

ICAIH 2021

THE 4TH INTERNATIONAL
CONFERENCE ON
ARTIFICIAL
INTELLIGENