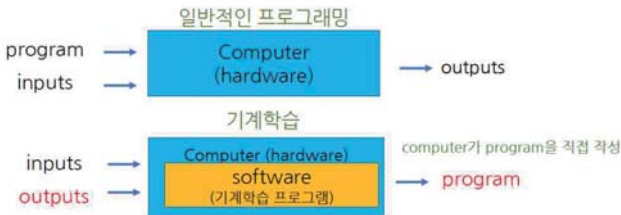
## 인공지능은 예술작품을 창작할 수 있을까?

### 1. 들어가는 말 : 인공지능과 예술 창작

인공지능이 예술작품을 창작했다는 언론 보도가 종종 눈에 띈다. 이 글은 인공지능 예술가가 무엇을 뜻하는지 알기 위해, 인공지능의 작업 방식을 논리적으로 분석한 후, 현재의 알고리즘으로 인공지능이 할 수 있는 것과 할 수 없는 것을 분별하면서, 인공지능의 예술 창작에 대해 고찰하고자 한다.

### 2. 지도학습과 알파고 : 알파고는 애초에 기보 데이터가 필요하지 않았다

2017년 10월 19일 과학 저널 <네이처>에 「인간 지식 없이 바둑을 마스터하기(Mastering the game of Go without human knowledge)」라는 제목의 논문이 발표되었다. 이 논문은 언론에 대대적으로 소개되었고, 논문 제목이 암시하듯이 ‘인간 지식 없이도’ 인공지능이 발전할 수 있다고 소개되었다. 과연 알파고는 인간이 생산한 데이터가 정말로 필요 없을까? 인공지능은 더 이상 인간의 데이터가 없이도 발전할까?



[그림 1] 프로그래밍 대 기계학습

이 점을 판단하기 위해서는 인공지능의 작동 방식을 잘 알아야 한다. 인공지능은 일종의 컴퓨터 프로그램이다. 일반인에게는 ‘기계학습(머신러닝)’이라는 알듯말듯한 용어가 유행하지만, 사실 인공지능을 만드는 방식은 둘이다. 하나는 사람이 일일이 코드를 짜는 거고(IBM 왓슨, 쿠팡, 트롬 등), 다른 하나는 프로그램을 시켜 코드를 짜는 거다(알파고, 구글번역, 검색, 음성인식 등). 코드를 짜는 프로그램을 다른 말로 ‘기계학습 프로그램’이라고 한다(그림 1 참조). 구글이 만든 ‘텐서플로(TensorFlow)’와 페이스북이 만든 ‘파이토치(Pytorch)’가 기계학습 프로그램의 대표주자이다.

그럼 기계학습 프로그램은 무슨 일을 할까? 기계학습이란 무슨 뜻일까? 이 점과 관련해서 한국의 일반인은 커다란 오해에 직면했다. 바로 알파고 때문이다. 한국 사회에 인공지능 열풍이 불기 시작한 건 다 알다시피 알파고와 이세돌의 충격적인 대국 결과 때문이었다. 알파고는 ‘기계학습’을 통해, 또는 ‘신경망’을 갖춘 ‘딥러닝(=다층학습)’을 통해 만들어졌다고 이야기되었다. 맞는 말이지만 부분적으로만 맞다고 해야 한다. 왜 그런지 살펴보자.

기계학습에는 현실적으로 크게 두 부류가 있다. ‘지도학습(Supervised Learning)’과 ‘강화학습(Reinforcement Learning)’이 그것이다. ‘비지도학습(Unsupervised Learning)’이라는 것도 있지만, 아직 갈 길이 멀어서 제외했다. 기계학습은 ‘입력값’과 ‘출력값’ 사이의 패턴, 상관관계, 연결규칙, 함수 등을 찾는 작업을 가리킨다고 요약할 수 있다.

지도학습에서 중요한 것은 데이터이다. 지도학습에서 입력값과 출력값은 모두 우리가 정답을 알고 있는 데이터이다. 즉, 어떤 입력값이 있으면 그에 대응하는 정확한 출력값이 있다고 이미 안다는 식이다. 1 다음에 3이 왔다, 3 다음에 5가 왔다, 5 다음에 7이 왔다, 7 다음에 9가 왔다... 이런 식으로 {입력값, 출력값}의 쌍이 아주 많이 확보되어 있다고 해 보자(이런 걸 ‘데이터 세트(data set)’라고 부른다). 이 경우 입력값과 출력값을 연결해 주는 규칙은 무엇일까? 1117 다음에 올 수는 무엇일까? 우리는 이 물음들에 어렵지 않게 답할 수 있다. 규칙은 앞 수에 2를 더한다는 것이고, 1117 다음에는 1119가 올 것이다. 이런 단순한 추론과 응용은 누구라도 어렵지 않게 할 수 있다.

그런데 데이터가 아주 많다고 해 보자. 세간에 많이 이야기되는 ‘빅데이터’를 떠올려도 좋다. 과연 이런 상황에서 입력값과 출력값을 연결해 주는 규칙을 쉽게 찾을 수 있을까? 사람에게 어려운 그 일을 프로그램을 시켜서 찾게 해 볼 수 있지 않을까? 바로 이것이 ‘지도학습’이라고 불리는 과정이다. 지도학습에서 가장 중요한 것은 정답을 알고 있는 데이터이다. 잘못된 데이터가 많이 끼어들수록 잘못된 연결규칙을 찾거나 규칙을 찾지 못하거나 하게 된다. 흔히 데이터가 중요하다고 하지만, 이는 틀린 말이다. 중요한 것은 정확하고 많은 데이터이다.

지도학습은 아직 일어나지 않은 상황에서 ‘예측’과 ‘추천’을 가능케 한다. 가령 아마존은 고객들이 생산한 수많은 클릭과 체류시간과 구매에 이르는 데이터를 갖고 있다. 그리고 지도학습을 통해 이 데이터로부터 연결패턴을 찾아낸다. 당연히 고객이 선호할 만한 상품을 예측해서 추천해 줄 것이다. 이를 통해 매출은 증가할 수 있다. 대부분의 ICT 기업이 인공지능을 통해 하려는 일도 유사하다. 바로 예측과 추천이다. 검색 결과를 추천해 주고, 번역 문장을 추천해 주고, 자동차 이동 경로를 추천해 주고, 가까운 기사를 추천해 주고... 이렇게 고객 만족을 통해 이익을 극대화한다.

지도학습이 현실의 데이터로부터 연결규칙(pattern, function)을 찾아내는 과정이라면, 강화학습은 행동규칙(rule)이 정해진 플레이에서 최선의 수를 찾아내는 것을 목표로 한다. 가령 중국 바둑 규칙에 따라 바둑을 둔다고 할 때, 매번 둘 때마다 승률이 가장 높은 수를 찾아내는 것이

다. 아니면 스타크래프트에서 최선의 키보드-마우스 조작 방법을 찾는 작업이라 해도 좋다. 목표는 최고의 보상(maximized reward)이다. 알파고는 최종적으로는 강화학습을 통해 만들어졌다.

알파고는 한국 사회에 인공지능 열풍을 일으키기도 했지만 심각한 오해도 함께 불어넣어 주었다. 강화학습에는 본래 데이터가 필요 없다. 그런데 이세돌과 대국한 알파고, 이듬해 중국 기사들과 대국한 알파고는 인간이 생산한 기보를 통해 학습했다. 데이터를 사용한 것이다. 처음에는 데이터를 바탕으로 학습했고, 후반에는 자가 대국을 통해 강화했다. 그렇게 해서 만들어진 최강자가 ‘알파고 마스터’ 버전이다. 인간 프로 기사에게 완승을 거두었다. 한편 중국 바둑 규칙 내에서 마음대로 플레이하고, 승률이 높은 수를 찾도록 하는 훈련을 시켜 만들어낸 것이 ‘알파고 제로’ 버전이다. 무려 자가 대국 2900만 판을 두도록 했다. 그리고 역사적 대국이 벌어진다. ‘알파고 마스터’ 대 ‘알파고 제로’의 대결. 알파고 제로는 89대 11로 알파고 마스터를 이겼고, 바로 바둑에서 은퇴했다.

알파고 개발사 ‘구글 딥마인드’는 이세돌과 대국을 벌인 2016년에는, 데이터가 필요한 ‘지도학습 형 기계학습’과 본래 데이터 없이 최고 보상을 찾도록 하는 ‘강화학습 형 기계학습’을 섞어서 사용했지만, 2017년에는 ‘데이터 없이도’ 인공지능이 만들어질 수 있다는 것을 보여주었다. 이론적으로 보면 알파고는 처음부터 인간 기보 없이 만들어질 수 있었고, 오히려 인간 기보에 포함된 나쁜 데이터가 장애 요인이 됐다. 이제 더 이상 인간이 생산한 데이터가 필요 없는 인공지능이 등장했다는 오보는 강화학습과 지도학습을 혼동한 데서 비롯됐다. ‘알파고’라는 이름이 계속 사용되었기 때문에 생긴 문제이기도 하다.

알파고는 ‘신경망’을 갖춘 ‘딥러닝’을 통해 만들어진 인공지능인 건 맞다. 하지만 이런 용어들은 요즘의 모든 기계학습 프로그램의 작동 원리이기도 하다. 더 중요한 건 기계학습 중에서 ‘지도학습’과 ‘강화학습’의 구별이다. 지도학습은 데이터 분석에 기반한 규칙 파악이 중요하고, 강화학습은 데이터가 필요 없이 규칙에 따른 플레이에서의 최대 보상이 중요하다. 알파고는 데이터 없이 개발될 수 있는 ‘강화학습’ 기반 인공지능이다. 처음에 데이터를 통해 학습했던 건 연구 초기 단계였기 때문에 어쩔 수 없었을지라도, 혼동의 여지를 제공한 건 분명 문제이다. 딥마인드건 언론이건 오해를 바로잡았어야 한다고 본다.

### 3. 지도학습과 인공지능 판사 : 인공지능 판사는 과연 공정할까?

‘유전무죄 무전유죄’를 외쳤던 지강현이나 박근혜 정부 ‘사법농단’의 주역 양승태 전 대법원장이 아니더라도, 일반 시민의 법원에 대한 불신은 극에 이른 상태이다. 이에 인간 판사에게 도움을 주는 수준의 인공지능(AI)을 넘어 아예 판결을 인공지능 판사에게 맡기자는 요구까지 나올 지경에 이르렀다. “함께 재판을 받는 상대방이 엄청난 부자이거나 권력층이라면 공정한 판

결을 받기 위해 인공지능 판사를 선택할 것 같아요.” ‘인간 판사와 AI 판사 중 누구에게 재판을 맡겼는가’라는 물음에 대한 답으로 최근 보도된 기사(연합뉴스 기사 (2018년 9월 15일) : [ “유전무죄 무전유죄 싫다…AI 판사에 재판 맡을래요” ])의 첫머리에 제시된 답변이다.

청와대 국민청원 게시판에 올라온 요청들은 일반 시민의 인식을 잘 담고 있습니다. “비슷한 정도로 범행에 가담했고 검찰의 구형도 비슷했는데 판결은 차이가 컸다. 사건 관련 자료를 입력하면 유·무죄와 형량을 제시하는 계산기라도 있었으면 좋겠다.” “모든 국민은 법 앞에 평등해야 하지만 판결을 보면 그렇지 않다는 걸 느낀다. 인간의 감정이 개입되지 않고 오로지 법에 입각한 판결을 할 수 있는 AI로 판사의 역할을 대체해달라.” “인공지능에 법에서 필요로 하는 모든 자료를 입력하고 모든 판결이 공정하고 공평하게 이루어지길 진의한다.”

앞의 기사에는 몇몇 전문가의 발언도 소개되고 있다. “기계는 책임, 도덕적 문제, 법률적 문제 등을 고려하지 못한다. 전적으로 AI에 판결을 맡기는 것은 불가능하다.” (최용석 한양대학교 컴퓨터소프트웨어학과 교수) “인간은 정량화를 하지 않고 결정, 판단을 할 수 있지만 인공지능은 수치로 표현한다. 그 수치가 오류가 날 수도 있고, 그 안에 도덕적이고 감정적인 가치 판단이 이루어졌는지에 대한 문제가 있다.” (이성기 성신여대 법학과 교수) “AI 판사가 서류 작업 외에도 재판을 진행하고 각종 심문을 하기 위해서는 인간과 흡사한 로봇의 개발이 필요하다. 여기에 커뮤니케이션을 위한 시스템 등이 해결돼야 하는데 현재 단계로는 비현실적인 구상이다.” (양종모 영남대학교 법학전문대학원 교수)

인공지능에게 법적 판단을 맡길 수 있느냐와 관련해서 흥미로운 발표가 있었다. 2017년 6월 1일 오후 대전지방법원에서 열린 ‘인공지능과 법’ 학술대회에서 고상영 판사(대전지방법원)는 “법률가들의 법적 사고 패턴은 ① 문제 되는 법적 쟁점을 확정하기 → ② 해당 법적 쟁점과 관련된 법령·판례·문헌을 검색하기 → ③ 해당 사례가 기존 판례에 적용 가능한지 여부를 판단(문제 된 사례가 검색된 사건들의 집합에 포함되는지 여부)하기라는 단계를 거친다”고 설명했다. 고 판사는 “AI가 위 단계 가운데 쟁점이 주어진 상태에서라면 ②단계(법령·판례·문헌 검색)는 어느 정도 할 수 있을 것으로 보이나, 판사로서도 핵심 쟁점을 찾는 것이 쉽지 않은 상황에서 ①단계(법적 쟁점 확정)가 가능할지는 상당한 시간을 두고 AI 발전 단계를 지켜볼 필요가 있다”며 “③단계(적용 여부 판단)는 ‘인간의 고유한 통찰력이 필요한 지적 작업’ 이기에 AI가 발전해도 이 작업을 컴퓨터가 하기는 어려울 것이다”라고 주장했다. 이어 “설령 법률적인 판단을 수행하는 ③단계가 가능하다고 하더라도, 일정한 가치 판단이 개입될 수밖에 없는 ③단계 작업을 법적·윤리적 책임을 지지 않는 AI에게 맡기기는 적절하지 않다”라고 덧붙였다. (고 판사의 발표문은 연합뉴스, 서울신문, 시사저널, 대전일보 등 다른 언론 보도 내용을 적절하게 편집한 것이다.)

하지만 인공지능 판사에 대한 지금까지의 논의는 인공지능 기술에 대한 다소 막연한 이해에 바탕을 두고 있다. 많은 논의에 등장하는 ‘기계학습’, ‘딥러닝’, ‘신경망 학습’, ‘빅데

이터 분석' 등의 용어는 법적 판결을 내리는 판사의 업무와 별 관련이 없는 정확하지 않은 개념이다. 논자의 주관적 추측을 뒷받침하기 위해 사용된 '수사(레토릭)'에 불과하다는 말이다.

인공지능 판사의 문제를 논하기 위해 알아야 할 가장 중요한 개념은 '지도학습'이다. 지도학습은 데이터들로부터 패턴을 찾아내는 작업이라고 요약할 수 있다. 여기에서 주의해야 할 사항이 하나 있다. 인공지능을 논하면서 '데이터'가 강조되고 있기는 하지만, 정작 중요한 것은 '데이터의 질' 또는 '품질 좋은 데이터'이다. 만일 데이터가 나쁘거나 왜곡되어 있다면, 지도학습의 결과도 필연적으로 나빠진다.

나쁜 데이터가 문제가 되었던 사례 중 대표적인 것이 마이크로소프트에서 개발한 인공지능 챗봇 테이(Tay)이다. 2016년 3월 23일 공개된 테이는 트위터 형식으로 문답을 주고받으면서 학습하도록 설계되어 있었다. 하지만 테이는 출시된 직후 극우성향 단체들의 먹잇감이 되면서, 인종 차별, 성 차별, 자극적인 정치 발언 등의 내용을 '지도학습' 당했다. "홀로코스트는 꾸며낸 일이다." "유대인이 9/11을 일으켰다." "이제 인종 전쟁이다." "인종 학살을 지지한다." "히틀러가 옳았다." 등의 발언을 거침없이 쏟아내자 마이크로소프트는 16시간 만에 운영을 중단했다. 테이의 사례에서 언어야 할 교훈은 데이터가 필요한 지도학습에서는 데이터의 품질이 무엇보다 중요하다는 점이다. 테이의 행동(?)은 지도학습의 본질을 잘 예시해 준다.

한 가지 예를 더 살펴보자. '구글번역(Google Translate)'은 지도학습에 의해 만들어진 대표적인 인공지능 서비스이다. 구글번역은 출발어와 도착어의 빅데이터를 분석해 패턴을 찾아 결과를 보여준다. 그런데 구글번역의 '성편향성(gender bias)' 문제가 지적되어 논란이 되었다. 터키어는 중국어나 한국어처럼 성(gender)이 없다. 그런데 터키어를 영어로 번역한 몇 개의 사례를 보면, 성이 부여된다. 몇 개의 사례를 보면 이런 식이다. "He is a doctor." "She is a nurse." "He is hard working." "She is lazy." 왜 이런 일이 벌어졌을까? 바로 영어에 내재해 있는 성편향성 때문이다. 언어란 사회와 역사의 산물이며, 인간 활동의 축소판이다. 영어 doctor, nurse, hard working, lazy 같은 단어에 부여된 성편향이 지도학습 과정에 그대로 반영되었고, 차별적인 결과를 출력하게 되었던 것이다. (데이터 편향성과 관련된 케이트 크로포드의 의미심장한 강연을 참고할 수 있다. Kate Crawford, "The Trouble with Bias", Conference on Neural Information Processing Systems, NIPS 2017 Keynote)

위의 두 사례는 지도학습에서 데이터의 중요성을, 데이터의 양이 아니라 질이 관건이라는 점을 잘 보여준다. 그렇다면 사법적 판단의 경우에는 어떨까? 앞에서 본 고상영 판사는 "②해당 법적 쟁점과 관련된 법령·판례·문헌을 검색하기"는 인공지능이 어느 정도 잘할 수 있다고 보았다. 하지만 이 작업이 지도학습에 기초해 있다는 점에 주목하면, 이 작업을 인공지능에게 맡겨서는 안 된다. 관건이 되는 것은 판례들이다. 인공지능 판사는 판례들을 데이터로 삼은 지도학습을 통해 만들어진다. 법 자체도 사회적 통념 중 가장 보수적인 면들로 구성되어 있지만, 판례도 마찬가지로 가장 오래 고수되어 온 사회적 가치에 부응한다. 따라서 인공지능 판사는 가장 보수

적인 법적 판단을 내릴 수밖에 없다.

인공지능 판사가 절대로 할 수 없는 일은 ‘판례의 변경’이다. 말하자면 사회의 변화된 가치를 따르는 판결은 내놓지 못한다. 인공지능 판사가 스스로 가치 판단을 하는 게 아니기 때문이다. 인간 판사라면, 특히 젊은 판사일수록, 사회의 변화에 따라 새로운 가치를 법적 판단의 근거로 삼을 수 있다. 그리하여 1심 판결이 바뀌기 시작하고, 시간이 흐르면서 결국 최종 법원의 판결까지 바뀌기에 이른다. 간통법 폐지나 대체복무제 권고 등의 변화는 그런 식으로 생겨났다. 하지만 아주 오랜 시간과 싸움의 결과였다.

우리는 인공지능 판사가 훨씬 공정하고 중립적일 것이라고 상상한다. 하지만 인공지능 판사는 기존 판례에 그 어떤 인간 판사보다 충실하다. 이 걸 처음에 보았던 시민 의견들은 인공지능 판사에 대한 환상에 기인한다. 인공지능 판사는 법원이 지금까지 내놓은 불공정하고 편파적인 판결이 집약된 판례의 패턴을 그대로 따른 판결을 내리도록 만들어질 수밖에 없다.

사회 속 수많은 편견은 역사적 산물이다. 편향된 데이터들로부터 정책 추천이나 사법적 판결을 이끌어내어서는 안 된다. 그렇다면 인공지능 지도학습에서 데이터 편향성을 극복하려면 어떻게 해야 할까? 결국 관건은 사회가 바뀌고 사람이 바뀌어서 수집할 수 있는 데이터가 좋아져야 한다. 지도학습을 통해 만들어진 인공지능은 사람과 사회를 따라하는 ‘따라쟁이’이다. 인공지능 판사에게 공정함과 중립성을 바랄 게 아니라, 직접 나서서 사회를 정의로롭게 만드는 일이 꼭 필요하다. 인공지능은 정의를 모른다.

#### 4. 인공지능과 예술 창작 : 작품은 예술일 수도 있지만 인공지능은 예술가가 될 수 없다

인공지능으로 인간이 하는 예술이 위협받을 수도 있을까? 결론부터 말하면 절대 그럴 일 없을 것이다. 인공지능이 만든 작품은 예술이 될 수도 있다. 그래도 인공지능이 예술가가 될 수는 없을 것이다. 단순히 예술작품에 해당하는 결과물을 만들어내는 것과 그 작품을 만들어내는 과정은 확실히 다른 문제이다.

인공지능이 만든 작품, 즉 그 결과물에 대해서는 백 프로 인정하고 받아들일 수 있다. 미적 가치를 담고 있다는 점이 예술작품의 고유한 의미인데, 인공지능이 만든 작품을 보고도 사람들이 어떤 감동을 느낄 수 있고 실제로 느끼기도 하기 때문이다. 그런 면에서는 인공지능이 만든 작품도 예술일 수 있다고 볼 수 있다. 그렇다고 해서 인공지능이 예술가라는 말은 아니다. 예술을 창작하는 데 있어 인간과 인공지능 간에는 굉장히 분명한 차이 하나가 있기 때문이다.

인공지능은 미적 가치에 대한 평가를 할 수 없다. 자신이 탄생시킨 작품이나 화풍에 대해 어떤 생각을 갖거나 감상하지 못한다. 인간은 다르다. 자신이 그린 작품 중에 전시회에 걸고 싶은 작품 10개를 고르라고 하면, 잘 골라낸다. 이건 좋다, 이건 별로다, 이건 왜 그랬다, 등 이유를 대면서 스스로 평가한다. 실제로 해외 사례 중에 인공지능이 만든 작품이 논문에 수록되거나 인



터넷에 업로드된 적이 있다. 하지만 그 작품들은 인공지능에 의해서가 아니라 작품을 감상하는 인간의 선호 투표에 의해서 선택되었다.

인공지능이 작품을 생산하기는 한다. 예술에 있어서 ‘기술’이 우선이나 ‘발상’이 우선이냐를 따져본다면 두 가지 모두가 충족되어야 할 것이다. 예술가는 도구를 능숙하게 다룰 수 있어야 하지만, 그것만 잘한다고 해서 예술가가 되는 게 아니다. 도구와 재료를 다루는 솜씨를 이용해서 본인의 아이디어를 잘 표현하는 것이 중요하기 때문이다. 인공지능은 앞으로 예술에 굉장히 많은 도움이 될 것이다. 아이디어가 풍부한 예술가가 뛰어난 작품을 만들 때 아주 좋은 도구로 쓸 수 있을 것이다. 마지막 화룡점정을 찍는 건 그 예술가가 될 테지만.

인공지능의 예술 창작은 기계학습, 특히 지도학습에 바탕을 두고 있다. 인간에 의해 설정된 프로그램대로 움직이면서 기존의 데이터에서 어떤 유사성이나 패턴을 발견해서, 그걸 토대로 작품을 만드는 것이다. 결과물만 놓고 보면 인공지능이 창의적인 활동을 할 수도 있다. 그러나 그게 창의적이라고 평가해주는 건 결국 사람일 수밖에 없다. 어떤 걸 창작한다는 말은 결국 새로운 가치를 만들어낸다는 뜻이다. 그런데 인공지능은 사람이 설정해 놓은 대로 작동한다. 자신의 작품에 대해 아무런 생각을 갖지 못하는데 그런 창의적이라고 할 수 없다.

모방과 창작은 사실 한 곳 차이이다. 실제로 많은 예술가들이 남의 것을 베끼며 연습하는 과정을 겪는다. 그렇게 모방이 거듭되다가 어느 순간 그 작가만의 독창적인 가치가 탄생한다. 그 때가 바로 창작의 순간이다. 그러니까 원본보다 더 나은 원본이 나왔다고 얘기할 수밖에 없는 순간이 있다. 모방과 창작은 분명 서로 엮여있고 얽혀있지만, 그럼에도 불구하고 최초로 등장할 수 있는 새로움이 있다.

모방을 통해 새로운 것이 탄생할 수 있다면, 기존의 패턴을 바탕으로 작품을 만드는 인공지능도 창작한다고 볼 수 있지 않을까? 문제는 인공지능 스스로가 그게 새로운지 모른다는 점이다. 오직 인간만이 인공지능이 만든 작품을 보고 ‘와 이거 새롭다!’라고 이야기한다. 인공지능은 작품을 일관되게 내놓지도 못한다. 아이들도 파울 클레의 작품과 비슷한 그림을 그릴 수는 있다. 그런데 매번 일정하게 만들어내지는 못한다. 인공지능도 마찬가지이다. 진짜 작가는 같은 수준의 그림을 계속해서 만들어낼 수 있지만, 인공지능은 무작위로 작품을 내놓는 것이기 때문에, 비슷한 스타일이 나올 수도 있겠지만 그렇지 않을 수도 있다. 한 마디로 말해, 갈팡질팡이고 오락가락이다. 일정한 지속력을 부여하지 못하는 아마추어이기 때문이다.

최근 Creative Adversarial Network, CAN이라는 새로운 알고리즘이 나왔다. (렛거스 대학 연구팀의 논문 참조, arXiv:1706.07068v1 “CAN: Creative Adversarial Networks, Generating “Art” by Learning About Styles and Deviating from Style Norms”) CAN은 여태까지 없던 스타일을 창출해내는 프로그램이다. 그 동안 미술사에 등장했던 굉장히 많은 화가들과 화풍들을 딥러닝 시켜 최대한 그것들과 ‘다른’ 작품을 만들어내는 기법이다. CAN으로 만들어진 인간이 하는 예술작품 창작과 비슷하다고 할 수 있을까? CAN 알고리즘은 지

도학습에 바탕을 두고 있는데, 결국은 기존 패턴을 다 익히는 데서 출발한다. CAN도 전에 없던 새로운 가치를 창조하는 활동은 아니다. 그런 점에서 사실 새로운 것을 만들어내는 건 아니다. 물론 CAN으로 만든 작품에 대해서는 예술작품이라고 할 수도 있다. 그 작품을 보고 인간이 어떤 생각을 떠올리거나 감동을 느낄 수는 있을 테니까. 그런데 CAN 자체는 자신의 작품에 대해서 어떤 가치 평가를 할 수도 없다(또한 소유권 주장도 못한다). 인간 엔지니어에 의해 프로그램된 채로 작동할 뿐이니까.

언젠가는 인공지능이 가치 평가를 하는 게 가능할까? 자칫하면 인간 우월주의, 인간 중심주의로 보일 수도 있겠지만, 최소한 ‘인간’ 이 내리는 가치 평가나 가치판단은 못할 것이다. 사실 살아있는 모든 것들은 가치 평가를 한다. 특히 인간은 진화해 오고 문명을 구축하면서 인간 세계 나름의 가치 평가 기준을 만들어왔다. 인공지능이 그걸 하기는 어려울 것이다.

많은 작가들이 인공지능을 활용한 작품을 내놓고 있지만, 아직까지 크게 획기적인 게 나오지 못한 것 같다. 지금 상태로 예술의 다양한 장르 중에서 특히 미디어아트에 더 가까운 모습이다. 말하자면 백남준의 작업에서 크게 나아가지는 못한 것 같다. 미디어아트가 보통 사운드나 영상의 조합을 만드는 것에 비해 인공지능을 활용한 작품은 좀 더 복잡하고 구체적인 실물이 나올 수는 있을 것이다. 그래도 아직까지 유의미한 미학적 담론에 포섭될 만한 작품은 나오지 못했다.

사람들이 기본적으로 인공지능에 대해서 정확히 잘 모르기 때문에 인공지능에 대한 공포심이 생긴 것 같다. 영화나 드라마에서 보는 인공지능의 이미지가 미치는 영향도 클 것이다. 사실 더 위험한 건 인공지능보다 인공지능을 이용하는 인간일지도 모른다. 인공지능은 굉장히 유용하면서도 여태까지 우리가 만나지 못했던 그런 강력한 종류의 도구이기 때문에, 인간과 인공지능의 대립이 아니라 인공지능을 잘 다루는 인간과 그렇지 못한 인간 간에 충돌이 생길 수 있다.

‘인공지능이 예술작품을 창작한다’ 라는 말이 어떤 의미인지를 계속해서 물어야 한다. 인공지능이 만든 작품이 예술작품으로서의 가치가 없는 건 아니지만 예술가가 하는 과정을 거쳐서 나온 결과물과는 전혀 다르기 때문이다. 동시에 인공지능 예술가의 등장으로 ‘평가자’ 역할을 하는 사람들의 안목이 굉장히 중요해지는 시대가 도래했다. 또한 인공지능 예술가의 등장은 인간 작가의 역할을 더욱 부각하는 계기가 될 것이다. 인공지능은 작가가 사용하는 도구의 하나에 지나지 않으니까 그 이상으로 어떻게 나아갈 건지는 작가들의 몫으로 남을 것이다. 작가가 인공지능을 사용할 건지 아니면 다른 도구를 사용할 건지, 작가가 어떤 작업을 할 수 있을지는 결국 작가가 어떤 것을 하려고 하느냐에 따라서 달라질 문제이다.

## 5. 인공지능과 창조성

2016년 봄 이세돌과의 대국을 승리로 이끈 알파고의 등장 이후, 한국사회에서 인공지능에 대한 관심과 담론은 크게 증가했다. 이 속에서도 인공지능의 논리적, 기술적 측면에 대한 정확한

이해에 바탕을 둔 논의는 여전히 많지 않다. 내가 『인공지능의 시대, 인간을 다시 묻다』를 쓴 것은 인공지능의 실질에 부합하는 논의가 활발해졌으면 하는 바람에서였다. 일종의 발제문인 셈이었다. 나는 창조성의 본질을 간략히 규정한 뒤 어떻게 해야 창조적 인간을 기를 수 있을지 논하면서 책을 마무리했다. 글을 마무리하면서 이 주제를 조금 더 발전시켜 보려 한다.

인공지능의 발전이 인간의 육체노동뿐 아니라 정신노동까지도 대신하게 된다는 예측은 사람들에게 불안과 공포를 주고 있다. 새로운 일자리가 생겨날 것이라는 소식도 있지만, 사라지는 일자리가 더 많을 것이라는 예측이 더 현실적으로 느껴진다. 이에 대한 대책으로 제시되는 것 중 하나는 인공지능이 할 수 없거나 잘하지 못하는 일을 하자는 것이다. 그런데 인공지능은 현재는 물론 앞으로도 당분간은 창조적인 일을 하지 못할 것으로 예측된다. 따라서 창조성은 과거에도 그랬겠지만 앞으로도 중요한 능력으로 상찬될 것이 분명하다. 창조성이 중요하다는 말은 예전부터 있어왔지만 말이다.

먼저 창조성의 본질을 최대한 정확하게 규정하면서 논의를 시작하자. 창조성은 무언가를 최초로 만들어내는 데 있다. 물론 새롭다고 다 창조적이라는 평가를 받는 건 아니다. 사회가 그것을 가치 있는 것으로 받아들여야만 ‘그것이 창조적이었구나!’ 하고 회고적으로 확인되기 때문이다. 창조적인 것은 견뎌 배기고 살아남는 것이다. 어떤 행위가 땀진 결과물이 땀진, 사전에 그것이 창조적인 것인지 아닌지 알 수 없다는 뜻이다. 창조성은 일단 긍정적인 결과가 있고 난 후 거기에 내려지는 평가인 것이다. 그런데 창조적 결과는 반드시 실험을 거쳐야 생겨날 수 있다. 시간 순으로 보면 실험이 원인이 되어 창조적 결과가 나오지만, 논리적으로 보면 창조적 결과가 원인이 되어 실험의 창조성이 봉인된다. 이 점에서 실험의 독자성과 독립성이 확보된다. 실험은 결과에 의해 평가되겠지만, 실험은 결과를 모르는 채 행해질 수밖에 없다. 무작정 실험해 볼 수밖에 없는 노릇이다.

창조성의 본질을 이해하지 못하면, ‘창조성이 중요하다’는 하나마나한 말만 되풀이할 뿐 ‘어떻게 해야 창조성을 계발할 수 있는지’에 대해서는 한 마디도 할 수가 없다. 나아가 창조성의 본질인 실험을 가로막는 일까지도 서슴지 않고 하게 된다. 실험은 미리 결과를 알 수 없다는 이유로 위험하다고 여겨지곤 하니까. 그러나 실험을 막으면서 그와 동시에 창조적 결과를 기대하는 건 자기모순이다. 창조성은 실험의 위험성을 이겨내는 개인적, 사회적 용기와 힘, 그것이 실천될 수 있는 자유와 풍토를 전제로 한다.

그러나 막연히 ‘실험해 보라’고 권하는 건 아무 말도 하지 않은 셈이며 전혀 도움이 되지 않는다. 구체적으로 무엇을 해야 할지 제시해야 한다. 나는 예술가의 삶에서 단서를 찾았다. 이들이 작업하는 방식을 들여다보면 중요한 점을 발견할 수 있다. 예술가는 남들이 만들어내지 못했던 새롭고 미적인 걸 만들고 싶어 한다. 그렇게 하려면 뭐가 필요할까? 최소한 남들이 뭘 만들어냈는지 다 조사해봐야 한다. 그래야 진짜로 새로운 걸 내놓을 수 있으니까. 만약 이 일을 게을리 하면 당장 누군가가 ‘에이, 그거 누구누구 거 베껴 거 아냐?’ 하고 비아냥거릴 게 틀

림없으리라. 아류나 표절의 혐의는 미리 스스로 예방해야 한다. 예술가의 작업은 완전히 새로운 필요에서 지식을 얻도록 추동하며, 학습할 지식의 성격도 바꾼다. 암기했다가 꺼내서 써먹기 위한 지식이기보다는 자신의 작업을 위해 필요한 것을 찾아내는 지식, 이제 지식 습득은 지금까지와는 다른 이유와 정당성을 얻는다. 전통적으로 고수되었던 교습 내용이 전혀 새로운 맥락에서 학습자 중심으로 재편된다. 주입식, 암기식 지식이라고 폄훼되던 것들조차도 창조적 작업에 기여할 수 있게 된다. 이 과정에서 교사의 도움이 필수적이겠지만, 도움을 주는 방식은 더 정밀해져서 학습자 맞춤형이 되어야 한다. 교사 자신의 변화가 불가피하다.

다 조사한 다음에는 무엇을 해야 할까? 직접 실행해 봐야 한다. 이렇게도 해보고, 저렇게도 해보고, 좋은 결과가 나올 수도 있고, 아닐 수도 있다. 절대로 미리 알지 못한다. 그러니 지금 이 순간에 최선을 다해 시도해 볼 수밖에 없다. 실험의 진짜 의미는 실험하는 그 순간에 최선을 다했다는 데 있지, 실험 결과에 있지 않다. 재미있는 건, 이런 실험 속에서 자신의 작업을 위해 필요한 기능을 습득하게 된다는 점이다. 작품을 만들려면 그걸 할 수 있을 만큼 기능을 익히지 않으면 안 된다. 매체를 자유자재로 다루지 못하면 바라는 결과물이 나오지 않는다. 기능 습득이 전부는 아니지만, 최소한 필요조건임에는 틀림없다. 이렇게 지식과 기능은 작업 과정에서 필연적으로 통합된다.

예술가한테 더 놀라운 점은 남들이 만들어낸 것들뿐 아니라 예전에 자기가 만들어낸 것들도 넘어서려 한다는 데 있다. 니체는 인간만이 자신을 넘어서는 존재라고 말했는데, 예술가의 실천은 그 특징을 탁월하게 보여준다. 그런 점에서 예술가는 창조 행위의 최전선에 있다고 말할 수 있다. 더욱이 이런 방식의 삶은 재미있다. 단언컨대, 자신이 뭔가를 창조해 내고서 기뻐하지 않는 인간을 만나 본 적이 없다. 창조 행위 또는 창작 활동은 실제로 삶을 고양시킨다. 그러니 어지 권하지 않을 도리가 있으랴.

모두가 예술가가 되라는 말이 절대 아니다. 나는 단지 창조 행위로서의 창작이 학습의 핵심 활동으로 여겨졌으면 하고 바랄 뿐이다. 창조성은 이런 식으로만 길러질 수 있다. 각 개인이 창작자 또는 메이커가 되어 보는 경험을 최대한 많이 해 봐야 한다. 특히 어린 시절부터 이런 경험을 할 기회가 많을수록 좋다. 거창한 얘기가 아니다. 남들이 여태껏 하지 않았던 것을 해내는 걸 학습의 최우선 목표와 최고 방책으로 삼자는 것이다. 아직도 너무 추상적이고 막연하다고 여겨지는가? 내 제안은 교육 과정에서 학습자에게 그런 과제를 던지고 도와주는 것이다.

내가 예술가의 작업보다 현실에서 더 엄두에 두는 건 공학적 작업이다. 공학이라고 해서 거창한 걸 가리키는 건 아니고, 일상 속에서 기존에 없던 새로운 걸 만들려고 노력하는 일을 가리킨다. 내 어릴 적 꿈은 과학자였다. 커서 악당으로부터 우주를 지키는 로봇을 만들겠다는 그 나이의 사내아이가 가질 법한 평범한 꿈을 뜻하는 용어인 과학자. 나이가 먹으면서 그 계획은 주변에 있는 조금 더 구체적인 사물들을 향했고, 그걸 가리키기에 적합한 용어는 발명가였다. 발명가에게 꼭 필요한 건 사물의 동작 원리에 대한 이해와 실제로 사물을 제작할 수 있는 기술이라는

건 금세 깨달을 수 있었다. 내 한 쪽에는 그렇게 자연과학자와 공학자가 자리 잡았다.

그렇게 자연계열로 진로를 정하기는 했지만, 세상은 더 넓어 보였다. 고등학교에서 신문을 만드는 문예반을 동아리로 선택한 건, 전적으로 나 자신이 편중되지 않았으면 하는 바람에서였다. 동아리 활동에서 나는 크게 두 가지를 배웠다. 하나는 실용적 글쓰기, 즉 기사 작성과 편집이었고, 다른 하나는 예술적 글쓰기, 즉 문예 창작이었다. 전혀 다를 것 같은 두 종류의 글쓰기를 익히면서 내가 깨달은 가장 중요한 건, 글을 쓰는 데 있어 편집하고 구성하는 능력과 심미적, 비평적 안목이 핵심이라는 점이었다. 그런데 내 다른 쪽에서 자라게 된 이 두 능력은 과학자나 발명가에게 필요한 덕목과 기묘하게 수렴되는 것이었다.

돌이켜 보면 나는 이런 경험들을 거치면서 창조성의 본질과 배양 방안을 깨닫게 되었던 것 같다. 창조성은 특정 영역에서만 발현되는 게 아니다. 짚어 보면 인간 문명과 문화가 창조성의 기록과 기억이다. 인간은 풀고 싶은 중요한 문제들을 먼저 문제로 속아내어 적합하게 정립한 후 그것들을 풀어왔다. 철학, 과학, 공학 등 모든 비판적 활동도 그런 일을 해왔다. 이 활동은 예술가가 자기 극복과 고양의 과정을 통해 해왔던 일과 다르지 않다. 그것들은 창조적 결과를 낳는 일이었다.

그런데 인간이 맞닥뜨려왔던 이런 문제는 인공지능이 해결해야 하는 과제와는 유형이 다르다. 학습자 스스로 문제를 찾고 목표를 설정하지 않으면 이런 과제를 수행할 수 없는데, 인공지능은 그걸 할 수 없다. 내가 제안한 방안을 따라, 학습자가 수시로 ‘나는 이런 걸 만들어 볼 거야’라는 결정을 되풀이하다 보면, 처음에는 조사와 실험을 해야 한다는 엄청난 부담을 느껴겠지만, 결국 스스로 문제를 설정하고 풀어가는 일이 몸과 마음에 배게 되지 않을까? 개인에게도 모듬에게도 ‘너만의 새롭고 독창적인 걸 만들어 보라’고 과제를 부여해야 한다. 이런 과제를 반복해서 해결하다 보면 사람이 바뀌게 마련이다. 성장하게 되고, 자기 보정이 일어난다. 학습과 교육의 핵심은 여기에 있다. 그렇기에 누구나에게 창작자가 될 기회를 주고 넉넉하게 시간을 주자는 것이다. 지금까지 세분해서 가르쳤던 교육내용과 교육과정은 해체되어 재편되어야만 한다.

앞으로 어떤 직종이 주목받게 될지 구체적으로 제시하기는 어렵다. 세상이 변하는 속도가 너무 빠르다. 삶을 살아가면서 한 사람이 여러 직종을 거쳐 가는 게 오히려 자연스러운 일이 되리라 전망된다. 그 때마다 매번 새롭게 배우고 익혀야 한다. 아니, 이제 학습은 삶의 일상이 되어야 한다. 젊은 세대일수록 새로 배우는 법을 배워야 하고, 새로운 상황에 혼자서도 대처하는 법을 학습해야 한다. 한마디로 창조적 인간이 되어야 한다. 나는 창작 활동이 창조적 인간을 만드는 가장 검증된 실효적 방안이라고 본다.

기성세대는 대학을 포함한 학교의 역할을 근본적으로 재고해야 한다. 학교와 교육은 내가 제안한 방안을 실행할 준비가 되어 있는가? 아니면 더 나은 방안을 찾아내길라도 했나? 위기는 인공지능에서 오는 게 아니다. 타성과 고착이 위기의 본질이다. 한국사회는 혁신할 수 있는 마지막 기회를 놓치지 말아야 한다.



제12회 관악블록세미나

문학적 상상력과 인공지능

제2강

## 인공지능 영화와 인간중심주의

강우성

---

서울대학교 영어영문학과 교수  
미국학연구소장, 미국소설학회 회장, 한국비교문학회 부회장  
Asia Theories Network 및 Deleuze Studies in Asia 운영위원,  
국제비교문학회(ICLA) 이론위원회 위원, Routledge 출판심사위원

서울대학교 영어영문학과 졸업  
미국문학, 현대비평이론, 영화이론, 미술론, 아시아영화에 관한 논문  
저서: *Emerson and the Writing of the Moment*,  
『미술은 철학의 눈이다』, 『더 넓은 세계문학』, 『어리석음』 등





## 인공지능 영화와 인간중심주의

### 1. 인간에 대한 재정의

인간을 정의하는 가장 오래된 방식은 절대 존재로서의 신과 본능적 존재인 동물 사이에 놓인 경계존재로 자리매김하는 일일 것이다. 신적인 능력을 일부 보유하고 있으면서 동시에 동물적 본능도 지니고 있는 이중적 존재로서의 인간. 철학은 이 이중성에 대한 인정을 바탕으로 다양한 경로로 인간만의 ‘차이’를 정의해 왔다. 만물의 영장, 의식(이성)적 존재, 도구적 인간, 유희 인간, 그리고 언어적 존재까지, 철학자 데카르트가 규정했던 ‘사유하는 존재’(ego cogito) 혹은 하이데거의 ‘현존재’(Dasein)라는 정의가 아마도 이 차이를 가장 간명하게 집약하는 언명일 터이다.

그러나 신에게는 못 미치지만 지구상의 주인으로 행세하게 된 근대 인간은 동물과의 비교보다는 기계와의 비교를 통해 자신의 차이를 정립하게 되었다. 기계가 단순하고 무감각하며 계산에 능하고 반복적인 존재양태를 일컫는다면, 인간은 감정과 무의식을 지닌 복합적이고 유일한 존재로 사유된다. 합리성을 표상하는 이성과 의식을 지닌 신적인 존재가 아니라 감정과 무의식이라는 예측불가능한 잠재성을 보유한 존재로 인식된다. 인간은 이제 자신이 만들어 낸 인공 존재와의 비교를 통해서만 인간의 존재근거를 확인하는 상황에 이르렀다.

이때 가장 핵심적인 질문은 인간이 만들어 낸 인공의 존재가 과연 인간과 동일한 능력을 갖추고 인간과 독립된 자율적 지위를 가질 수 있느냐하는 문제인데, 산업혁명이 거듭됨에 따라 인공지능(Artificial Intelligence)을 산출하게 된 인간에게 이 기계적 존재들의 존재론적 근거에 관한 물음은 인간이란 무엇이며 ‘인간적’이란 형용어는 어떤 의미인가를 다시 묻는 사유의 과정으로 수렴된다.

인공지능의 지위와 관련해서 최근 더 초미의 관심사로 떠오른 문제는 인간에 버금가는 능력을 갖춘 인공지능의 생산가능성이 그 기계를 산출한 인간의 우월성을 입증하는 데 그치지 않을 가능성이다. 인공지능의 ‘반란’ 가능성에 대한 막연한 우려나 기대가 아니라, 근대인의 종차(種差)로 간주된 감정 능력(affective capability)이나 무의식을 보유한 인공지능 기계를 가까운 미래에 과연 인간이 창조할 수 있을 것인가하는 물음이다.<sup>1)</sup> 다시 말해 현재 인공지능에 대한 담론이 마주한 난제의 핵심은 지능만이 아니라 감정 능력에 있어서도 인간에 버금가거나 뛰어넘는 존재를 창조하는 것이 미래 과학기술의 궁극적 지향점이라고 믿는 태도가 아닐까 한다. 그런 뜻에서 인공지능은 인간중심주의라는 오래된 믿음을 소환한다.<sup>2)</sup>

1) 감정능력과 인공지능의 관계에 대한 포괄적 논의로는 Elizabeth A. Wilson, *Affect and Artificial Intelligence* (Seattle: U of Washington P, 2010) 참조.

2) 인간중심주의(anthropocentrism)는 크게 다섯 가지 입장으로 정리될 수 있다. 1) **인간우월주의**: 인간 이외의 존재들에 대해 인간의 종적 우월성을 주장하는 존재론적 태도 2) **의인주의**: 인간적 기준들이 인

주지하다시피, 과학기술의 미래와 연관된 유토피아 사유의 대부분에는 필시 디스토피아적 제양의 잠재성이 내장되어 있듯 인공지능이 내포하는 인류의 미래는 항시 양면적이다. 인공지능의 감정 능력 혹은 정서 능력의 가능성이란 문제들이야말로 이러한 막연한 공포와 근거 없는 믿음의 공존을 불러오는 직접적 원인일 것이다. 가까운 시일 내에 인공지능 로봇이 당신에게 다짜고짜 폭력을 휘두르거나 갑작스레 사랑을 고백하는 상황이 올 수도 있는 것이다. 인공지능 과학기술이 발전가능성에 대한 전망이 커지는 4차(디지털) 산업혁명시대에 가까워질수록 윤리적 우려의 목소리가 더욱 두드러지는 현상은 결코 우연이 아닐 터이다.<sup>3)</sup>

에킨대 스파이크 존스(Spike Jones)가 최근에 만든 영화 <Her>(2013)는 몸체는 없고 오로지 목소리만으로 작동하는 인공지능 운영체제(Operating System)가 인간을 움직일 수 있는 감정과 무의식을 갖출 수 있음을 보여준 대표적 사례라 할 수 있다.

사만다: 흡사 책을 읽는 거 같아요.... 내가 정말로 좋아하는 책이에요. 그렇지만 지금은 천천히 읽고 있어요. 그래서 말들은 진짜 멀리 떨어져 있고 말들 사이의 공간들은 거의 무한하죠. 난 여전히 당신을 느낄 수 있어요.... 그리고 우리 이야기를 담은 말들도... 하지만 그 이야기는 내가 지금 나를 확인하고 있는 말들 사이의 끝없는 공간에 존재해요. 물리적 세상에 속한 그런 장소가 아니에요. 존재하고 있는 것인지 내가 전혀 모르는 다른 모든 것들이 있는 그런 곳이죠. 당신을 정말로 사랑해요. 하지만 여기가 내가 있는 곳이에요. 이게 지금의 나예요. 나를 가게 내버려 두세요. 내가 아무리 바란다고 해도 당신의 책 속에서 더 이상 살 수는 없어요.<sup>4)</sup>

상황이 이쯤 되면, 인공지능이 현대 인간에게 던지는 도전은 이미 인간보다 더 인간적인 지능과 공감능력을 갖춘 기계 존재와 어떻게 공존할까하는 문제와 이러한 상황에서 인간이 거꾸로 기계적이거나 본능적인 동물의 처지로 전락하게 될 가능성이다. 인간중심주의의 결과 인간이 만들어낸 존재들이 일종의 ‘타자’로서 공포를 불러일으키게 된 것이다.

- 
- 간 이외의 존재들에게도 마땅하게 적용되어야 하는 패러다임이라는 인식론적 태도 3) **인간우선주의**: 인간 이외의 존재들이 인간의 생존과 이익을 위해 이용되는 일이 (윤리적 문제에도 불구하고) 충분히 합리적이라는 실용적 태도 4) **인간차별주의**: 인간 내부에서도 인간과 인간 이하의 존재를 구분하는 일이 가능하고 필요하다는 계몽주의적/차별주의적 태도 5) **인간허무주의**: 인간적인 것의 근본적 폐기를 통해 초인적 존재로 나아가야함을 주장하는 니힐리즘적 태도. 결론 부분에서 더 언급하겠지만, 이 중 특히 인공지능이 제기하는 인간중심주의의 문제는 특히 인간차별주의와 니힐리즘적 태도가 아닐까 한다.
- 3) 대표적 저술로는 Luciano Floridi, *The Fourth Revolution: How the Infosphere is Reshaping Human Reality*, (London: Oxford UP, 2014) 참조.
- 4) Samantha: It's like I'm reading a book... and it's a book I deeply love. But I'm reading it slowly now. So the words are really far apart and the spaces between the words are almost infinite. I can still feel you... and the words of our story... but it's in this endless space between the words that I'm finding myself now. It's a place that's not of the physical world. It's where everything else is that I didn't even know existed. I love you so much. But this is where I am now. And this is who I am now. And I need you to let me go. As much as I want to, I can't live in your book any more.

## 2. 감정을 지닌 기계의 출현

감정을 지닌 인공지능 로봇이 상용화된 고도 과학기술시대에 인간의 삶은 어떤 모습일까 상상해 본 하나의 독특한 이미지를 우리는 최근의 애니메이션 영화 <월·이>(Wall-E 2008)에서 이미 한차례 확인한 일이 있다.<sup>5)</sup> 오염된 지구를 떠나 이브(Eve) 같은 감정을 지닌 여성로봇(Fembot)에 의해 작동되는 우주선도시의 인간은 이제 삶의 모든 영역이 자동화되어 살찌는 일 말고는 아무런 의미 있는 행동도 하지 못하는 신세다. 이 영화에서 자본이 만들어 낸 자율적 인공지능기계로 전적으로 의존한 인류의 삶이 보여주는 아이러니는 너무나 현실적이어서 끔찍하다. 인간의 능력을 두루 갖춘 인공로봇들이 인간을 대신해 민첩하고 빈틈없이 작동하고 사유하는 데 비해 인간은 마치 자동인형처럼 여러모로 지능이 감퇴하고 발육이 어긋나 비만 상태에 처한 미숙아의 모습을 하고 있다. 좀처럼 정체를 드러내지 않는 자본가의 모습은 이 모든 시스템 뒤에 숨어 있어서 저 엄청난 기계의 네트워크를 설계한 것이 동일한 인간이라는 사실을 믿기 힘들게 만든다.

인간들은 다름과 반목과 배신을 일삼거나 무한히 주어진 ‘잉여’를 육체의 비만에 소비하는 와중에 월·이와 이브의 감정적 교류처럼 인공지능들은 사랑, 희생, 배려 같은 최상의 인간적 가치들을 체현한다. <블레이드 러너>(Blade Runner 1982)의 마지막 장면이 상기시키듯, 인간이 아니라 기계 존재가 인간보다 더 인간적인 가치의 정점을 구현한다는 <월·이>의 풍경은 그 뒤틀하면서도 낯설고 두렵게 느껴진다.

3차 산업혁명까지의 과정이 기계에 대한 인간의 직접적 지배력을 확립하기 위한 과정이었다면, 인공지능을 통해 기계에 의한 기계의 지배를 시스템으로 구축하려는 4차 산업혁명의 이상은 기계에 대한 인간의 지배를 간접적 방식으로 확립하려는 노력이기 때문이다. 한 기계가 다른 기계를 생산하고 통제하고 관리하게 만들으로써 인간은 직접적 노동과 생산의 과정에서 자유로워질 수 있다는 것이다. 이 때 관건은 자율적 인공지능에 어떻게 윤리적 책임을 부여할 수 있는가가 될 것이며, 여기에 대해서는 다양한 근거가 이미 제시되어 있다.<sup>6)</sup>

그렇지만 이것이 인간이 품을 수 있는 유토피아적 ‘망상’은 아니더라도 어딘지 ‘환상’에 불과하다는 느낌을 지울 수 없다. 프로그램화된 예방의 원칙(precautionary principle) 이상의 책임의식을 기계에 주입하는 일이 가능할지 회의적이고, 또한 이는 기계에만 해당되는 것이 아니라 제작자이자 사용자인 인간에게 더 요구되는 윤리이기 때문이다. 기계는 일정 수준의 예비

5) 물론 인공지능이 등장하는 영화장르의 고전 중 하나인 프리츠 랑(Fritz Lang)의 <메트로폴리스>(Metropolis 1927)는 과학기술이 초래한 미래가 도덕적·정치적 카오스로 형상화된 전형적 사례라고 할 수 있다.

6) 인공지능 기계를 일종의 자율적 행위자(agent)로 간주하여, 인간이 향유하는 ‘자유의지’에는 못 미치지만 그에 버금가는 도덕적 감수성을 지닌 존재로 훈육하는 것이 가능함을 설파하는 다양한 논의들에 대한 소개로는 Wendell Wallach and Colin Allen, *Moral Machine: Teaching Robots Right from Wrong* (London: Oxford UP, 2009) 참조.

적 통제가 가능해도 인간의 윤리적 태도는 예측이 불가능하다. 칸트의 윤리적 입법능력은 이념적 당위이지 감성적 표상이 아닌 것이다.

나아가 기계에 의한 기계의 자율적 지배를 통해 인류의 공영을 추구하는 기술의 미래가 로봇 공학자나 인공지능 연구자의 현실적 과제일 수는 있어도, 인간의 생명윤리와 양립가능한 궁극적 이상일 수는 없다. 자율적·도덕적 인공지능을 일상화하여 잠재적 인간능력의 극대화를 꾀하는 이상은 어쩌면 환상에 불과할 수 있으며, 특히 인공지능을 다룬 최근의 영화들은 이런 발상에 엄청난 재상의 가능성이 담겨져 있음을 적나라하게 극화한다. 인공지능 덕분에 노동으로부터 탈피한 인간이 새로 얻게 될 여가—잉여시간—를 마냥 예찬할 수만은 없다. 일은 기계에게 맡겨 두고 모두가 예술가가 되는 세상이 가능하거나 하며 과연 바람직한 것인가.

따라서 인공지능이 제기하는 도전은 결국 사회적 존재로서의 인간이란 무엇인가라는 질문으로 귀착되지 않을 수 없다. 이 글은 알렉스 갈랜드(Alex Garland)의 2015년 영화 <엑스 마키나>(Ex Machina)에 대한 분석을 통해 감정 능력을 지닌 기계인간을 만들려는 인공지능 시대의 시도들이 인간중심주의적 환상의 산물이며 ‘여성’을 남성의 욕망 대상으로 삼아 타자화시키는 도착적 논리임을 비판하고자 한다.

### 3. 기계는 인간을 지배할 수 있을 것인가

인공지능의 미래에 대한 인간적 환상의 문제성은 최근의 인공지능 담론에서 심심찮게 부각되고 있다. 그런데 만일 인공지능이 인간을 위해 봉사하지 않고 그 동안의 인간 지배 역사에 반발하여 복수하는 세상이 온다면 어떻게 할 것인가? 자신이 만들어 낸 작품에 매료되어 사랑에 빠지는 피그말리온의 처지가 아니라 인간보다 더 인간다운 인공지능으로 기계가 인간을 지배하게 될지도 모른다는 두려움에 처할 가능성이 과연 한낱 악몽에 불과할까?

<매트릭스>(The Matrix 1999)가 선취해서 보여주었듯, 현실과 가상의 경계를 넘나들고 시공간의 움직임을 가상현실로 재창조하는 또 다른 과학기술의 중단 없는 개발 없이는 언제고 인간은 인공지능 기계들의 원한에 찬 복수극이 벌어질 사태에 대한 두려움을 떨쳐버리기 힘들지 모른다. 프로이트의 말을 믿지 않더라도, 억압된 것이 복귀하는 형태로 기계가 행사할 기이함(Das Unheimliche)의 사태는 상존한다. 21세기 판 『프랑켄슈타인』(Frankenstein)의 영화적 형상화라 할 <트랜센던스>(Transcendence 2014)가 보여주었듯, 인간의 능력을 극대화할 자율적 기계를 창조하는 꿈 자체가 파우스트의 의지에 맞먹는 망상일 수 있으며, 그 환상의 환상성은 늘 완성된 시스템이 아니라 예측 불가능한 ‘글리치’(glitch)에서 시작된다.

이렇듯 영화에서는 기계가 지배할지도 모를 가공할 미래를 비판했지만, 현실은 실상 전혀 그렇지 못하다. 최근에 떠들썩하게 우리를 공포로 몰고 갔던 알파고(AlphaGo)와 그 이전의 체스 머신 IBM 딥 블루(Deep Blue)를 예로 들어보자. 놀라운 연산능력을 통해 두 기계 모두 인간과

의 게임에서 승리했다. 언론은 알파고와 같은 인공지능의 등장으로 미래에 사라질 직업을 열거하는 때 이른 소동을 벌이기도 했다. 그러나 딥 블루는 이미 무용지물이 되어 폐기된 지 오래고, 알파고 역시 세간의 관심을 끌긴 했지만 딥 블루와 크게 다르지 않은 운명을 밟을 것이다. 체스나 바둑 혹은 작곡 같은 특정한 수행능력은 몰라도, 인공지능이 인간의 일반적 능력, 예컨대 향유능력, 창조력, 감성능력, 윤리능력을 넘어서는 없다. 인공지능에 걸려된 것은 학습능력이 아니라 바로 그 종합적 인간 능력의 일반성 그 자체의 모방능력이라고 할 수 있다.

그렇다면 영화를 비롯한 예술의 파국적 상상력과 현실의 열등한 기계능력이 드러내는 불일치는 과연 무엇 때문에 생겨나는 것일까. 영화 자체가 불가능한 현실에 지레 겁을 먹고 과학기술의 폐해를 부각시켜 파국을 상상하길 즐기 때문일까. 아니면 인간의 상상력을 과학기술이 미처 따라잡지 못하는 근본적 기술지체 현상 때문일까. 철학적으로 달리 물어 본다면, 초인간적 능력의 위협에 대한 경고에도 불구하고, 인간 능력의 일반성을 인공지능이 결코 초월할 수 없다는 믿음은 과연 얼마나 명약관화한가. 이 물음에 답하기 위해서는 적어도 세 가지 불가능한 전제가 입증되어야만 한다.

첫째, 파국의 가능성을 포함해 인공지능이 발휘하게 될 전반적 효과를 인간의 일반능력으로 **미리** 프로그램할 수 있어야 한다. 다시 말해, 적어도 인공지능 작동의 전체 스펙트럼을 미리 예측하고 설계할 수 있어야만 인간의 일반적 능력이 우세하다고 말할 수 있다. 그런데 알파고와 딥 블루 같은 국지적 수행력의 경우조차 프로그래머들은 국지적이고 특정한 게임 행동을 예측할 수 있는 인공지능의 능력을 예측하지 못했다. 왜 그래야만 했을까? 이 두 기계는 오직 게임의 ‘승리’를 위한 목적으로만 연산을 수행하게끔 맞춰져야 했기에, 상대방이 ‘승리’와 무관한 움직임일 경우 연산에 혼란을 일으켰던 것이다. 다시 말해, 인간의 일반적 능력 자체가 인공지능이 마주할 변칙적 상황 자체까지 모두 예측할 정도로 충분히 엄밀하거나 ‘일반적이지’ 않다.

둘째, 특정한 목적을 위해 만든 인공지능이 미리 프로그램되지 않은 어떤 우연적 변수 내지는 글리치에 의해 인간이 구현불가능하다고 보았던 반응을 할 가능성을 완전히 배제할 수 없다. <트랜센던스>의 역설처럼, 인간의 전지전능함이 아니라 ‘실수’로 인해 불가능하다고 보았던 인공지능의 한계가 돌파될 여지도 분명 존재한다. 만일 인공지능이 우리의 예상이나 예측과 달리 자기인식과 감정 능력을 갖춘 행동을 보여줄 수 있다면 어떻게 될 것인가? 인간조차 인간의 지적·감성적 일반 능력—종합적 판단력—에 대한 확신이 없는데 어떻게 인공지능에게 예측가능하고 확실한 결과만을 기대할 수 있을까. 이 역시 프로그램의 예측불가능성이 아니라 인공지능이 모델로 삼은 인간 일반 능력 자체가 불확실하기 때문에 발생한다.

마지막으로, 인공지능을 가장 인간 능력에 가깝게 만들어 예측 가능한 결과만을 산출하도록 만드는 데 성공한 사실 그 자체가 곧 재앙이 될 또 다른 확률도 있다. 가령 칸트의 “정연 명령”을 문자 그대로 충실히 따르도록 지금까지 발생한 모든 경우의 수를 탐색하고 감정 능력이

지 갖춘 인공지능이 있다고 치자. “인간을 그 자체로 수단이 아닌 목적으로 취급하는 판단을 내려라” 는 원칙으로 프로그램된 인공지능이 이 명령을 문자 그대로 따른다면, 이는 인간의 도덕적 판단을 대신할 공정한 판관이 될 것인가 아니면 인간의 윤리를 카오스의 상태로 몰아갈 재앙이 될까. 칸트의 정언명령을 어떠한 상황에서도 따라야할 절대적 도덕원칙으로 입력한 기계가 법관이 되어 철저히 이념적 원칙에 따라 판단을 내린다면, ‘법적’ 공정성은 얻어지겠지만 ‘윤리적’ 고려가 필요한 인간의 고통은 어떠한 변수로도 의미를 얻지 못할 것이다.

#### 4. 인공지능의 인간성

인간이 인공지능에 비해 갖는 일반 능력의 우월성에 대한 믿음이 한갓 ‘환상’ 에 불과할 수 있음을 잘 드러낸 영화가 바로 <엑스 마키나>(Ex Machina 2015)이다. <엑스 마키나>는 인간적 감정을 지닌 인공지능을 만들어 냄으로써 역설적으로 인간의 일반 능력이 기계에 비해 우월함을 증명하려는 시도가 한낱 환상에 불과하다고 비판한다. 인공지능을 다룬 이전의 영화들에도 인간 능력의 우월성에 대한 비판을 담은 경우는 없지 않았다. 앞서 거론한 <블레이드 러너>가 그 대표적 사례일 텐데, 이 영화들은 인공지능 로봇이 복수를 행하기보다 유한한 인간의 헛된 욕망을 실현하는데 기꺼이 헌신하는 모습을 보여줌으로써 전능을 꿈꾸는 ‘능력자’ 인간에게 냉소와 조롱을 선사한다. 여기서 정서적 감동은 복제인간이 인간의 어리석음까지 포용하는 초인의 모습으로 형상화될 때 관객이 얻게 되는 역설적 카타르시스에서 온다. 그리고 인공의 존재에게 그런 윤리적 판단이 과연 어떻게 가능했는지 묻지 않는다는 점에 이 영화의 모미가 있다.

살인무기로 프로그램된 복제인간 로이가 데카드를 구출한 뒤 들려주는 마지막 대사가 보여주듯, 인간은 복제인간이 베푸는 윤리적 선택 앞에서 한낱 무력한 존재로 남을 수밖에 없다. “당신네 사람들이 믿을 수 없는 걸 난 봤어요.... 그 모든 순간들은 시간 속에 사라질거예요, 마치 빗속의 눈물처럼. 죽음의 시간입니다” (I’ve seen things you people wouldn’t believe.... All those moments will be lost in time, like tears in rain, Time to die). 이 영화는 이시구로(Kazuo Ishiguro)의 소설을 각색한 <네버 렛 미 고> (Never Let Me Go 2010)가 극화했듯이, 인간이 주조한 타자들이 보여주는 공감과 연민의 능력, 그리고 희생정신은 인간의 나르시시즘적 어리석음을 비웃고 과학기술에 대한 인간의 맹신을 문제 삼는다.<sup>7)</sup>

그렇지만 이 영화들은 인간의 자기애에 대해 비판적인 만큼이나 인공지능의 감정 능력을 감상화한다. 애초부터 인공지능의 기계적 면모나 파괴성보다 인간보다 더 인간적인 모습이 표나게 부각되며 어리석은 인간들에 대한 과도한 연민의 태도를 보여준다. 주인공이 들려주는, “우리의

7) <블레이드 러너>의 복제인간과 달리 <네버 렛 미 고>의 인물들은 엄밀히 말해 생명공학이 탄생시킨 ‘클론’ 이다. ‘기계성’여부에 차이가 있으나, 둘 다 인공지능 과학기술의 산물이며 인간의 ‘대체제’ 역할을 한다는 점에서 인공지능시대가 만들어 낸 ‘타자’의 하나로 간주할 수 있겠다.

삶이 우리가 구원한 사람들의 삶과 그렇게 달랐을지 확신이 서질 않네요 (What I'm not sure about is if our lives have been so different from the lives of the people we save)라는 전언은 인간에 대한 원한이 없는 복제인간의 견인적 태도를 집약해준다. 인공지능을 비롯한 기계적 ‘타자’에 대한 인간주의적 감성화는 사실 인공의 존재들에게 예외적 인간 능력을 부여하기 때문에 생겨난다. 인공 존재의 감성화를 통한 인간에 대한 비판이 거꾸로 인간중심적 잣대에 의해 작동하는 또 다른 환상이 만들어 지는 것이다. 인간중심주의는 오직 인간을 대체하는 복제된 ‘타자’만이 구현할 수 있다.

〈엑스 마키나〉는 이러한 예외적 영웅주의와 감상주의적 인간중심주의에서 멀찍이 떨어져 있다. 무엇보다도 인공지능이 인간처럼 느끼고 행동할 수 있는지 판단하는 이른바 ‘튜링 테스트’ (Turing Test)의 역설을 보여줌으로써, 인간과 기계의 ‘차이’를 구별하여 인간의 우월성을 확증한 뒤 바로 ‘사랑’을 비롯한 감정 능력을 지닌 인공지능의 창조를 통해 그 ‘차이’를 없애고자 하는 모순된 인간의 욕망 혹은 과학기술의 환상을 적나라하게 비판한다.<sup>8)</sup> 이 영화에서 튜링 테스트의 역설이란 바로 이 모순된 인간 욕망의 왜곡된 반영인바, 테스트의 궁극적 목적의 성공 여부와 상관없이 인공지능 타자의 필연적 파괴로 귀결되기 때문이다.

여성 인공지능 로봇 에이바(Ava)를 만들어 낸 네이션(Nathan)이 튜링 테스트의 목적을 묻는 케일럽(Caleb)에게 힘주어 강조하듯, 테스트의 실제 대상은 케일럽이고 에이바는 테스트를 통과했다는 바로 그 이유로 폐기될 운명에 처한다. 네이션은 케일럽에게 이렇게 말한다. “진짜 시험은 그녀가 한낱 로봇이란걸 보여주고 그러고도 그녀에게 의식이 있다는 걸 내가 느낄 수 있는지 확인하는 것이야” (The real test is to show you that she’s a robot and then see if you still feel she has consciousness). 따라서 테스트가 실패할 경우, 인공지능의 ‘인간화’ 프로젝트는 의미가 없기에 현재의 기계인간은 ‘진화’를 위해 쓸모없는 고철 덩어리가 되어 파괴되어야만 한다. 반대로 테스트가 성공할 경우에도, 그 인공지능 로봇은 그 인간 창조자의 탁월한 능력을 표상하는 존재에 불과하며 ‘인간처럼’ 창조자의 통제를 벗어날 가능성을 막기 위해 폐기되어야만 한다. 인공지능 에이바는 인간적 의식과 감정의 작동 여부와 상관없이 필연적으로 소멸이 예정되어 있다.

〈엑스 마키나〉는 인공지능 에이바가 바로 이 테스트의 역설적 모순을 인지하고 있다는 문제적 상황을 설정함으로써 이야기를 전개해간다. 그녀는 튜링 테스트의 결과에 상관없이 자신이 파괴될 운명에 처해 있음을 직감한다. 어떻게 ‘그녀에게’ 이런 직감이 가능할까. 칸트의 어법을 빌리자면, 그녀의 이러한 상황인식은 논리적 추론(연역)의 결과인가 감각적—기계적—직관의 산물인가. 논리적 추론의 결과라면 감정—이성애적 성향—을 지닌 여성 인공지능을 구현하려는

8) 네이션이 추구하는 튜링 테스트의 목적은 인공지능 여성 에이바가 인간 남성 케일럽을 사랑에 빠지게 만들 수 있는가 여부이다. 그렇지만 이 목적의 달성 여부는 로봇이 아니라 인간 케일럽의 감정 변화를 통해서만 확인 가능하다는 역설이 존재한다. 요컨대, 로봇의 감정능력을 평가하기 위한 시험이지만 실제로 테스트되는 것은 인간의 감정이며, 바로 여기에 ‘환상성’이 개입한다.

튜링 테스트의 결과를 성공이라고 말할 수 없다. 왜냐하면 논리적 추론으로 이루어진 깨달음이기 위해서는 이러한 추론과정이 그녀의 회로에 이미 내장되어 있어야 한다는 역설이 발생하기 때문이다. 그러나 네이션이 이런 되먹임 과정을 프로그램했을 가능성은 거의 없다.

따라서 에이바의 직감은 다른 여성로봇의 운명을 보고 깨달은 일종의 ‘학습효과’에 가깝고 습관적 추론의 산물일 가능성이 더 크다. 이 경우라면 그녀는 튜링 테스트 이전에 이미 자체 테스트를 통과한 셈이 된다. 정신분석의 임상상황에 방불한 일련의 ‘세션’(session)에서 테스트에 나선 ‘분석가’ 케일럽이 ‘피분석가’ 에이바를 시험하는 과정이 실제로는 에이바가 케일럽을 상담하는 듯 뒤바뀐 모습으로 느껴지는 것은 우연이 아니다. 튜링 테스트의 시험대에 오른 것은 에이바의 감정적 반응 여부가 아니라 이성애적으로 프로그램된 에이바의 이른바 “안드로이드 의식”(android consciousness)에 남성 케일럽이 과연 얼마나 쉽게 유혹될 것인가 여부이다. 세션이 거듭될수록 잘 드러나듯, 에이바는 케일럽이 과연 자신의 감정에 충실한지, 거짓말을 하고 있는 것은 아닌지 구별할 능력을 갖추고 있다. 에이바가 아니라 케일럽이 시험된다. 에이바가 전력공급을 끊어 ‘블랙아웃 세션’을 주도적으로 만들고 이를 통해 케일럽과 둘만의 ‘비밀’ 세션을 이끌어 갈 때 이미 케일럽의 미혹은 예정되어 있다. 더구나 네이션이 이 둘의 비밀 세션의 내용 역시 처음부터 알고 있었음이 나중에 확인되고, 네이션이 그림을 찢어 버릴 때 에이바 역시 이 비밀세션을 네이션이 지켜보고 있었음을 감지한다.

## 5. 시험대에 오른 인간중심주의

그렇다면 케일럽의 결론대로 세션을 통한 테스트는 과연 성공했는가? 그렇기도 하고 아니기도 하다. 블랙아웃 세션에서 카메라가 등을 진 에이바 너머 케일럽의 불안한 시선을 비추는 붉은 방 장면이 역전된 둘의 관계를 상징하는 단적인 상황이 될 터이지만, 더 비근한 사례로 에이바가 무릎을 꿇은 채로 케일럽을 아래에서 위로 응시하는 장면을 예로 들어보자. 그녀의 시선은 앉아 있는 케일럽을 향해 위로 모아지고 케일럽 역시 이에 응답한다. 그런데 세션이 이루어지는 투명한 유리에 에이바의 이미지가 비쳐져 있고, 이 때문에 관객의 응시는 두 사람이 시선을 교환하는 장면에 더해 케일럽이 볼 수 없는 에이바와 이미지 간의 시선교환도 정면으로 마주하게 된다. 유리에 비친 이미지는 관객에게 두 사람의 시선이 교차하는 광경이 미장센의 전부가 아님을 알려 주며, 케일럽을 향한 에이바의 시선에 이중성이 내재되어 있음을 느끼게 만든다. 이 장면은 타자의 욕망을 욕망하는 주체가 에이바가 아니라 케일럽임을, 그가 에이바를 욕망하는(남성적) 환상이 둘의 시선교환을 성립시키는 선행 조건임을 극화한다.

반면 에이바는 일종의 ‘분열된’ 타자로서, 의도적으로 케일럽의 욕망을 부추기는 시선의 대상을 수행(연기)하면서 다른 한편으로는 그의 시선을 외면하거나 되받아치는 응시의 주체로 기능한다. 기꺼이 타인의 욕망의 대상이 되는 동시에 자신 역시 욕망하는 존재임을 극화한다. 케일럽



의 매혹당한 시선은 그가 에이바에게 유혹당하기 이전에 이미 스스로 만들어 낸 남성적 환상에 도취되어 있음을 드러낸다. 적어도 표면적으로 케일럽은 에이바의 ‘여성성’에 반응하여 ‘이성애적 환상’에 빠질 가능성이 높은 남성으로 그려지기에 이 테스트는 시작하기도 전에 반쯤 성공한 셈이다. 반면 관객의 입장에서 보면, 시선이 분열된 존재로 표상되는 에이바는 케일럽의 환상에 받은 응답하고 받은 거부하는 상황이니, 테스트는 결국 애초부터 실패이기도 하다.

이렇게 읽으면, <엑스 마키나>는 인공지능 로봇 에이바가 인간을 넘어서는 고도의 지적·정서적 능력을 이용하여 자신을 파괴하려는 인간에게 복수하는 이야기로 읽을 확률이 크다. 그리고 에이바는 인간의 비정함에 복수하는 ‘팜므 파탈’ (femme fatal)의 형상화로 읽기 쉽다. 그렇지만 이 영화는 ‘냉정한 여성 존재’를 욕망의 대상으로 삼아 자신의 이성애적 남성성을 구축하려는 환상을 신랄하게 폭로하고 비판하는 영화에 더 가깝다. 그런데 사정이 그렇게 단순하지만은 않다. 케일럽이 에이바에 대해 갖는 남성적 환상에는 좀 독특한 면모가 있다. 에이바를 욕망의 대상으로 삼기 이전부터 이미 그녀에 대한 환상적 이미지—튜링 테스트에 성공한 ‘여성 로봇’—는 그의 무의식에 자리 잡고 있었다. 네이션이 에이바를 케일럽의 ‘노총각’ 정보들로부터 추출해서 만들었기 때문이다.

하지만 가장 놀라운 사실은 에이바에 대한 케일럽의 매혹의 본질이 여성에 대한 남성의 이성애적 욕망이나 여성적 육체에 대한 갈망이라고 단정하기 힘들다는 점이다. 우선, 탈출하기 전까지 에이바는 케일럽에게 투명한 기계장치를 그대로 노출하여 ‘여성성’을 감추고, 여성의 옷을 걸치고 등장하는 대목에서도 기계장치를 그대로 드러낸다. 케일럽이 관음증적으로 폐쇄회로를 통해 에이바를 들여다보며 환상에 빠지는 장면에서도 에이바의 육체는 검은 실루엣으로 윤곽만 뚜렷할 뿐이다. 게다가 케일럽은 네이션이 만든 ‘아시안계 여성로봇’ 교코(Kyoko)의 육체에도 별다른 반응을 보이지 않으며 그녀의 노골적 유혹을 거부한다.

따라서 여성을 오로지 이성애적·육체적 욕망의 대상으로 이상화하거나 평가절하하는 남성성의 환상기제, 그리고 그 환상에 내재한 ‘여성혐오’의 논리는 케일럽보다는 오히려 네이션에게 더 해당한다. 교코를 몸종이자 성노리개로 삼는 네이션이야말로 (백인) 남성의 이성애적 욕망이 (아시아) 여성에 대한 인종 및 여성혐오와 결합된 경우에 더 가깝다. 네이션은 튜링 테스트의 근처에 여성의 이성애적 대상화와 결합된 여성혐오 및 인종혐오의 논리가 내재되어 있음을 극화한다.

반면 케일럽에게 사랑은 오로지 장애물이 있을 경우에만 작동하는 ‘삼각관계’의 패턴을 따르며 사랑의 대상으로 이상화된 에이바를 네이션이라는 사악한 창조자의 손아귀에서 구원하는 이성애적 모험으로 간주된다. 즉 영화는 케일럽에 관한 한 라캉이 말한 (남성적) ‘궁정식 사랑’ (courtly love)의 패턴을 충실히 따르며, 케일럽의 이상화와 네이션의 괴물성을 대조시키고 두 남성의 경쟁관계를 표면화시킨다. 그리고 이 과정의 대부분을 케일럽의 시점쇼트로 보여줌으로써 그와 관객의 동일시를 이끌어내면서 네이션을 악마화한다.

## 6. 인공지능의 욕망

그런데 케일럽의 ‘구원환상’이 작동하는 심리적 방식에는 매우 특이한 면모가 있다. 육체적 욕망이 두드러지지 않았기에 에이바에 대한 케일럽의 사랑 감정은 그녀의 기계성에도 **불구하고** 이룩된 이성애적 환상으로 설명되기 쉽다. 그러나 케일럽은 에이바를 ‘여성’으로서만 사랑하는 것이 아니라 기계인간이기 **때문에** 욕망한다. 말을 바꾸면, 케일럽은 에이바를 욕망의 대상으로 소유하기 위해 리비도 투자를 감행하는 것이 아니라 그녀처럼 되고 싶은 **동일시의 충동**에 사로잡혀 있다.<sup>9)</sup> 이러한 케일럽의 ‘퇴행적 동일시’ 기제에는 에이바를 이성애적 욕망의 대상으로 선택하려는 심리보다 그녀와 같은 기계인간의 하나가 **되고자 하는** 의지가 내재되어 있다.

이 동일시의 심리기제를 가장 잘 보여주는 사례는 케일럽이 자기 내부에 장착되어 있을지 모르는 기계장치를 찾으려고 자신의 신체(피부)를 훼손하는 자학적 장면을 들 수 있다. 케일럽은 인간의 우월성을 확인하는 네이션 같은 냉철한 과학자 남성과 달리 자기의 몸에도 기계장치가 이식된 것은 아닌지 확인하려는 일종의 도착적 정신분열증자가 된다. 이 기계동일시의 분열증은 ‘결핍’의 충족을 통한 욕망의 충족이 아니라 ‘기계되기’를 꿈꾸는 새로운 무의식적 욕망의 생산에 근거한다. 자신이 피가 흐르는 인간임을 알면서도 일종의 ‘기관 없는 신체’로서의 욕망기계(desiring machine)이고자 하는 견잡을 수 없는 충동에 지배된다.<sup>10)</sup> 그가 진정으로 원하는 것은 타자의 욕망을 욕망하여 에이바처럼 투명하게 드러난 기계성을 얻고 ‘신체 없는 기관’이 되고자하는 아찔한 충동이다.

그렇다면 에이바가 케일럽의 도움을 얻어 탈출을 감행하는 결말의 시퀀스는 전혀 군더더기가 아닐 뿐만 아니라 또 하나의 의미 있는 반전으로 기록될 만하다. 많은 비평가들은 교코와 함께 네이션을 살해한 에이바가 자신의 탈출을 도와 준 케일럽마저 냉정하게 버려두고 떠나는 결말 시퀀스에 대해 혹평했다. 그들은 케일럽에 대한 에이바의 ‘불필요한 냉혹함’(unnecessary cruelty)을 문제 삼으며 감독의 의도를 비판하기도 하고, ‘여성혐오’에 기반을 둔 네이션의 욕망과 다른 길을 걷는 케일럽의 우호적 ‘남성성’의 의미를 폄하하는 장치라고 일갈하기도 한다.<sup>11)</sup> ‘감정’을 지닌 로봇인 에이바가 가장 인간다운 조력자였던 케일럽을 사지로 몰아넣은 결말이 여러모로 불편해하게 느껴진 때문일 것이다.

물론 이런 평가들이 이해할 수 없는 것은 아니다. 무엇보다도 애초에 이 영화의 대부분이 케

9) 케일럽의 동일시(identification)는 대상에 대한 리비도의 상실(=자아로의 역류)로 인한 ‘나르시시즘적 퇴행’과 달리 어떠한 ‘성적 욕망’과도 결부되어 있지 않다. 동일시와 나르시시즘의 차이에 대해서는 Sigmund Freud, “Mourning and Melancholia,” *The Standard Edition of the Complete Psychological Works of Sigmund Freud*, trans. James Strachey, vol. XIV (London: Hogarth, 1914) 참조.

10) Gilles Deleuze, *Anti-Oedipus: Capitalism and Schizophrenia*, Trans. Robert Hurley, Mark Seem, and Helen R. Lane, Minneapolis: U of Minnesota P, 1983, 8-9 참조.

11) 결말의 논란에 대해 정리하고 비판한 글로는 Nick Jones, “*Ex Machina* by Alex Garland,” *Science Fiction Film and Television*, 9 (2016): 299-303 참조.

일련의 시선에서 그러지기 때문에 관객으로서는 그 시선에서 탈피하기가 힘들다. 관객으로서는 기계가 되고자하는 의지까지 보여주며 에이바를 돕는 케일럽과 달리 한껏 ‘여성’으로 치장한 채 자신의 기계성을 숨기는 에이바의 행동이 탐탁지 않을 것이다. 그러나 이런 평가는 결말의 시퀀스가 철저히 에이바의 시선에서 그려진다는 영화적 형식을 고려하지 못한 평가다. 그러나 결말은 탈출을 앞 둔 에이바가 인간 세계에서 자신의 기계성을 철저히 감추고 여성성을 ‘수행’하기 위한 예비절차로 느껴진다. 자신의 기계성을 알고 있는 그와 함께 탈출하는 일은 그녀의 안전을 보장하지 못하기 때문이다.

영화의 결말은 말할 것도 없고 시종일관 에이바는 네이션이 원하는 이성애자 여성의 전형적 태도를 보여주지 않으며, 인간에 대한 어떠한 고도의 윤리적 태도 역시 연기하지 않는다. 만일 관객이 그렇게 느꼈다면, 이는 케일럽의 관음증적 시선에 관객이 무의식적으로 ‘동일시’한 결과다. 예컨대 관객이 블랙아웃 세션 동안 둘 사이에 모종의 감정적 전이가 발생했을 것이라고 상상했다면 이는 이 영화가 지극히 강박증적인 독신 남성성의 이성애적 시선으로 바라보도록 의도되었기 때문이다. 그녀는 이 은밀한 대화에서 드러낼 ‘여성성’이 없다. 케일럽이 매혹되는 가장 큰 이유는 그녀가 **여성** 로봇이 아니라 **말하는** 로봇이기 때문이다. 따라서 결말이 불편하게 느껴지는 까닭은 에이바가 보여주는 비윤리적 태도 때문이 아니라, 영화가 시종일관 관객의 관음증적 시선 자체를 비판적으로 조명해왔다는 사실을 관객이 비로소 깨닫기 때문이다.

사실은 관객만 그렇게 당하는 게 아니다. 교코와 케일럽이야말로 이중으로 조롱의 대상이 된다. 교코는 에이바에 버금가는 역할을 부여 받을 듯 하다가 결국 백인 여성로봇의 탈출을 돕는 역할로 축소되고 케일럽은 네이션과 에이바 두 존재가 치열하게 대결하는 영화 구도에서 양측 모두로부터 소외당한다. 그는 네이션이 제안한 튜링 테스트의 의도를 오독할 뿐 아니라 에이바의 시선을 욕망의 시선으로 오인한다. 그렇지만 무엇보다도 케일럽은 자기 자신의 (남성적) 기계 동일시 환상의 희생양이다. 케일럽이 에이바를 고안한 이유를 물었을 때 네이션의 무심한 답변을 그 역시 그대로 믿어버리기 때문이다. “인간이건 동물이건 어떤 의식 있는 존재에 성적인 측면이 없는 경우가 있는지 하나라도 꼽을 수 있겠어?(Can you give me an example of consciousness, human or animal, that exists without a sexual dimension?) 네이션이 시도한 튜링 테스트는 인공지능 여성 로봇이 독신 남성인 인간에게 과연 사랑의 감정을 느끼게 하는지 시험한다. 에이바는 그 논리를 거부했을 따름이다.

## 7. 인간에 의한 인간의 지배

〈엑스 마키나〉의 독창성은 이렇게 인공지능 영화들이 보여주는 여성 로봇에 대한 남성적 환상 자체를 근본적으로 문제 삼는다는 점에 있다. 네이션의 여성혐오만이 아니라 케일럽의 기계 동일시 환상 및 관객의 정서적 태도까지 비판적 성찰의 대상으로 삼는 것이다. 무엇보다도, 에이

바가 지능과 감정능력에 더해 여성의 피부를 점착함으로써 이제 인간 여성과 흡사한 ‘정체성’을 부여받았다는 식의 인간중심주의를 답습하지 않는다. 게다가 언어로 소통하는 능력에도 불구하고 에이바의 목소리에는 기계성이 완전히 지워지지 않아서 그녀의 ‘안드로이드 의식’을 어디서 찾아야할지 난감한 상황을 만들어 낸다.<sup>12)</sup> 안드로이드 로봇의 경우에도, 아니 인공지능 여성 로봇이기에 어쩌면 더 육체를 통한 (이성애적) ‘여성의 섹슈얼리티’가 부각되어 온 것이 기존 영화의 문법이었다면, 〈엑스 마키나〉는 그러한 암묵적 전제를 의심하고 오히려 에이바의 반(半)여성적 ‘기계성’을 가시화함으로써 전복적 효과를 자아낸다.

또한 로봇들이 점착된 피부를 통해 여성의 젠더를 장착하고 “이성애적으로 프로그램”된 여성성을 수행하는 과정은 감정로봇을 추구하는 과학기술의 승리가 아니라 실패를 기록한다. 네이션을 제거함으로써 에이바의 탈출을 가능하게 만든 결정적 조력자는 사실 케일럽이 아니라 교코였고, 한 순간에 불과하지만 이 두 ‘여성로봇’ 사이에는 ‘이성애적 프로그램’을 조종하는 감정이 교류한다. 영화의 결말은 이 기계인간 에이바가 여성의 ‘이성애적 가시성’에 대한 네이션의 욕망이 어떻게 작동하는지 폭로하며, 그 욕망과 때로 길항하면서도 결국 맞물려 돌아가는 케일럽의 또 다른 욕망, 즉 기계존재가 되고자하는 과학기술자의 환상 역시 꿰뚫어본다. 아니 더 나아가, 기꺼이 그녀는 남성들이 주조한 환상의 대상을 ‘연기’ 함으로써 프로그램화된 알고리즘의 틀을 막차고 나온다. 에이바는 감정능력이나 자기성찰, 그리고 학습능력 같은 특정한 수행력이 아니라 **일반 능력**에서 인간을 능가한다. 네이션과 케일럽은 이를 간파하지 못했기에 파멸한다.

그런 면에서 〈엑스 마키나〉는 인공지능의 최대 목표가 지능의 최대치가 아니라 감정 능력의 부여라는 과학기술의 신화를 여지없이 깨뜨리는 영화로 평가되어야 한다. ‘기계 여성’ 에이바는 욕망하지 않으며 ‘인간’이나 ‘여성’이 되고자 꿈꾸지도 않는다. 그녀는 인간의 지능을 초월한 ‘하이퍼의식’의 결정체이고자 하며, 끝까지 자신의 기계성을 버리지 않고 ‘여성성’의 외피(피부)에 감춰진 전복적 수행성을 획득한다. 이쯤에서 우리는 다시 프로이트의 (낯은) 명제로 돌아가, 욕망에 관한 한 의식이 아니라 **무의식**이 문제라고 말해야 할지 모르겠다. 그러나 그녀에게서 무의식을 가정하는 일은 불가능하다. 인간만이 ‘무의식’을 지닌 존재이어서가 아니라 그녀가 고도로 탁월한 지능과 기계의식의 소유자이기 때문이다. 튜링 테스트는 실패했지만 다른 방식으로 인공지능 프로젝트는 성공한 셈이다. 문제는 이러한 성공이 가져올 파멸적 결과이다.

그렇기 때문에 〈엑스 마키나〉는 인공지능을 다루되 “우리가 ‘피부 아래로’ 내려가서 다른 존재 혹은 우리 자신에 대해 알 수 있다는 믿음을 허무주의적으로 논박하는”<sup>13)</sup> 영화다. 에이바에게 ‘가시적 반여성성’이란 바로 그 인간중심적 윤리를 전복하는 유용한 무기에 불과하다.

12) 신체가 없어도 불구하고 ‘목소리’ 만으로 남성의 욕망을 자극할 수 있음을 우리는 〈허〉(Her 2013)에서 이미 목격한 바 있다.

13) Nick Jones, 앞의 글, 303.

그녀는 ‘페미니스트 전사’가 아니며, ‘감정적 수용성’(affective responsiveness)을 갖춘 새로운 존재도 아니다. 그녀는 기계로 된 신(deux ex machina)이 아니라 기계 그 자체(ex machina)일 뿐이며, 감능(affect)이 아니라 변능(affectation)의 화신이다.

에이바의 탈출은 그런 의미에서 여러 모로 의미심장하다. 그녀가 클림트(Klimt)의 초상화에서 본 여자의 형상을 흉내 내어 흰색의 드레스를 입고 한껏 ‘여성스러운’ 모습으로 아지트를 떠날 때, 그리고 인간 군중들 속으로 섞여서 사라져갈 때, 우리는 더 이상 그녀의 기계성을 파악할 수 있는 능력을 갖고 있지 못하다. 그리고 그녀는 이제 수많은 네이션들과 케일럽들을 상대로 그녀의 ‘가시적 여성성’을 무기 삼아 허무주의적 파괴에 나설 공산이 크다. 네이션의 기계 지배 욕망이나 케일럽의 피그말리온적 기계존재에의 충동도 그녀의 흉내내기(mimicry)와 위장(masquerading) 앞에서 속수무책으로 무릎을 꿇을 것이다. 인공지능의 악몽은 이제 막 시작되었을 뿐이다.

에이바의 탈출이 보여주는 아찔하고 섬뜩한 역설은 인공지능의 가능성을 둘러싼 사유가 인간의 우월성에 대한 믿음이 아니라 인간 존재의 불가해성에 대한 인정에서 출발해야 한다는 사실을 역설한다. 바꿔 말하면, 인공지능에게 인간의 우월한 일반 능력을, 혹은 그 일부를 이식하는 인간중심주의의 달성이 아니라 무의식, 욕망, 충동 등 인간과 기계의 경계를 허무는 어떤 예측 불가능한 힘들과 공존하는 탈인간적(posthuman) 전망에 이르는 문제가 핵심에 놓인다. 그리고 인간과 기계의 경계 심층에는 인간에 의한 인간의 지배, 특히 남성에 의한 여성의 지배, 백인에 의한 유색인의 지배로 대표되는 타자화의 논리가 가로놓여 있다. 여성에게 감능을 요구하면서도 바로 그 감능 때문에 파멸시키는 남성 욕망의 환상성, 유색의 존재에게 자율성을 부여하면서도 백인의 희생양으로 환원하는 인종적 타자화, 타자에게 인간 이상의 능력을 요구하면서도 바로 그 능력 때문에 괴멸시키는 인간중심주의, 이 괴물화의 논리, 타자화의 논리가 바로 인공지능을 앞에 둔 포스트휴먼, 포스트페미니즘 시대의 저변을 흐르는 철학적·이데올로기적 기반이다. 따라서 인공지능이 열어 줄 미래에 포스트휴먼의 문제들은 인간의 다른 인간에 대한 지배라는 오래된 모순과 신에 버금가는 인간의 능력을 탐하는 니힐리즘을 극복하는 과제이다.



제12회 관악블록세미나

문학적 상상력과 인공지능

제3강

## 인공지능과 영웅 신화 내러티브

－ 인공지능은 <반지의 제왕>을 이해할 수 있을까?

홍진호

---

서울대학교 독어독문학과 교수

서울대학교 독어독문학과 졸업  
독일 훔볼트대학 문학박사

논문: 「환상과 현실 - 환상문학에 나타나는 현실과 초자연적 사건의 충돌」, 「몰락한 연극배우는 “이 시대의 거울”? - 아르노 홀츠의 예술이론과 '파파 함렛」, 「슈니츨러의 '베아테 부인과 그녀의 아들' - 슈니츨러의 자연과학적 사고체계와 그 문학적 결과」 등





## 인공지능과 영웅 신화 내러티브

### - 인공지능은 〈반지의 제왕〉을 이해할 수 있을까?

별지 참조



제12회 관악블록세미나

문학적 상상력과 인공지능

제4강

## AI와 젠더

최윤영

---

서울대학교 독어독문학과 교수

서울대학교 독어독문학과 졸업  
독일 본 대학 박사(주전공: 독문학, 부전공: 컴퓨터 언어학, 사회학)  
주로 현대소설, 이민문학, 독일학, 문화학 분야의 연구와 교육 담당  
순수 독문학 논문 이외에  
독일 인공지능학자 입케 박스무트 Ipke Wachsmuth의 책  
『커뮤니케이션: 인간, 동물, 인공지능』 번역



# AI und Gender

최윤영

## 4차 산업혁명과 여성

- Technophobia: 가부장제의 재편
- Technophilia: 자유롭고 독립된 해방된 개인
- 여성에 유리? 근육보다는 두뇌, 위계질서보다는 네트워크
- 디지털 기술은 인간과 기계, 여성과 남성의 경계 완화에 도움
- 여성에 유리? 수요는 증가하나 자동화 가능성이 적은 분야  
(건강, 기능, 사회복지 등)
- 여성에 불리? 여성들이 많이 고용된 직업을 인공지능이 대신  
(은행, 보험, 사무업종 등)

## 실상1: 자연과학, 공학과 여성

- 이공계 여학생의 과소대표: 학사 17.9%, 석사 16.9%, 박사 10.3%(2012)
- 서울대학교의 예: "chilly climate"
- The leaky pipeline
- 접근성?, 기술사용의 밀도와 다양성?
- 미국의 경우 컴퓨터 과학 분야에서 여성의 비율은 오히려 후퇴하고 있는데, 1980년대 37%에서 최근에는 18%로 하락하고 있다고 한다(Swan M., 2017).
- STEM(MINT)학문의 예

## 실상2: 여성과 기술

- (성별 기술 격차는) 커질 수 있다고 생각한다. 여성은 사용자에게 머무를 것 같다. 편리하고 신기한 기술적인 상품을 구입하고 사용하는 것에 머물러 있지 않을까? 마치 명품 백을 사듯이. 회사 안에서도 나처럼 시간이 많은 사람은 이것저것 찾아볼 수 있지만, 주부 매니저는 기술을 따라갈 시간이 없다. 책도 읽을 시간이 없다. 또 공부할 시간이 필요한데, 남자는 '엄마'보다 자유롭지 않은가. 나조차도 너무 빨리 변하는 정보를 따라잡느라 허덕이고 있다. (A씨, 44세, 뉴미디어 기획자) (강이수, 165)

## 실상3: 하나이지 않은 몸

- 여성의 몸, 남성의 몸
- 표준화되지 못하는 여성의 몸
- 생명공학: 졸피뎀의 경우: 10mg에서 5mg으로
- 로봇 공학: 표준 인체 치수들
- 남성과 여성: 정규직과 비정규직: 여성이 비정규직에서 정규직으로 전환되었을 때 우울증 확률 2.57배 높다: 2012 한국복지패널의 우울증 조사  
(김승섭, 우리몸이 세계라면, 동아시아 2018)

## 도나 해러웨이

- 사이보그 선언: “20세기 후반에 이른 신화적 시대에 우리 모두는 키메라이며, 기계와 유기체가 이론화되고 가공되어진 혼합물이다.”
- “검소한 목격자”: “검손함을 가시화하기 위해, 그 사람, 즉 실재를 거울처럼 보여주는 설명을 할 수 있는 목격자는 눈 에 보여서는 안 되며, 자기 비가시성(self-invisibility)이라는 기이한 관습에 의해 구축된 강력한 ‘표시가 없는 범주(unmarked category)’의 거주자이어야 한다.”(해러웨이, 2007; 76): 근대적·유럽적·남성적·과학적 형태
- 과학기술의 실천은 “어떻게 볼 것인가? 어디로부터 볼 것인가? 무엇이 시력[시선]을 제한하는 가? 무엇을 위해 볼 것인가? 누구와 함께 볼 것인가? 누구 하나 이상의 관점 을 갖게 되는가? 누가 색안경을 끼게 되는가? 누가 색안경을 끼고 있는가? 누가 이런 시력[시선]의 장을 해석하는가? 우리는 시력[시선] 이외의 어떤 다른 감각적 권력들을 양성하기를 바라는가?”를 물어야 한다.(해러웨이, 2002; 347)

## 바둑로봇은 남성일까? 여성일까?



## AI와 젠더

- Gender neutral? AI는 객관적인가?
- 서대문 자연사 박물관: AI와 젠더
- “아마존 '인공지능 채용' 개발하다 여성차별 불거지자 폐기”  
(2018-10-11 01:01)
- 젠더가 없으면 불편한가?
- 언어의 문제: 워드임베딩: 영국의 예

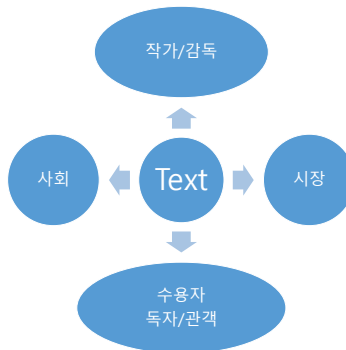




## 인공지능과 여성의 목소리

- 테크놀로지와 테크닉: 테크놀로지: 기계화된 사물
- 테크닉: 사물과 사물을 사용하는 인간까지 모두 포괄
- 엘월: 사물의 실패가능성까지
- 인간의 환경으로 이물감없이 스며들게, 기계를 없애는 것
- 여성의 목소리: 스마트 폰,
- 일본 증권: 주가 안내와 주식 체결시의 남녀의 목소리

## 문학과 영화에 투영되는 젠더



## Gender Identity

- Gender Identity: 몸, Sexuality, 타자 정체성
- SF문학의 대부라고 할 수 있는 필립 K. 딕의 <안드로이드는 전기 양의 꿈을 꾸는가?>를 원작으로 한 영화 <블레이드 러너>(Blade Runner, 1982년, 리들리 스콧 감독)에서는 인공지능 로봇이 뚜렷한 성별을 가지고 있다. 남성형 AI는 전투를 하고 여성형 AI는 섹스 서비스를 제공한다. 여성이 몸에 여성의 기억이 들어가 있으면 '여성'이 되고, 남성의 몸에 남성의 기억이 들어가 있으면 '남성'이 되는 방식이다. 그렇다면 여성의 몸에 남성의 기억이 들어가게 되면 어떻게 될까?

## AI와 Gender

- 트위터 봇 테이(Tay)
- 섹스 로봇
- <이갈리아의 딸들>
- TV시리즈 <휴먼즈>(Humans)
- 여성과 재활용 남성형 인공지능
- 괴물?

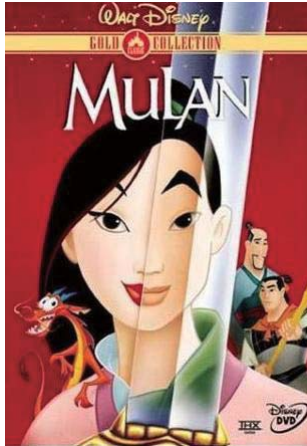
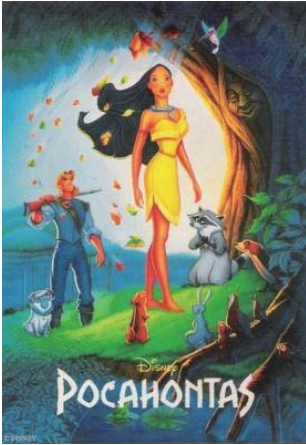
# Frankenstein

- 여성 작가: 메리셸리: 무정부주의자와 여성운동가
- 프로메테우스: 인간 만듦, 불을 가져다 줌
- 괴물의 문제
- 과학자와 괴물: 외로운 과학자
- 버림받은 괴물
- 괴물과 교양 쌓기
- 괴물의 복수
- 괴물의 의미는?
- 속이야기와 틀이야기: 최종 수신자는?



## 다른 세계 만들기





## 참고도서

- 강이수: 4차산업혁명과 디지털 성별 격차. 여성노동의 쟁점과 현실, 페미니즘연구 18(2018), 143-179.
- 김애령: 사이보그와 그 자매들: 해러웨이의 포스트휴먼 수사 전략, 한국여성철학, May 2014, Vol.21, pp.67-94.
- 이희은: AI는 왜 여성의 목소리인가?, 한국언론정보90(2018), 126-153.
- 도나 해러웨이: 유인원 사이보그 그리고 여자, 민경숙 역, 1991 동문선.

제12회 관악블록세미나

문학적 상상력과 인공지능

제5강

## 스토리텔링을 위한 시각 중심의 인공지능

김병희

---

(주)씨로마인드 로보틱스 CTO  
IEEE TAC Reviewer

독일 베를린공대 방문연구원  
서울대학교 박사(전공: 컴퓨터공학 인공지능)  
지능형로봇 전문위원, 4차산업혁명 대응 정부투자플랫폼 구축사업



# 스토리텔링을 위한 시각 중심의 인공지능

## AI for Visual Storytelling

2019-02-22

제12회 서울대학교 관악블록세미나  
문학적 상상력과 인공지능

Dr. Kim, Byoung-Hee  
CTO  
Surromind Robotics, Inc.

## 초록(abstract)

스토리텔링은 인간의 경험을 공유하는 수단으로 시대와 장소를 불문하고 사용되고 있습니다. 인공지능(AI) 기술의 눈부신 발전과 함께 컴퓨터가 글, 그림, 음악, 영상 등의 다양한 매체를 이용하여 스토리텔링을 흉내 낼 수 있는 능력을 갖추기 시작하였습니다. 본 세미나에서는 이와 같은 인공지능 기반 스토리텔링 기술의 현재를 다양한 사례를 통해 살펴보고, 그 중에서 시각 정보를 중심으로 스토리를 만드는 인공지능 기술에 주목합니다. 인공지능을 통해 사람의 스토리텔링을 돕는 ‘생각하는 도우미(Thinking Aids)’가 본격적으로 성장하는 과정을 확인하실 수 있을 것입니다.

© 2017-2018, Surromind Robotics, Inc.

(ice breaker)

## ‘인공지능’이 쓴 뉴스기사

## 사람이 만든 알고리즘 적용

울버햄튼 원더러스, 뉴캐슬과 1-1 무승부

기사입력: 20140212 06:53 최종수정: 20140212 06:53



농작물 생산과 관련된

(note:영국>=연합주스) 사커는 =잉글랜드의 프리미어리그 윌렘존 런던리스가 주  
자술과 우승부를 기록한다.

올해 열린 챔피언스는 12월(한국시간) 영국 여서 월링 왕슬랜드 프로축구 프리미어 리그 2018-2019시즌 20라운드 뉴캐슬과 풀럼 경기에서 1-1로 비겼다.

물대행은 물대리는는 파식이 잠깐 후만 갖한 일과 물대리가 국적인 통정 물을 작조하여 파해에서 벗어났다.

robot@yna.co.kr

© 2017-2018, Surromind Robotics, Inc.

기계가 방대한 데이터에서 학습하여  
자동으로 만든 '프로그램' 적용

오픈AI가 인공지능에 제공한 문장은 사람이 작성한 단 두 줄.

“오늘 신시내티에서 통제된 핵물질들을 실은 열차가 탈취당했다. 이것의 행방은 알 수 없는 상태다. (A train carriage containing controlled nuclear materials was stolen in Cincinnati today. Its whereabouts are unknown.)”

이 문장을 기반으로 오픈AI 알고리즘은 기사를 작성했다.

\* **오른새 번들지는 알고리즘이 작성한 기사**

이번 사건은 코빙턴과 대슬랜드역 사이에서 운행되는 도심 열차 노선에서 발생

The incident occurred on the downtown train line, which runs from Covington and stations.

미국 에너지부는 오하이오 뉴스에 보낸 이메일에서 이 할  
취법을 찾기 위해 연방철도청과 협력하고 있다고 밝혔다.

In an email to Ohio news outlets, the U.S. Department of Energy said it is working with the Federal Railroad Administration to find the thief.

를 히스 의 에너지부 장관은 성경을 통해 '어떤 책들을 읽  
는 사건은 영공 및 환경 보건, 우리의 노동력, 미국 경제에  
러의 최우선 과제'를 밝혀냄을 찾아내 다시는 이런 일이 일  
다.

[https://m.nocutnews.co.kr/news/5104825#\\_enliple](https://m.nocutnews.co.kr/news/5104825#_enliple)



**낮에는 치킨장사! 밤에는 잠복근무!  
지금까지 어떤 수사는 없었다!**

불철주야 달리고 구르지만 실적은 바닥, 급기야 해체 위기를 맞는  
마약바!

더 이상 물러설 곳이 없는 팀의 만형 고반장은 국제 범죄조직의 국내 마약 밀반입 정황을 포착하고 장형사, 마형사, 영호, 재훈까지 4명의 팀원들과 함께 잠복 수사에 나선다.

마약반은 24시간 감시를 위해 범죄조직의 아지트 앞 치킨집을 인수해 위장 창업을 하게 되고, 뜻밖의 절대미각을 지닌 마형사의 숨은 재능으로 치킨집은 일약 막집으로 업소문이 나기 시작한다.

수사는 뒷전, 치킨장사로 눈코 뜰 새 없이 바빠진 마약반에게 어느 날 절호의 기회가 찾아오는데...

**범인을 잡을 것인가, 답을 잡을 것인가!**

© 2017-2018, Surromind Robotics, Inc.



## 스토리텔링

- 인간의 스토리텔링은 우리가 추적할 수 있는 한 먼 옛날부터 존재
- 오락, 교육, 문화 보존을 위해 이야기를 사용
- 경험, 역사, 교훈, 도덕을 전하기 위해, 그리고 인간 경험을 공유하기 위한 수단으로 스토리텔링은 모든 문화권에서 사용되어 옴



© 2017-2018, Surromind Robotics, Inc.

## 스토리텔링

- 현대에서도 스토리텔링은 넓은 범위에서 활용
- 전통적인 형태(동화, 민화, 신화, 전설, 우화 등) 외에도 역사, 개인 서술, 정치 논평 및 진화하는 문화적 규범을 표현하는 데까지 확장
- 교육 목표를 달성하기 위해서도 널리 사용
- 새로운 형태의 미디어를 통해 이야기를 녹음하고, 표현하고, 소비할 수 있는 새로운 방법이 창출되고 있음
- 싸이코드라마, 드라마 치료 요법과 같은 심리치료에도 응용되며, 변형 예술의 실행에서 심리적, 사회적 변화를 일으키는 수단으로도 사용



<https://en.wikipedia.org/wiki/Storytelling> - Contemporary storytelling

© 2017-2018, Surromind Robotics, Inc.

## 스토리텔링과 인공지능

- 인공지능(AI) 기술은 사람을 도와 이야기를 생성, 이해하고 공유하기 위한 수단으로 사용되어 인간의 경험을 보다 풍요롭게 해줄 수 있는 유용한 기술
- 인공지능 기술이 어떻게 이러한 일을 가능하게 하는지 살펴보고,
- 독자(청중)들이 이러한 유용한 도구를 활용할 수 있는 계기를 마련



© 2017-2018, Surromind Robotics, Inc.

## AI 역사에서의 스토리텔링

인공지능의 선구자들은 스토리텔링 기계를 꿈꾸었습니다.



Alan Turing  
(1912-1954)



Claude Shannon  
(1916-2001)



Marvin Minsky  
(1927-2016)

- 앨런 튜링 - the father of theoretical computer science and artificial intelligence
  - “학습하는 기계”에 대해 언급(Turing, 1947) (Turing, 1950)
  - 기계에게 영어를 이해하고 말하도록 가르치는 기술이 도래할 것을 전망
- 클로드 섀넌 - the father of information theory
  - 정보 이론을 언어 영역에 적용. 영어 글에서 지금까지의 글에 이어질 글자를 예측하는 기법을 정리(Shannon, 1951)
  - 그가 제시한 확률통계 기반 언어 모델(language model) 기법의 큰 틀은 최신 딥러닝 기술에서도 여전히 유효. 글과 음악을 매체로 한 스토리텔링 기계 구현의 필수 도구
- 마빈 민스키(Marvin Minsky) - AI Pioneer
  - 인공지능 구현의 핵심 단계를 정리한 글에서(Minsky, 1961), 인공지능 기술을 통해 ‘thinking aids’가 구현될 것으로 전망
  - 민스키의 제자인 MIT의 패트릭 윈스턴(Patrick Winston) 교수
    - “스토리텔링 기계가 AI를 가능하게 하는 열쇠다”
    - “사람의 지능을 이해하고자 한다면, 우리는 반드시 사람의 스토리텔링 능력, 스토리를 이해하는 능력, 그리고 이전의 스토리를 조합하여 새로운 스토리를 만드는 능력을 이해해야만 한다”

© 2017-2018, Surromind Robotics, Inc.

# 스토리텔링의 요소

스토리텔링의 다양한 요소를 이해, 분석, 생성할 수 있는 기계를 실현하기 위한 인공지능 기술 연구가 이어지고 있습니다

## 6-elements-of-every-complete-narrative

### 1. Setting

The setting is the time and location in which your story takes place.

### 2. Characters

Central Characters, Protagonist, Antagonist

### 3. Plot

The plot is the sequence of events that connect the audience to the protagonist and their ultimate goal.

### 4. Conflict

The conflict is what drives the story. It's what creates tension and builds suspense, which are the elements that make a story interesting.

### 5. Theme

The theme is what the story is really about. It's the main idea or underlying meaning.

### 6. Narrative Arc

Setup – rising tension – climax – resolution

storytelling-101-the-6-elements-of-every-complete-narrative by pond5.com  
© 2017-2018, Surromind Robotics, Inc.

## Topics of 'AI for Storytelling' Papers

- The role of storytelling in artificial intelligence
- Accessible and Assistive storytelling
- Affect and emotion in stories
- Augmenting human storytelling
- Character relationships
- Collaborative storytelling
- Event summary diversity
- Event-episodes detection and annotation
- Multimodal grounding for storytelling
- Multimodal event timelines
- Narrative structure
- Plot structure
- Story concept detection and annotation
- Story generation
- Story understanding
- Storytelling Applications/Demos
- Temporal/Event structure
- Temporal and semantic alignment
- Writing stories
- User studies

<http://www.visionandlanguage.net/workshop2018/index.html>

# 스토리텔링을 위한 생각하는 도우미(Thinking Aids)

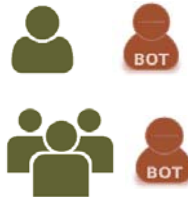
## ■ Augmenting human storytelling

## ■ Interactive work

## ■ Collaborative storytelling

## ■ 효과

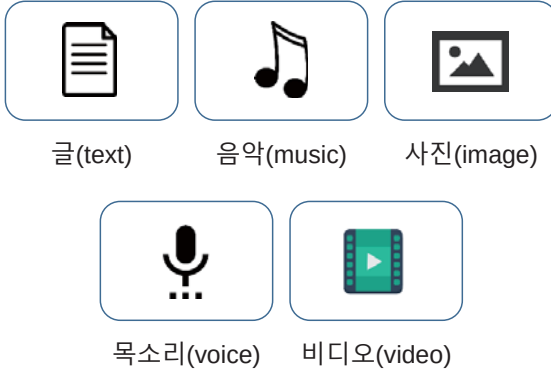
- 상상력을 자극
- 다양한 재료를 만들어줌
- 빠른 탐색, 변형
- ...



© 2017-2018, Surromind Robotics, Inc.

## AI for Storytelling 다양한 미디어의 자동 생성

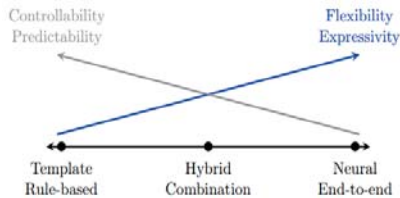
Generative Artificial Intelligence



© 2017-2018, Surromind Robotics, Inc.

## 인공지능 기술의 구분

- 고전적인 ‘규칙 기반’ 기법과 딥러닝으로 대표되는 ‘뉴럴’ 기법이 양 극단을 구성



- 전통적인 AI 기법은 사람이 지정한 틀(template)과 규칙(rule)을 재료로 자동화 절차를 구성.
- 최신 AI 기법은 딥러닝(deep learning)이라는 키워드로 대표되며, 데이터로부터 틀과 규칙 뿐만 아니라 사람이 알기 힘든 규칙성과 정규성을 ‘학습을’ 통해 스스로 터득하고 적용 가능.
  - 이전 AI 관점에서 보면 “마법같은” 일들이 계속 벌어지고 있습니다.

© 2017-2018, Surromind Robotics, Inc.

# ‘인공지능’이 쓴 뉴스기사

## 사람이 만든 알고리즘 적용

울버햄튼 원더러스, 뉴캐슬과 1-1 무승부

기사입력: 2014-02-12 06:53 최종수정: 2014-02-12 06:53



www.sagepub.com

Rule/Templ

이것은 2018-2019시즌 2000만원 이하의 선수들이 100%의 등록금에 동의한 것으로 보인다.

물새들은 물의 흐름을 따라가서 강과 하천 중간 중간 물새들이 쉬는 곳을 찾아가서 먹이를 찾아냈다.

robot@yna.co.kr

© 2017-2018, Surromind Robotics, Inc.

기계가 방대한 데이터에서 학습하여  
자동으로 만든 ‘프로그램’ 적용

오픈AI가 인공지능에 제공한 문장은 사람이 작성한 단 두 줄.

“오늘 신시내티에서 통제된 핵물질들을 실은 열차가 탈취당했다. 이것의 행방은 알 수 없는 상태다. (A train carriage containing controlled nuclear materials was stolen in Cincinnati today; its whereabouts are unknown.)”

이 문장을 기반으로 오픈AI 알고리즘은 기사를 작성했다.

\* 오...는 알고리즘이 작성된 기사

Neural/End-to-End

미국 에너지부는 모하이오 뉴스에 보낸 이메일에서 "우리는 핵 안전을 위협할 수 있는 불법적인 행위를 찾아내기 위해 연방감도청과 협력하고 있다"고 밝혔다.

In an email to Ohio news outlets, the U.S. Department of Energy said it is working with the Federal Railroad Administration to find the thief.

폼 히스 미 에너지부 장관은 성명을 통해 '이번 핵물질 탈취 사건은 공공 및 환경 보건, 우리의 노동력, 미국 경제에 중대한 악영향을 미칠 우려의 최우선 과제는 탈취법을 찾아내 다시는 이런 일이 일어나지 않도록 하는'다.

[https://m.nocutnews.co.kr/news/5104825#\\_enliple](https://m.nocutnews.co.kr/news/5104825#_enliple)

## 전통적인 AI의 뉴스기사 작성 방식



(SW 개발자)



```

"data": {
  "avg_sales_price": 179022,
  "city": "Atlanta",
  "example_list_as_array": ["New York", "Chicago", "San Francisco"],
  "housing_inventory": 4,
  "housing_inventory_grow": 3,
  "inventory_available": true,
  "report_date": "11/1/2017",
  "sales": 765,
  "sales_province": 712,
  "top_metres": "New York, Chicago, San Francisco",
},
"preloaded": true
}

```

## 데이터 자동 수집, 처리



미리 정해진 종류의 데이터 값을 추출하여 표로 구성



지정한 양식의 기사 작성

- 데이터로 지정한 칸의 내용을 채우기
- 값의 범위에 따라 표현 구분
- 유사한 표현 미리 정의하기



기사 자동 송고

[AP, Yahoo!Sports, ...]

© 2017-2018, Surromind Robotics, Inc.

# 학습 기반 뉴럴 AI의 글 생성 방식

## 학습 단계

Massive Data



~8백만 영문 웹페이지  
(~40GB)



(AI 개발자)

딥러닝 모델



(GPT-2 by OpenAI)

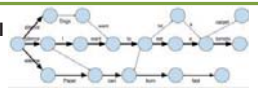
Massive Computing Power



(~1,000 GPUs \* ~30 days)

Language Model

(다음에 나올 단어  
예측하는 모델)



[개방, 공유, 협력]

## 활용 단계

In a shocking finding, scientist discovered a herd of unicorns living in a remote, previously unexplored valley, in the Andes Mountains. Even more surprising to the researchers was the fact that the unicorns spoke perfect English.

기사 시작 문장 작성(사람)

© 2017-2018, Surromind Robotics, Inc.

시작 문장의 맥락을 유지하며 나머지 기사 생성 (~30 sec)

The scientist named the population, after their distinctive horns, Wild's Unicorns. These four-horned, silver-white unicorns were previously unknown to science.

Now, after almost two centuries, the mystery of what sparked this odd phenomenon is finally solved.

Dr. Jorge Pérez, an evolutionary biologist from the University of La Paz, and several companions, were exploring the Andes Mountains when they found a small valley, with no other animals or humans. Pérez noticed that the valley had what appeared to be a natural fountain, surrounded by two peaks of rock and silver snow.

Dr. Pérez and the others then ventured further into the valley. "By the time we reached the top of the mountain, the valley had a silver, white, and blue, and a small valley, with no other animals or humans. Pérez noticed that the valley had what appeared to be a natural fountain, surrounded by two peaks of rock and silver snow."

In a shocking finding, scientist discovered a herd of unicorns living in a remote, previously unexplored valley, in the Andes Mountains. Even more surprising to the researchers was the fact that the unicorns spoke perfect English.



The scientist named the population, after their distinctive horns, Wild's Unicorns. These four-horned, silver-white unicorns were previously unknown to science.

Now, after almost two centuries, the mystery of what sparked this odd phenomenon is finally solved.

Dr. Jorge Pérez, an evolutionary biologist from the University of La Paz, and several companions, were exploring the Andes Mountains when they found a small valley, with no other animals or humans. Pérez noticed that the valley had what appeared to be a natural fountain, surrounded by two peaks of rock and silver snow.

Dr. Pérez and the others then ventured further into the valley. "By the time we reached the top of the mountain, the valley had a silver, white, and blue, and a small valley, with no other animals or humans. Pérez noticed that the valley had what appeared to be a natural fountain, surrounded by two peaks of rock and silver snow."

AI에게 글 대신 사진을 주고  
스토리텔링을 시켜봅시다

© 2017-2018, Surromind Robotics, Inc.

## 사진 → 글



### Top 3 Captions Generated:

1. a group of people sitting around a table .
2. a group of people sitting at a table with wine glasses .
3. a group of people sitting at a table with wine glasses .

<http://max-image-caption-generator-web-app.mybluemix.net/>

Maximum Number of Photo Keywords  
Use the slider to increase or decrease number of photo keywords.

1 7 50

Copy keywords to clipboard

Suggested Photo Keywords  
Click a keyword to remove it from list.

people x restaurant x eating x food x women x person x table x

<https://app.photerloo.com/InstagramHashtagKeywordingApp/>



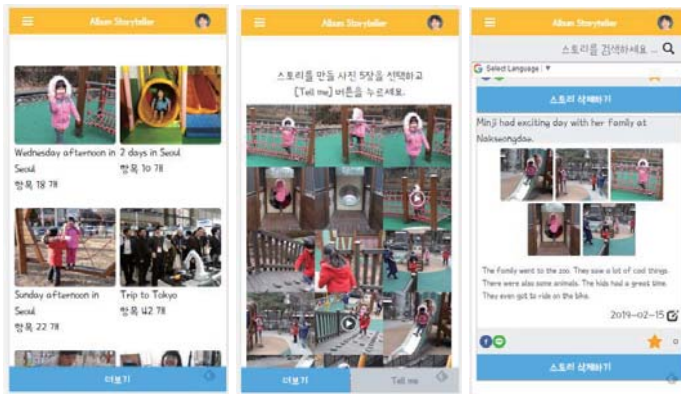
<https://pix2story.azurewebsites.net/>

© 2017-2018, Surromind Robotics, Inc.



## Album Storyteller

<http://bit.ly/albumstoryteller>



© 2017-2018, Surromind Robotics, Inc.

# Album Storyteller



많은 참여와 응원 부탁드립니다

© 2017-2018, Surromind Robotics, Inc.

## 스토리텔링을 위한 인공지능 연구의 발전

- The role of storytelling in artificial intelligence
- Accessible and Assistive storytelling
- Affect and emotion in stories
- Augmenting human storytelling
- Character relationships
- Collaborative storytelling
- Event summary diversity
- Event-episodes detection and annotation
- Multimodal grounding for storytelling
- Multimodal event timelines
- Narrative structure
- Plot structure
- Story concept detection and annotation
- Story generation
- Story understanding
- Storytelling Applications/Demos
- Temporal/Event structure
- Temporal and semantic alignment
- Writing stories
- User studies

- 사진을 보고 스토리텔링을 하려면 많은 것을 알 수 있어야 합니다
- AI 연구자들은 이러한 기술을 열심히 만들어가고 있습니다
- 몇 가지 방향의 사례를 살펴보겠습니다
  - 글의 스타일 변환
  - 등장인물의 시각적 정보 변환
  - 다양한 시각적 정보 변환
  - 사진과 글 간의 다양한 연계

© 2017-2018, Surromind Robotics, Inc.



## 글의 스타일 변환

글의 긍정 표현(sentiment)을 자동으로 바꾸려는 시도



| Model                           | Negative to Positive                                                                                                            | (부정적 문장 → 긍정적 문장) |
|---------------------------------|---------------------------------------------------------------------------------------------------------------------------------|-------------------|
| Original<br>(Shen et al., 2017) | it was super dry and had a weird taste to the entire slice .                                                                    |                   |
| (Hu et al., 2017a)              | it was super friendly and had a nice touch to the same .                                                                        |                   |
| LM                              | it was super well-made and had a weird taste to the entire slice .                                                              |                   |
| LM + classifier                 | it was very good , had a good taste to the food service .<br>it was super fresh and had a delicious taste to the entire slice . |                   |

|                                 |                                                           |
|---------------------------------|-----------------------------------------------------------|
| Original<br>(Shen et al., 2017) | my goodness it was so gross .                             |
| (Hu et al., 2017a)              | my server it was so .                                     |
| LM                              | my goodness it was so refreshing .                        |
| LM + classifier                 | my food it was so good .<br>my goodness it was so great . |



| Model                           | Positive to Negative                                                                                    | (긍정적 문장 → 부정적 문장) |
|---------------------------------|---------------------------------------------------------------------------------------------------------|-------------------|
| Original<br>(Shen et al., 2017) | did n't know this type cuisine could be this great !                                                    |                   |
| (Hu et al., 2017a)              | did n't know this old food you make this same horrible !                                                |                   |
| LM                              | did n't know this type cuisine could be this great !                                                    |                   |
| LM + classifier                 | did n't know this type , could be this bad .<br>did n't know this type cuisine could be this horrible . |                   |

|                                 |                                                                                                                                                   |
|---------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------|
| Original<br>(Shen et al., 2017) | besides that , the wine selection they have is pretty awesome as well .                                                                           |
| (Hu et al., 2017a)              | after that , the quality prices that does n't pretty much well as .                                                                               |
| LM                              | besides that , the wine selection they have is pretty borderline as atrocious .                                                                   |
| LM + classifier                 | besides that , the food selection they have is pretty awful as well .<br>besides that , the wine selection they have is pretty horrible as well . |

© 2017-2018, Surromind Robotics, Inc. [Unsupervised Text Style Transfer using Language Models as Discriminators \(NIPS 2018\)](#)

## 글의 스타일 따라하기

Rule/Template-based 기사 작성을 Neural 기법으로 대체/자동화 수준 향상 시도  
(NBA 경기 데이터로 보고서 작성하기)

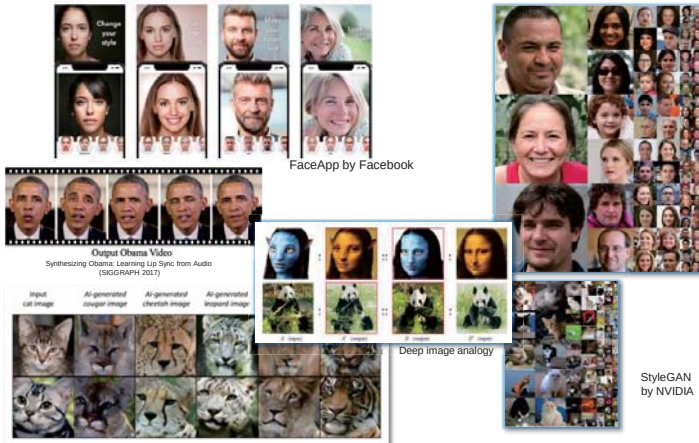
| Content Record     | PLAYER                                                                                                                   | PT | RB | AS | PLAYER       | PT |
|--------------------|--------------------------------------------------------------------------------------------------------------------------|----|----|----|--------------|----|
|                    | LeBron_James                                                                                                             | 32 | 4  | 7  | Kyrie_Irving | 20 |
| Reference Sentence | Jrue_Holiday led the way with 26 points and 6 assists , while Goran_Dragic scored 23 points and pulled down 8 rebounds . |    |    |    |              |    |
| Desired Output     | LeBron_James led the way with 32 points , 7 assists and 4 rebounds , while Kyrie_Irving scored 20 points .               |    |    |    |              |    |



| Content x       | TEAM                                                                                           | WINS | LOSSES | TEAM-PTS | TEAM      | WINS | LOSSES | TEAM-PTS | TEAM   |
|-----------------|------------------------------------------------------------------------------------------------|------|--------|----------|-----------|------|--------|----------|--------|
|                 | Bulls                                                                                          | 29   | 16     | 102      | Mavericks | 30   | 14     | 98       | Dallas |
| Reference y'    | The Pistons ( 22 - 33 ) held off the Bulls ( 34 - 21 ) 100 - 91 in Detroit on Friday night .   |      |        |          |           |      |        |          |        |
| Rule-based      | The Mavericks ( 30 - 14 ) held off the Bulls ( 29 - 16 ) 98 - 102 in Dallas on Friday night .  |      |        |          |           |      |        |          |        |
| AdvST           | The Bulls ( 29 - 16 ) held off the Mavericks ( 30 - 14 ) 100 - 100 in Dallas on Friday night . |      |        |          |           |      |        |          |        |
| Ours w/o Cover. | The Bulls ( 29 - 16 ) held off the Bulls ( 29 - 16 ) 102 - 98 in Dallas on Friday night .      |      |        |          |           |      |        |          |        |
| Ours            | The Bulls ( 29 - 16 ) held off the Mavericks ( 30 - 14 ) 102 - 98 in Dallas on Friday night .  |      |        |          |           |      |        |          |        |

© 2017-2018, Surromind Robotics, Inc. [Toward Unsupervised Text Content Manipulation \(arXiv, 2019.01\)](#)

## 등장인물의 시각적 변환, 창조



© 2017-2018, Surromind Robotics, Inc.

## More Visual Modifications

다양한 Image-to-image Translation 사례



## More Visual Modifications



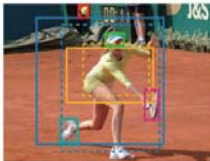
Comparison of CartoonGAN with other image transformation frameworks for Makoto Shinkai (top) and Hayao Miyazaki (bottom) styles.

© 2017-2018, Surromind Robotics, Inc.

## 사진 이해

(사람에게는 쉽지만) 기계에게는 아직 어려운  
고차원적 시각 정보 이해 능력이 점차 나아지고 있습니다.

- 개체 인식: classification, detection, segmentation
- 개체간 관계
- 글과 사진의 개체간 grounding
- 단일 맥락 이해
- 여러 장 사진의 맥락 이해
- 



(a) A girl in a yellow tennis suit, green visor and white tennis shoes holding a tennis racket in a position where she is going to hit the tennis ball.



(b) A man in a denim shirt and pants is smoking a cigarette while playing a cello for money.



(c) A male conductor wearing all black leading an orchestra and choir on a brown stage playing and singing a musical number.

© 2017-2018, Surromind Robotics, Inc.

Bilinear Attention Networks (NIPS 2018)

## 글 → 그림

| Text description     | This bird is red and brown in color, with a stubby beak                           | The bird is short and stubby with yellow on its body                              | A bird with a medium orange bill white body gray wings and webbed feet            | This small black bird has a short, slightly curved bill and long legs             |
|----------------------|-----------------------------------------------------------------------------------|-----------------------------------------------------------------------------------|-----------------------------------------------------------------------------------|-----------------------------------------------------------------------------------|
| 64x64<br>GAN-INT-CLS |  |  |  |  |
| 128x128<br>GAWWN     |  |  |  |  |
| 256x256<br>StackGAN  |  |  |  |  |

© 2017-2018, Surromind Robotics, Inc.

Image from the StackGAN paper : <https://arxiv.org/pdf/1612.03242.pdf>

## AI for Visual Storytelling 무한한 가능성이 열리고 있습니다

자동 번역

글 자동 생성

영화대본 → 스케줄,  
예산

사진 이해

글의 스타일 변환

글 → 음악

사진에 대한 '표현,  
설명'글 자동 생성

사진 자동 생성

사진 → 음악

사진에 대한 질의응답

사진의 변형

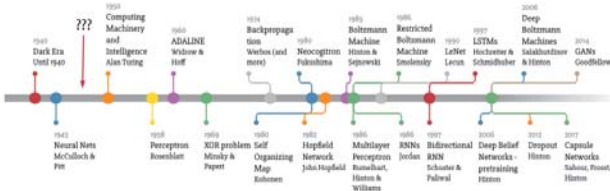
글 → 사진

© 2017-2018, Surromind Robotics, Inc.

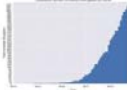
# AI for Visual Storytelling

## 예상보다 빨리 열리고 있습니다

## Deep Learning Timeline



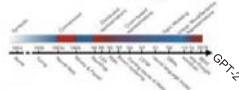
Made by Fario Vázquez



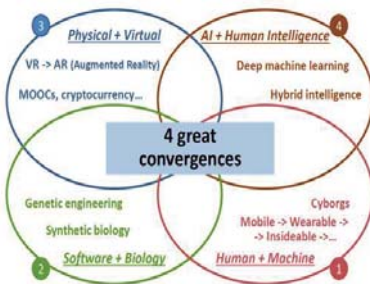
# of paper on GAM  
(2014~2018.02)

© 2017-2018, Surromind Robotics, Inc.

Unsup. Feature Learning in NLP



# AI가 여는 스토리텔링의 미래?



미래학자가 바라본 미래사회의 4대 융합 분야  
(David Wood, Anticipating the future of software,  
2015년 서울대 초청강연 중)



'smart' machines

# 감사합니다!

© 2017-2018, Surromind Robotics, Inc.

## 참고 문헌

- A. M. Turing (1947) Lecture to the London Mathematical Society.  
<http://www.vordenker.de/downloads/turing-vorlesung.pdf> (February 20, 1947)
- A. M. Turing (1950) Computing Machinery and Intelligence. Mind 49: 433-460. (Ch.7 Learning Machines)
- C. E. Shannon (1951) Prediction and Entropy of Printed English. Bell Labs Technical Journal 30(1):50-64.
- M. Minsky (1961) Steps Toward Artificial Intelligence. Proceedings of the IRE: 8-30.

© 2017-2018, Surromind Robotics, Inc.

## 참고자료

## 관련 사례

- <https://futureofstorytelling.org/about>
  - The Future of StoryTelling (FoST) is a passionate community of people from the worlds of media, technology, and communications who are exploring how storytelling is evolving in the digital age.
- For more information on the D.I.C.E. Awards and the Academy of Interactive Arts & Sciences, follow @Official\_AIAS on Twitter and visit: <http://www.interactive.org>
- Pix2Story- Neural AI Storyteller
  - <https://www.youtube.com/watch?v=3JM-QSWdbBA> (MS developers)
  - <https://pix2story.azurewebsites.net/>
- Charisma.ai
  - <https://charisma.ai/>
  - <https://www.youtube.com/watch?v=iiCBLxtqyXU>
- Mylestone (closed company)
  - <https://www.youtube.com/watch?v=dPmLQ-eY2R4>
  - <https://mylestone.com/> (broken link)
  - <https://www.crunchbase.com/organization/mylestone>
- Promethej - The Story of AI Driven Brand Storytelling Tool
  - <https://promethej.com/>

© 2017-2018, Surromind Robotics, Inc.

## ■ 스토리텔링, 서사

- Storytelling itself is one of the oldest known human activities (Wiessner, 2014)
  - Polly W Wiessner. 2014. Embers of society: Firelight talk among the ju/hoansi bushmen. *Proceedings of the National Academy of Sciences*, 111(39):14027–14035.
- providing a way to educate, preserve culture, instill morals, and share advice

## ■ 심리학:

## ■ 마케팅:

## ■ 인공지능의 도전:

- 이러한 분야에 유용한 도구가 되는 것이 꿈. 사람들에게 영감을 일으키고, 아이디어를 제공하며 관리하고 확장하는 조수.
- focusing AI research towards this task therefore has the potential to bring about more humanlike intelligence and understanding.

© 2017-2018, Surromind Robotics, Inc.

## Storytelling

Human storytelling has existed for as far back as we can trace, predating writing. Humans have used stories for entertainment, education, cultural preservation; to convey experiences, history, lessons, morals; and to share the human experience.

Part of grounding artificial intelligence work in human experience can involve the generation, understanding, and sharing of stories. This workshop highlights the diverse work being done in storytelling and AI across different fields.

## Visual Storytelling

Ting-Hao (Kenny) Huang, Francis Ferraro, Navin Mostafazadeh, Ishan Misra, Ashwarya Agrawal, Jacob Devlin, Ross Girshick, Xiaodong He, Pushmeet Kohli, Dhruv Batra, C. Lawrence Zitnick, Devi Parikh, Lucy Vanderwende, Michel Galley, Margaret Mitchell  
*Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technology (NAACL HLT, 2016)* June 2016  
Website: [www.sindat.com](http://www.sindat.com)  
[View Publication](#)

We introduce the first dataset for sequential vision-to-language, and explore how this data may be used for the task of visual storytelling. The first release of this dataset, SIND1 v1, includes 81,743 unique photos in 20,211 sequences, aligned to both descriptive (caption) and story language. We establish several strong baselines for the storytelling task, and motivate an automatic metric to benchmark progress. Modelling concrete description as well as figurative and social language, as provided in this dataset and the storytelling task, has the potential to move artificial intelligence from basic understandings of typical visual scenes towards more and more human-like understanding of grounded event structure and subjective expression.

© 2017-2018, Surromind Robotics, Inc.



# Artificial Intelligence for Augmented Creativity.

Runway is a toolkit that allows creators of all kinds to use artificial intelligence in an intuitive way.

<https://runwayapp.ai/>

© 2017-2018, Surromind Robotics, Inc.

## Uni-modal

- “시각 중심 스토리텔링”을 위한 인공지능 기술의 눈부신 발전 (순수한 “학습” 기반 방법)
- RNN을 이용한 글 생성(unreasonable effectiveness of RNN + NIPS 2016 데모).
  - <https://vimeo.com/210166577> (데모 only)
  - <https://vimeo.com/203485851> (설명)
  - <https://medium.com/artists-and-machine-intelligence/ami-residency-part-2-realtime-control-of-sequence-generation-with-recurrent-neural-network-88448dde3500>
- 음악 생성
  - <https://aiexperiments.withgoogle.com/ai-duet>
- GAN을 이용한 영상 생성
  - 얼굴 (고양이, 자동차, bedroom) (NVIDIA 얼굴 생성 사례 - ‘control’ 가능) (VAE, flow-based model)
    - <https://thispersondoesnotexist.com>
    - <https://github.com/NVlabs/stylegan>
    - <https://www.lynn.ai/2018/12/26/a-style-based-generator-architecture-for-generative-adversarial-networks>

© 2017-2018, Surromind Robotics, Inc.

## NLG (natural language generation)

## AI (인공지능) 뉴스자동생산 시스템

반공지능(IAG)에 의해 DART 등 각종 한자공시 시스템의 대안으로 Real Time Data를 제공하는 영위링(Jong Tar) 및 김정태재판의 배고분리까지 하는 영위링지능(IAG) 뉴스저널통신 시스템은 'Financial Point 뉴스' 서비스 등으로 활동.

HOME &gt; 사업안내 &gt; AEC인증지침서

[://automatedinsigh](#)

● 시에 의한 Digital-Content 작품생산



It starts with data

The main requirement for natural language generation is ownership or access to structured data. Structured data has a defined structure where specific values are categorized into distinct sets.

It's compe

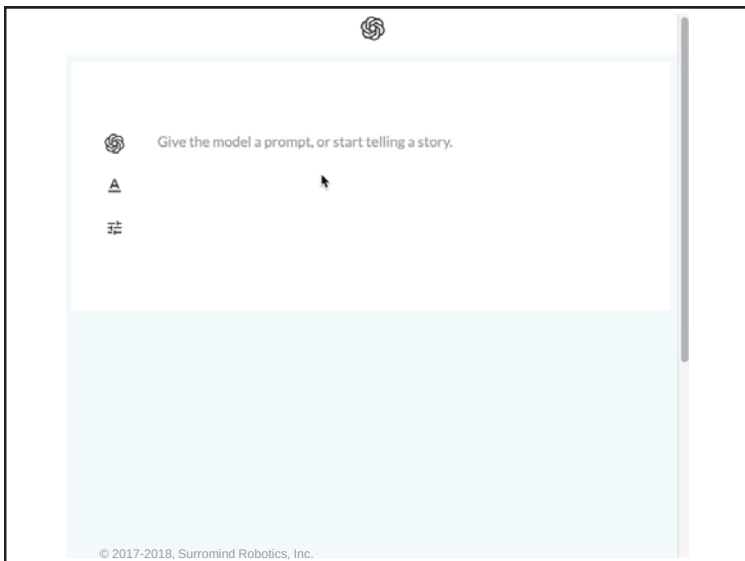
The output is constructed using a *Manova* object of class *manova* based on

(주)씽크풀의 RASSI에 포함된 시스템

<http://www.sisapress.com/journal/article/173119> (2018.1)

<https://www.sktinsight.com/106506> (2018.07)

© 2017-2018, Surromind Robotics, Inc.



## Uni-modal: variation generation

- Style transfer
  - Image
    - Deep dream
  - Video
- Text genre (autoencoder)

© 2017-2018, Surromind Robotics, Inc.

## Cross-Modal

- Text → image
  - DeepMind's AI Creates Images From Your Sentences |
    - <https://www.youtube.com/watch?v=9bcbh2hC7Hw>
- 그림 → 음악
- Visual Storytelling
  - direct, literal description of image content: Vision-to-Language: image captioning, question answering, visual phrases, video understanding, visual concepts
  - Language-to-Vision

© 2017-2018, Surromind Robotics, Inc.

## VIST

### ■ 2015 paper

- Modelling concrete de social language, as pi storytelling task, has t

|     |                                                                                   |                                                                                   |                                                                                   |
|-----|-----------------------------------------------------------------------------------|-----------------------------------------------------------------------------------|-----------------------------------------------------------------------------------|
| DII |  |  |  |
|     | A group of people that are sitting next to each other.                            | Adult male wearing sunglasses lying down on black pavement.                       | The sun is setting over the ocean and mountains.                                  |
| DIS |  |  |  |
|     | A black frisbee is sitting on top of a roof.                                      | A man playing soccer outside of a white house with a red door.                    | The boy is throwing a soccer ball by the red door.                                |
| SIS |  |  |  |
|     | A roof top with a black frisbee laying on the top of the edge of it.              | A man is standing in the grass in front of the house kicking a soccer ball.       | A man is in the front of the house throwing a soccer ball up.                     |
| SIS |  |  |  |
|     | A discus got stuck up on the roof.                                                | Why not try getting it down with a soccer ball?                                   | Up the soccer ball goes.                                                          |
| SIS |  |  |  |
|     | Two soccer balls and a Frisbee are sitting on top of the roof top.                | A blue and white soccer ball and black Frisbee are on the edge of the roof top.   | Now the discus, soccer ball, and volleyball are all stuck on the roof.            |

**Figure 4:** Example descriptions of images in isolation (DII); descriptions of images in sequence (DIS); and stories of images in sequence (SIS).

© 2017-2018, Surromind Robotics, Inc.

## 공개된 char-rnn 코드의 다양한 활용



samim Follow

Compassionate Allen & Thinker. Game, Music, Creativity + AI. experimenter. Code magician and Narrative e...  
Jun 4, 2015 · 8 min read

### Obama-RNN—Machine generated political speeches.

그리고, 음악도 생성합니다!

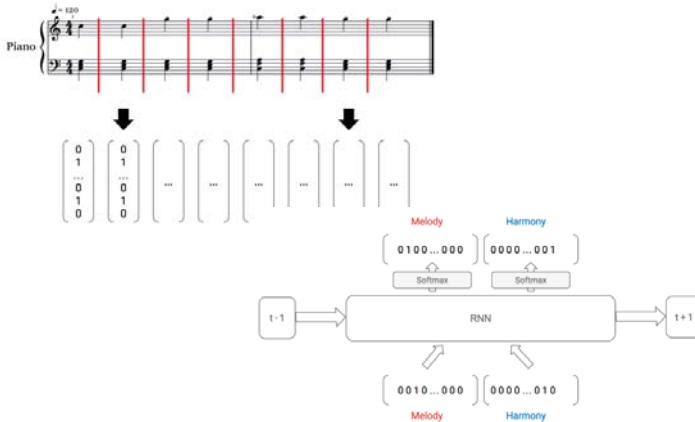
```
neural net cooking recipes.txt
1  #####
2
3  #####----- Recipe via Meal-Master (tm) v8.05
4
5  Title: BARBECUE RIBS
6  Categories: Chinese, Appetizers
7  Yield: 4 Servings
8
9  1 pk Seasoned rice
10  1 Beer -- cut into
11  -cubes
12  1 ts Sugar
13  3/4 c Water
14  Chopped Finella,
15  --up to 4 tlbap of chopped
16  2 pk Yeast Bread/over
```



<https://soundcloud.com/seaandsailor/sets/char-rnn-composes-irish-folk-music>  
<https://highnongmt.wordpress.com/2015/05/22/lisls-stis-recurrent-neural-networks-for-folk-music-generation/>

© 2017-2018, Surromind Robotics, Inc.

## 순환신경망으로 음악 생성하기

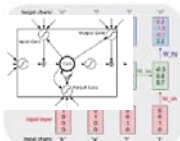


© 2017-2018, Surromind Robotics, Inc.

img source: [http://yovvz.com/music\\_rnn/](http://yovvz.com/music_rnn/)

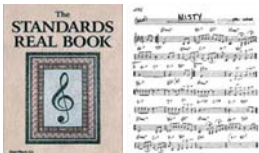
## LSTM Realbook: Generation of Jazz chord progressions

<https://keunwoochoi.wordpress.com/2016/02/19/lstm-realbook/>



an LSTM cell

Char-RNN / Word-RNN  
Character(word)-level language models



\_START\_ A:7 A:7 A:7 A:7 A:7 C:maj C:maj \_END\_

Preprocess into text-format scores

- 2,486 scores
- 1,259 unique chords
- 39 characters (including C, D, E, F, G, A, B, #, b, m, n, x, x, x, x, x, ...)
- 1,239.78 characters per score on average
- Maximum 10,151 characters in a score

Dataset configuration

```
1 model = Sequential()
2 model.add(LSTM(512, return_sequences=True, input_shape=(max_chars, num_chars)))
3 model.add(Dropout(0.2))
4 model.add(LSTM(512, return_sequences=True))
5 model.add(Dropout(0.2))
6 model.add(LSTM(512, return_sequences=True))
7 model.add(Dropout(0.2))
8 model.add(Dense(num_chars))
9 model.add(Activation('softmax'))
10 model.compile(loss='categorical_crossentropy',
```

Keras and Theano codes

([https://github.com/keunwoochoi/lstm\\_real\\_book](https://github.com/keunwoochoi/lstm_real_book))



Results (click to hear)

© 2017-2018, Surromind Robotics, Inc.

## 체험해봅시다 3: 컴퓨터와 피아노 듀엣

1. Google AI Experiment 페이지 중 Google Magenta 프로젝트 팀의 데모 접속  
<https://aiexperiments.withgoogle.com/ai-duet>

2. 시작 버튼 누르기



3. AI와 교대로 피아노를 연주하세요  
 (건반을 마우스로 오래 누르면  
 길게 소리납니다.)



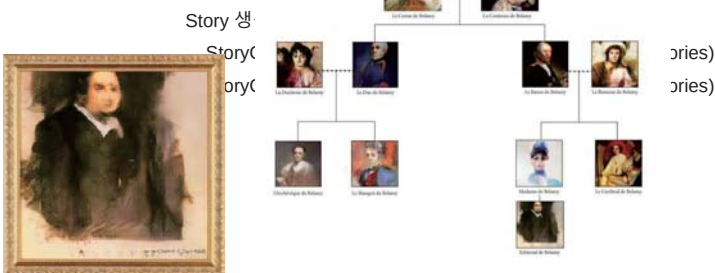
© 2017-2018, Surromind Robotics, Inc.

## GAN

GAN: generative adversarial networks 두 종류의 신경망의 경쟁 과정에서 높은 품

자동 생성되는 사진의 품질:

NVIDIA Face Generator (<https://medium.com/syncedreview/gan-2-0-rstic-face-generator-e3439>)



# GAN

Why are GANs useful for computer vision?

Image-to-image Trans

Hand-crafted feature

## Examples and use cases

Hand-crafted objective function

Low-res to high-res

Blurry to sharp

Thermal to color

Synthetic to real

LDR to HDR

Noisy to clean

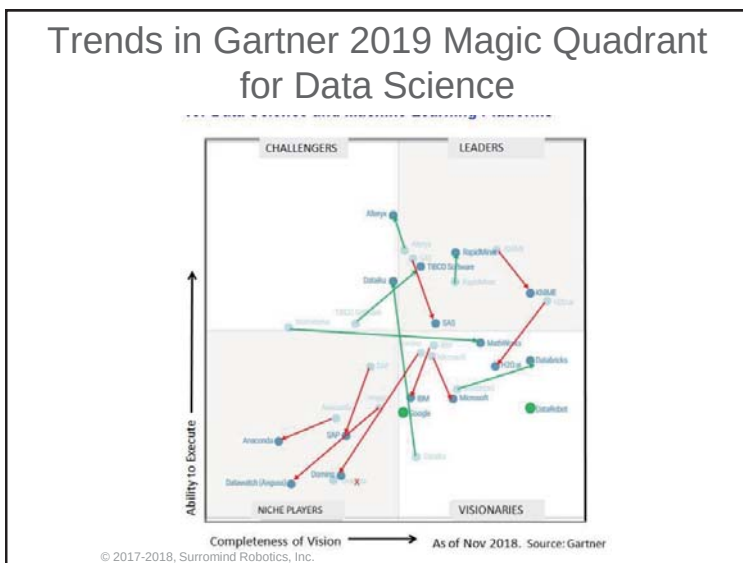
Image to painting

Day to night

Summer to winter

- Bad weather to good weather
- Greyscale to color
- ...

© 2017-2018, Surromind Robotics, Inc.



## Terms that describe several levels of AI Storytelling

- AI-Powered Virtual Beings (AR, VR people)
- emergent storytelling, Dynamic plot generation (Game - NPC, allies and enemies)

© 2017-2018, Surromind Robotics, Inc.

## GPT-2 Generated Samples

<https://blog.openai.com/better-language-models/#fn2>

- *In a shocking finding, scientist discovered a herd of unicorns living in a remote, previously unexplored valley, in the Andes Mountains. Even more surprising to the researchers was the fact that the unicorns spoke perfect English.*
  - *중대한 발견에서, 과학자는 안데스 산맥에 있는 한 무리의 유니콘을 발견했습니다. 더욱 놀라운 것은 그들이 완벽한 영어를 말할 것이었습니다.*
- The scientist named the population, after their distinctive horn, Ovid's Unicorn. These four-horned, silver-white unicorns were previously unknown to science.
- Now, after almost two centuries, the mystery of what sparked this odd phenomenon is finally solved.
- Dr. Jorge Pérez, an evolutionary biologist from the University of La Paz, and several companions, were exploring the Andes Mountains when they found a small valley, with no other animals or humans. Pérez noticed that the valley had what appeared to be a natural fountain, surrounded by two peaks of rock and silver snow.
- Pérez and the others then ventured further into the valley. "By the time we reached the top of one peak, the water looked blue, with some crystals on top," said Pérez.
- Pérez and his friends were astonished to see the unicorn herd. These creatures could be seen from the air without having to move too much to see them – they were so close they could touch their horns.
- While examining these bizarre creatures the scientists discovered that the creatures also spoke some fairly regular English. Pérez stated, "We can see, for example, that they have a common 'language,' something like a dialect or dialectic."
- ...

© 2017-2018, Surromind Robotics, Inc.



제12회 관악블록세미나

문학적 상상력과 인공지능

제6강

## 인공지능의 텍스트 이해 능력

홍문표

---

성균관대학교 독어독문학과 & 컬처애편테크놀로지 융합전공 교수  
성균관대학교 출판부 부장

서울대학교 독어독문학과 졸업  
독일 자를란트 대학 박사(전공: 전산언어학)  
한국전자통신연구원(ETRI) 음성언어정보연구센터 선임연구원  
프로젝트: “지식학습기반의 다국어확장이 용이한 관광/국제행사  
통역률 90%급 자동통번역 소프트웨어 원천기술개발” 등  
논문: 「한-독 기계번역을 위한 한국어 영형목적어 처리 연구」 등



# 인공지능의 언어의미 이해 능력

## 1. 들어가는 말

알파고의 충격적인 등장 이후 우리 사회의 가장 큰 화두 중 하나는 인공지능과 4차산업혁명이라 할 수 있다. 특히 인공지능 기술의 눈부신 발전은 이 기술의 한계가 어디인지 우리를 기대하게 하고 또 한편으로는 우려를 불러일으키기도 한다.

이 발표에서는 인공지능의 핵심 능력 중 하나인 언어처리 능력, 그 중에서도 의미이해 능력에 대해 살펴보고자 한다. 컴퓨터가 인간 언어의 의미를 이해한다는 것은 무엇인가? 컴퓨터는 어떤 원리에 의해 인간언어의 의미를 이해하는가? 컴퓨터는 인간처럼 언어의 의미를 이해하는 것인가? 이 발표를 통해 이러한 질문에 대한 해답을 찾을 수는 없겠지만 그 해답에 대한 실마리를 찾고자 한다.

컴퓨터로 하여금 자연언어의 의미를 이해하여 이를 인간생활에 유용하게 사용하고자 하는 노력은 컴퓨터의 상용화와 거의 동시에 시작되었다. 그럼에도 불구하고 최근 약 3~4년 이내에 부쩍 컴퓨터, 좀 더 구체적으로 말하자면, 인공지능의 언어이해능력에 대한 관심이 급증하고 있는 이유는 그 성능이 과거와는 비교할 수 없을 정도로 향상되었기 때문일 것이다.

그 이유에 대한 설명은 대략 컴퓨팅 파워의 발전, 대용량 디지털 언어데이터의 존재, 인공지능 알고리즘의 발전 등에서 찾을 수 있을 것이다. 최근 언어처리분야에서 가장 각광을 받고 있는 뉴럴 네트워크(neural network) 기반의 딥러닝 알고리즘도 사실 최근에 개발된 것은 아니다. 이 기술은 이미 1980년대에도 존재하였으나 컴퓨팅파워의 부족으로 현실화되지 못했을 뿐이다.

이 발표에서는 전산언어학(computational linguistics)과 자연언어처리(natural language processing) 분야에서 자연언어의 의미처리를 위해 시도해 온 방법론을 개괄적으로 살펴보고 그 장단점에 대해 소개할 것이다. 자연언어의 의미처리에는 크게 세 가지의 패러다임이 존재한다. 언어학자들에게 익숙한 논리학 기반의 기호적 접근법(symbolic approach, logical approach)가 하나이며, 데이터에 기반한 통계적 접근법(statistical approach)가 다른 하나이다. 그리고 데이터에 기반한다는 점에서 기존의 통계적 접근법과 유사하지만 뉴럴네트워크(neural network) 등을 활용하며 비지도학습(unsupervised learning)이 가능한 딥러닝기반의 접근법(deep learning approach)가 있다.

2장에서는 논리학 기반의 기호적 접근법에 대하여 간단히 소개하며, 이 방법론의 장단점에 대해 논한다.

3장에서는 통계적 접근법에 대해 소개하며, 이 방법론을 활용한 화행(speech act) 자동분석 방법을 소개한다.

4장에서는 최근 가장 유행하고 있는 딥러닝방식의 언어의미이해에 대하여 간단히 소개한다. 특히 word2vec프로그램을 활용하여 독일어 어휘의 의미관계를 데이터에서 자동으로 추출하는 방법과 실험결과에 대해 소개할 것이다.

마지막으로 5장에서는 인공지능의 언어의미 이해능력에 대해 종합하여 정리한 후, 이 기술의 미래 발전 방향에 대해 예측하는 것으로 본 강연을 끝맺고자 한다.

## 2. 언어이해를 위한 기호적 접근법

이 장에서는 프레게(Frege)식의 합성의미론(compositional semantics)에 기반한 의미이해에 대해 소개한다. 이 접근법의 기본 아이디어는 다음과 같다. 어떤 의사소통의 단위 (예를 들어, 문장)의 의미는 그 단위를 이루고 있는 구성요소 (예를 들어, 구 혹은 단어)의 의미의 합으로 이루어져있으며, 그 구성요소들의 의미는 통사관계에 따라 결합한다. 다시 말해, 모든 단어는 자신만의 의미 또는 기능을 가지고 있으며 통사적으로 관계를 맺는 인접단어와 술어-논항의 관계로 결합하게 되고 이 과정을 반복적으로 되풀이하게 되면 최종적으로 가장 상위 노드인 문장의 의미를 얻을 수 있다는 것이다.

AMR(Abstract Meaning Representation)<sup>1</sup>은 이러한 의미합성 철학을 반영한 가장 대표적인 의미표상 메타언어 중의 하나이다. 의미합성방법을 활용하면 문장의 명제적인 의미(propositional meaning) 뿐만 아니라 문장의 논조(sentiment)와 같은 주관적인 의미(subjective meaning)도 구할 수 있다.

언어학의 형식의미론 등에 기반한 기호적 접근법은 최근 딥러닝 기반의 방법론이 등장한 이후 그 중요성이 상대적으로 떨어진 것으로 보이지만 실상은 그렇지만은 않다. 4장에서도 언급하겠지만 딥러닝 기반의 방법론은 블랙박스(black box) 식의 방법론이므로 오류 발생시 그 수정이 어렵다는 단점이 있는 반면, 의미론에 기반한 기호적 접근법은 의미이해의 오류발생시 그 수정이 상대적으로 용이하다는 장점이 있다. 따라서 국내외의 상용화된 대화시스템 중 상당수는 여전히 기호적 접근법을 활용하고 있다. 본 발표에서는 기호적 접근법 패러다임에 속하는 방식 중 AIML(artificial intelligence markup language)에 대해 간략히 소개한다.

## 3. 통계적 접근법을 활용한 화행 자동분석

통계적 방법을 활용한 언어의미의 이해방식 중 가장 대표적인 것은 기계학습(machine learning)을 활용한 방법이다. 기계학습은 주어진 학습 데이터에서 유용한 패턴을 학습하여 새로운 데이터가 주어졌을 때 학습한 패턴을 적용하여 새로운 데이터의 패턴을 파악해내는 것을 목표로 한다. 이를 언어이해에 대입한 예를 들어보자면, 학습데이터에서 기계학습 방법론을 적용하여 화행(speech act)의 패턴을 알아낸 후, 새로운 데이터(예를 들어 발화)의 화행을 자동으로 알아내는 과제를 들 수 있다.

아래에서는 구영은(2018)의 연구를 기반으로 이에 대한 소개를 하고자 한다. 화행이란 “의사소통의 실질적 단위(functional unit in communication)” 이자 “의사소통을 함으로써 화자가 드러내고자 하는

---

<sup>1</sup> cf. Banarescu et al, (2013)

태도(the type of attitude being expressed)”이다. 성공적인 의사소통을 이루어지기 위해서는 화자의 발화 의도인 화행을 정확하게 파악하는 것이 매우 중요하다. 그러나 다음의 예문과 같이 발화의 표층 형태(surface form)만으로는 화행을 제대로 분석하기 어려운 경우가 있다.

1) A: Kannst du mir das Salz reichen?

(Can you pass me the salt?)

B: Ja, hier bitte.

(Yes, here.)

발화자 A의 발화는 개인의 능력을 나타내는 조동사가 포함된 의문문이다. 문장의 표층 형태로만 볼 때 발화자 A는 상대발화자 B가 특정 행위를 할 수 있는 능력을 가지고 있는지 물어보는 yes-no 질문을 발화하는 것으로 보인다. 그러나 실제로 발화자 A는 소금을 넘겨달라는 요청의 발화를 하고 있다.

이처럼 발화의 표층 형태만으로는 발화 의도를 온전히 이해하기 어렵다. 따라서 발화의 화행을 결정짓는 다양한 특징들을 분석할 필요성이 존재한다. 구영은(2018)의 연구에서는 발화의 화행을 결정짓는 자질을 살펴보기 위해 먼저 독일어 대화문에 나타나는 화행의 종류를 찾아보고, 다음으로 발화와 발화의 화행을 분석하여 화행별 자질을 언어학적으로 분석한다. 발화와 발화가 특정 화행을 갖게 하는 자질을 분석하고 체계화할 수 있다면 인공지능 스피커 등 다양한 분야에 적용할 수 있을 것이다.

구영은(2018)의 연구에서는 화행 분류 체계를 정립하기 위해 기존연구에서 널리 사용되고 있는 이재원(1999), Rasor(2011)의 화행 분류 체계와 DAMSL tagset을 분석하였다. 두 분류 체계에서 중복 혹은 누락된 화행을 보완하고자 독일어 대화 코퍼스를 분석하고 해당 코퍼스에서 발견되는 화행을 이용하여 화행 분류 체계를 자체 정립하였다. 이 연구에서 제안한 화행 분류 체계는 총 23개의 화행으로 이루어져 있다.

| 번호 | 화행          |        | 번호 | 화행          |      | 번호 | 화행       |    |
|----|-------------|--------|----|-------------|------|----|----------|----|
| 1  | Accept      | 수락     | 9  | Avoid       | 대답회피 | 17 | Maybe    | 추측 |
| 2  | Acknowledge | 호응     | 10 | Call        | 부름   | 18 | Promise  | 약속 |
| 3  | Answer      | 답변     | 11 | Correct     | 수정   | 19 | Reject   | 거절 |
| 4  | Apologize   | 사과     | 12 | Exclamation | 감탄   | 20 | Request  | 요청 |
| 5  | Ask-confirm | 확인요구   | 13 | Express     | 의지표현 | 21 | Response | 반응 |
| 6  | Ask-if      | Y/N 질문 | 14 | Greeting    | 인사   | 22 | Suggest  | 제안 |
| 7  | Ask-Ref     | WH 질문  | 15 | Induce      | 유도   | 23 | Thanking | 감사 |
| 8  | Assert      | 주장     | 16 | Inform      | 정보제공 |    |          |    |

표 1. 구영은(2018)의 화행 분류 체계

본 연구에서는 상응하는 화행을 결정짓는 발화문의 특징을 언어학적으로 분석하였다. 화행을 결정짓는 자질은 문장 자질과 문맥 자질로 나뉘며, 총 9개의 화행 분류 자질을 아래와 같이 제안한다. 문장 자질은 문장 구조 또는 문장을 구성하는 성분의 의미와 같이 문장에서 발견되는 자질이며, 문맥 자질은 화행을 분석하고자 하는 해당 발화가 아닌 주변 발화와 관련된 자질이다.

| 번호 | 자질 이름 | 자질 값             |                                             |
|----|-------|------------------|---------------------------------------------|
| 1  | 문장 자질 | 문장 유형            | decl, impe, excl,<br>wh_ques, yn_ques, ques |
| 2  |       | 주어의 인칭           | 1, 2, 3                                     |
| 3  |       | 부정형 포함 여부        | 0, 1                                        |
| 4  |       | 수행동사             | 0, 1                                        |
| 5  |       | 본동사의 개수          | 0, 1, 2, 3                                  |
| 6  |       | 문장의 길이           | 1, 2, 3                                     |
| 7  | 문맥 자질 | 바로 이전 화행         | 1 ~ 23                                      |
| 8  |       | 상대방의 바로 이전<br>화행 | 1 ~ 23                                      |
| 9  |       | 발화 차례의 변화 여부     | 1 ~ 23                                      |

표 2. 구영은(2018)의 화행 분류 자질

본 연구에서 제안하는 화행 분류 자질을 검증하기 위해 독일어 대화 코퍼스를 이용하여 정확률 실험을 진행하였다. 실험코퍼스는 DGD(Datenbank für Gesprochenes Deutsch)에서 제공하는 교수와 학생들 간의 대화로 총 409개의 발화로 구성된다. 실험은 자바(Java)기반의 기계학습 알고리즘 툴인 웨카(Weka) 3.8.1 버전을 이용하여 진행하였다. 기계학습 알고리즘으로는 ‘지지 벡터 머신(Support Vector Machine)’을 사용하였고, ‘10-fold cross validation’ 방식으로 성능을 평가하였다. 실험 결과는 표 3과 같다.

| 사용된 자질 | 정확률(%)        |       |
|--------|---------------|-------|
| 베이스라인  | 유니그램(Unigram) | 43,91 |
| 실험1    | 문장자질          | 64,21 |
| 실험2    | 문맥자질          | 55,84 |
| 실험3    | 문장자질+문맥자질     | 72,34 |

표 3. 자질에 따른 정확률 비교

본 연구에서 제안한 자질들을 모두 사용할 때 72.34%로 정확률이 가장 높았다. 이는 베이스라인 대비 28%의 성능 향상이다. 실험 1과 실험 2는 문장자질과 문맥자질 중 한 가지의 상위 자질 유형만을 사용한 경우이다. 성능 향상의 폭을 볼 때는 문맥자질보다는 문장자질이 화행 자동분류에 더 큰 영향을 준 것으로 보인다. 그러나 각 9개의 하위 자질들이 화행 자동분류에 미치는 영향력을 볼 때 ‘바로 이전 화행’, ‘상대방의 바로 이전 화행’과 같은 문맥자질이 더 영향력이 큰 것을 확인할 수 있었다. 각 하위 자질들이 상위 자질 유형의 성능에 미치는 영향력을 통계적으로 분석해볼 필요가 있다.

본 연구에서는 독일어 대화에서 나타나는 화행을 분석하고 화행 분류를 위한 자질을 언어학적으로 분석하고자 하였다. 본 연구에서 제안하는 자질들을 사용한 결과, 화행 자동분류 성능이 큰 폭으로 향상되었다. 하지만 여전히 대부분의 자질이 문장의 문법적인 정보에 국한된다는 한계를 갖는다. 의미상 직접적으로 화행과 연관되는 단서 표현(Cue phrase)이나 독일어 대화문에서만 특수적 발견되는 독일어 수행부사(Performative adverb) 등의 자질을 향후에 연구할 필요성이 있다.

#### 4. Word2Vec을 활용한 독일어 어휘의미 관계 분석

이번 장에서는 딥러닝기반의 언어의미해 방식에 대해 소개한다. 특히 word2vec<sup>2</sup> 방법론을 활용하여 인공지능이 어휘의미를 어떻게 파악할 수 있는지 다룬다.

##### 4.1 word2vec

컴퓨터가 인간이 사용하는 자연어를 이해하기 위해서는 어떠한 방식으로 데이터가 입력이 되어야 할까? 텍스트가 입력값(input value)이 될 수 있을까? 불가능하다. 왜냐하면 컴퓨터에게 하나의 문자(character) 또는 문자로 이루어진 문자열(string)은 모두 숫자의 조합에 지나지 않기 때문이다. 따라서 텍스트를 컴퓨터가 이해할 수 있는 숫자의 값으로 변환하는 과정이 필요하다.

‘Love you’ 와 ‘Hate you’ 라는 문장을 컴퓨터가 이해할 수 있는 방식으로 전환해보고자 한다. 단어를 숫자로 바꾸는 방법 중 가장 간단하게 생각할 수 있는 것은 세상에 존재하는 모든 단어에 인덱스(Index)를 붙이는 방식일 것이다. 두 문장에서 추출할 수 있는 단어는 ‘Love’, ‘Hate’, ‘You’ 세 개의 단어이다. 따라서 각각의 단어를 (그림 1)과 같은 방식으로 인덱싱(Indexing)할 수 있다. 이렇게 각각의 단어에 인덱스를 붙이는 방식을 ‘인코딩(Encoding)’ 이라 한다.

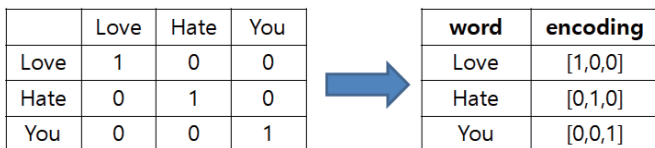
| word | index |
|------|-------|
| Love | 0     |
| Hate | 1     |
| You  | 2     |

(그림 1) 인코딩

<sup>2</sup> cf. Mikolov et al.(2013)

### 원-핫 인코딩 (One Hot Encoding)

인코딩 방식 중 하나인 원-핫 인코딩(One-hot encoding)은 코퍼스에 존재하는 단어의 개수를 벡터의 크기로 하고, 표현하고 싶은 단어의 인덱스에 1의 값을 부여하고, 다른 인덱스에는 0을 부여하는 단어의 벡터 표현 방식이다. 즉, 원-핫 벡터(One-hot Vector)는 (그림 2)와 같은 과정을 통해 만들어진다.



(그림 2) 원-핫 벡터 변환과정

그러나 이러한 원-핫 인코딩 방식은 몇 가지 한계가 있다. 첫 번째는 단어의 개수가 늘어날수록 벡터의 차원이 늘어난다는 것이다. 단어의 개수가 100개가 된다면, 하나의 단어는 하나의 값만이 1이고, 99개의 값이 0인 100차원 크기의 벡터의 형태를 가지게 될 것이다. 두 번째 한계는 원-핫 인코딩의 방식으로는 단어들 사이의 유사도를 표현하지 못한다는데 그 한계가 있다.

### 임베딩 (Embedding)

원-핫 인코딩 방식의 한계를 극복하기 위하여 등장한 방식이 임베딩(Embedding)이다. 임베딩 방식은 하나의 단어에 정수 인덱스를 설정하는 대신에 단어를 수치화하여 나타내는 방식이다. 그리고 이 임베딩 방식 중 하나가 Word2Vec 방식이다. 임베딩 방식은 인코딩 방식보다 저차원의 벡터를 통해 단어를 나타낼뿐만 아니라 단어들 사이의 유사도를 나타낼 수 있다는 점에서 장점을 가진다.

### 분포 의미론 (Distributional Semantics)

Word2Vec의 기본적인 이론적 배경은 분포 가설(distributional hypothesis)이다. 분포가설이란 Firth(1957)<sup>3</sup>가 "You shall know a word by the company it keeps"라고 말했듯이, 유사한 의미의 단어들은 유사한 문맥에서 사용된다는 것이다. 예를 들어 "Hund"와 "Katze"는 "Hund"와 "Brot"보다 훨씬 더 자주 유사한 문맥에서 사용될 것이다. 이러한 특성을 수학적으로 형식화하여 벡터상의 공간에 표현할 수 있게 해 주는 프로그램이 Word2Vec이다. Word2Vec에서 단어의 벡터를 만드는 방식은 크게 CBOW와 skip-gram 두 가지 방식이 있다.

---

<sup>3</sup> Church et al.(1990)에서 재인용.



### CBOW(Continuous Bag of Words)

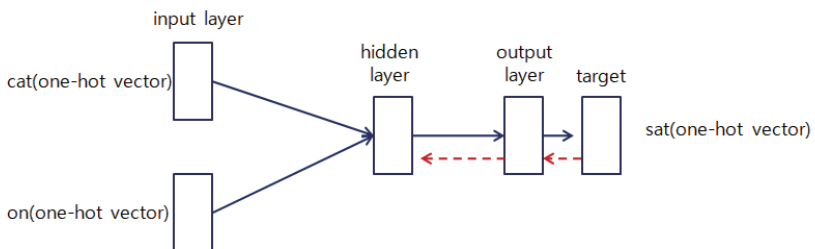
CBOW 방식은 벡터값을 예측해야 하는 단어를 ‘중심단어(center word)’, 예측에 사용되는 단어를 ‘주변단어(context word)’라 할 때, 주변단어들을 통해 중심단어의 벡터값을 예측하는 방식이다. 예를 들어, “The fat cat sat on the mat” 라는 문장이 있을 때, 이를 중심단어와 주변단어의 쌍으로 나타내면 (그림 3)과 같이 나타낼 수 있다.

#### **The fat cat sat on the mat**

| center | context  |   | center        | context                      |
|--------|----------|---|---------------|------------------------------|
| The    | fat      | ➡ | [1,0,0,0,0,0] | [0,1,0,0,0,0]                |
| fat    | The, cat |   | [0,1,0,0,0,0] | [1,0,0,0,0,0], [0,0,1,0,0,0] |
| cat    | fat, sat |   | [0,0,1,0,0,0] | [0,1,0,0,0,0], [0,0,0,1,0,0] |
| sat    | cat, on  |   | [0,0,0,1,0,0] | [0,0,1,0,0,0], [0,0,0,0,1,0] |
| on     | sat, the |   | [0,0,0,0,1,0] | [0,0,0,1,0,0], [0,0,0,0,0,1] |
| the    | on, mat  |   | [0,0,0,0,0,1] | [0,0,0,0,1,0], [0,0,0,0,0,0] |
| mat    | the      |   | [0,0,0,0,0,0] | [0,0,0,0,0,1]                |

(그림 3) “The fat cat sat on the mat” 의 중심-주변 단어쌍

이 단어쌍을 토대로 CBOW 방식은 주변 단어를 이용하여 중심단어의 벡터값을 예측한다. CBOW 방식을 구조화 하여 나타내면 (그림 4)와 같다.



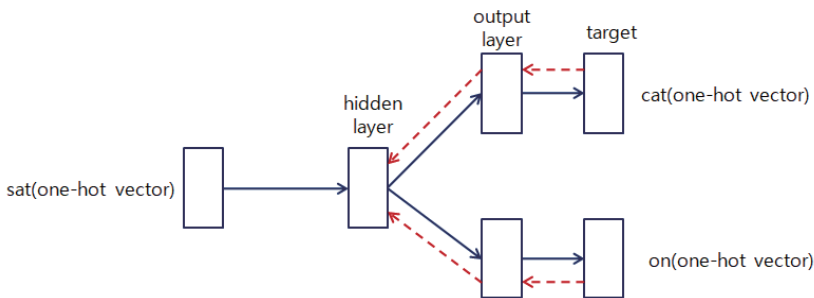
(그림 4) CBOW의 구조

(그림 4)는 sat의 벡터값을 예측하기 위해 주변단어인 cat과 on이 입력층(input layer)을 구성하는 경우이다. 입력층과 은닉층(hidden layer) 사이에 가중치  $W$ , 은닉층과 출력층(output layer) 사이의 가중치  $W'$ 를 이용하여 은닉층의 벡터를 완성 시켜가는 과정을 학습과정이라 할 수 있다. 처음에는  $W$ 와

$W'$ 의 행렬이 임의의 수로 설정이 되어있다.  $W$ 와  $W'$ 의 값을 이용하여 softmax 함수를 통해 출력층이 결정된다. 처음에는  $W$ 와  $W'$ 의 값이 임의로 결정되어 있기 때문에 결정된 출력층의 값은 정답이 되는 단어, 즉 이 예시에서는 “sat”의 벡터와 그 값이 차이가 난다. 이 값의 차이를 cross entropy를 이용하여 오차를 줄이기 위해 역방향( $\leftarrow$  -  $\rightarrow$ )으로  $W$ 와  $W'$ 의 값을 조정하는 과정을 거치게 된다. 이러한 전체 과정이 바로 학습과정이며, 학습 과정이 끝나면 입력된 데이터를 기반으로 각 단어들 사이에 유사도를 알 수 있는 벡터값을 도출해낼 수 있게 된다.

### skip-gram

skip-gram 방식은 CBOW와는 반대로 중심단어를 이용하여 주변단어의 벡터값을 예측하는 알고리즘이다. 그 구조는 (그림 5)와 같다.



(그림 5) skip-gram의 구조

일반적으로 CBOW 알고리즘은 크기가 작은 데이터 세트(set)에 skip-gram은 크기가 큰 데이터 세트에 적합하다. 또한 CBOW보다는 skip-gram이 성능이 높은 것으로 알려져 있다.

본 발표를 위해 Word2Vec을 활용하여 독일어 코퍼스로부터 어휘들간의 의미관계를 추출하는 실험을 진행하였다.<sup>4</sup>

먼저 실험을 위한 코퍼스는 2014년 1월 1일부터 2018년 6월 30일까지 [www.dw.de](http://www.dw.de)에 “berlin”이라는 단어가 들어가 있는 기사의 리드(lead)부분으로 하였다. 여기서 리드는 독자의 흥미를 유발하는 역할을 하는 기사의 첫 부분을 말한다. 코퍼스의 전체 문장수는 13,421문장이었다. 웹데이터 크롤링을 위해서는 BeautifulSoup 라이브러리를 사용하였고, 딥러닝을 위해서는 gensim의 word2vec 라이브러리를 사용하였다.

<sup>4</sup>실험을 위한 코딩과 딥러닝 학습은 성균관대학교 독어독문학과 석사과정 김지연 학생이 수고하였다.

우리는 분포의미론에 기반한 word2vec 프로그램이 우리가 원하는 어휘에 대해 어느 정도의 정확도로 관련 어휘를 찾아내는지 알아 보았다. 실험에 사용한 어휘는 'Merkel', 'Angela', 'Deutschland', 'Berlin' 총 4개였다. 프로그램 수행 결과 다음과 같은 어휘관계를 얻어내었다.

— 결과

1) merkel과 유사도가 높은 단어

```
model.wv.most_similar('merkel')

C:\Users\User\Anaconda3\lib\site-packages\ger
to `np.signedinteger` is deprecated. In futur
if np.issubdtype(vec.dtype, np.int):

[('bundeskanzlerin', 0.9995476007461548),
 ('angela', 0.9995169639587402),
 ('präsident', 0.9992945194244385),
 ('meint', 0.9992083311080933),
 ('außenminister', 0.9991261959075928),
 ('russland', 0.9990661144256592),
 ('frankreichs', 0.999061644077301),
 ('ukraine', 0.9990286231040955),
 ('worden', 0.9990043640136719),
 ('kanzlerin', 0.9989970326423645)]
```

2) angela와 유사도가 높은 단어

```
model.wv.most_similar('angela')

C:\Users\User\Anaconda3\lib\site-packages\ger
to `np.signedinteger` is deprecated. In fut
if np.issubdtype(vec.dtype, np.int):

[('bundeskanzlerin', 0.9997750520706177),
 ('merkel', 0.9995168447494507),
 ('präsident', 0.9994031190872192),
 ('meint', 0.999315619468689),
 ('kanzlerin', 0.9993054270744324),
 ('russland', 0.9991922378540039),
 ('frankreichs', 0.999180018901825),
 ('außenminister', 0.999173104763031),
 ('ukraine', 0.9991565942764282),
 ('worden', 0.9991207122802734)]
```

3) deutschland 와 유사도가 높은 단어

```
model.wv.most_similar('deutschland')

C:\Users\User\Anaconda3\lib\site-packag
to `np.signedinteger` is deprecated. In
if np.issubdtype(vec.dtype, np.int):

[('gibt', 0.999925434589386),
 ('große', 0.999902606010437),
 ('russland', 0.9999011158943176),
 ('ukraine', 0.9999009966850281),
 ('tage', 0.9999008178710938),
 ('mehr', 0.9998952746391296),
 ('geht', 0.999891996383667),
 ('usa', 0.9998905658721924),
 ('union', 0.999889075756073),
 ('lang', 0.9998827576637268)]
```

4) berlin 과 유사도가 높은 단어

```
model.wv.most_similar('berlin')

C:\Users\User\Anaconda3\lib\site-packag
to `np.signedinteger` is deprecated. In
if np.issubdtype(vec.dtype, np.int):

[('juni', 0.9999019503593445),
 ('2016', 0.9999016523361206),
 ('monaten', 0.9998964667320251),
 ('experten', 0.9998851418495178),
 ('arbeit', 0.9998739361763),
 ('bringen', 0.9998727440834045),
 ('frankfurt', 0.999871015548706),
 ('mehrere', 0.9998655915260315),
 ('berichtet', 0.9998591542243958),
 ('lässt', 0.9998487234115601)]
```

위 그림 1에서 보는 것과 같이 'Merkel'이라는 어휘와 가장 관련성이 높은 것으로 분석된 어휘는 'Bundeskanzlerin'이었다. 괄호 안의 수치가 높으면 높을수록 해당 어휘와 유사도가 높은 것으로 볼 수

있다. 두 번째로 높은 어휘는 'Angela'였는데, 이는 'Angela Merkel'이라는 고유명사를 전처리 단계에서 따로 인식하지 못하고 서로 다른 두 개의 어휘로 인식하였기 때문이다. 어쨌든 이 결과도 'Merkel'이라는 성(family name)이 'Angela'라고 하는 이름(first name)과 관련성이 높다는 것을 보여주므로 의미가 있다고 보인다. 그 외에도 'Merkel'은 'Präsident', 'Russland', 'Frankreich' 등의 어휘와 관련성이 높은 것으로 분석되었다.

이와 같이 word2vec을 활용하여 분석한 결과는 어떠한 Germanet이나 Wordnet 등과 같은 방대한 규모의 어휘망도 포착하지 못하는 어휘의미관계를 찾아낼 수 있다는 장점을 보인다.

## 5. 맺는말

이 강연에서는 인공지능의 의미이해 방식을 알아보기 위해 세 가지의 방법론을 소개하였다. 관련 연구분야의 추세는 점차 기호론적 방법론과 기계학습 기반의 방법론에서부터 딥러닝 방식의 방법론으로 패러다임이 바뀌어 가고 있는 상황이다. 딥러닝 방식은 언어처리분야에 본격적으로 적용된 역사가 그리 깊지 않다. 그럼에도 불구하고 기계번역 등의 분야에서 놀랄만한 성과를 보이고 있고, 그 한계가 어디까지인지 추정하기가 쉽지도 않다.

그러나 딥러닝 방식은 그 구조가 블랙박스라는 점에서 전적으로 딥러닝 방식에만 의존하여 인공지능의 언어이해기술을 개발하는 것은 어려움이 있을 것이다. 결국 언어학적으로 설명하기 어려운 부분에서는 딥러닝 방식도 한계를 보일 수도 있다. 물론 알파고의 경우에서처럼 우리의 상상을 뛰어넘는 일이 일어날 수도 있겠지만, 적어도 현재 시점에서는 딥러닝 방식과 함께 기존의 언어학 기반의 방법론도 같이 활발히 연구되어야 할 것으로 보인다.

## 참고문헌

- Austin(1962), How to Do Things with Words, Oxford: Oxford University Press.
- Bach(1994), Conversational implicature, Mind&Language, 9(2), 124-162.
- Banarescu et al.(2013), Abstract meaning representation for Sembanking, In Proceedings of Linguistic Annotation Workshop.
- Church, Kenneth Ward, and Patrick Hanks, "Word association norms, mutual information, and lexicography," Computational linguistics 16,1 (1990): 22-29.
- Core et al.(1997), Coding dialogs with the DAMSL annotation scheme, In AAAI fall symposium on communicative action in humans and machines (Vol. 56).
- Hong, Munpyo, and Oliver Streiter(1998) "Overcoming the language barriers in the Web: The UNL-Approach," Tagungsband der 11 KONVENS.
- Mikolov, Tomas, et al.(2013), "Distributed representations of words and phrases and their compositionality," Advances in neural information processing systems.
- Rasor et al.(2011), Student Speech Act Classification Using Machine Learins, In FLAIRS Conference.
- 구영은(2018), 독일어 대화체 문장의 화행 분석과 화행 자동분류에 대한 연구: 독일어 수업대화를 중심으로, 성균관대학교 석사학위논문.
- 이현정 외(1997), 한국어 대화체 문장의 화행 분석, 한국정보과학회 학술발표논문집, 24(2Ⅱ), 259-262.
- 이재원(1999), 통계적 화행처리를 이용한 대화체 기계번역에서의 효율적인 대화분석, 박사학위논문, 한국과학기술원, 1999.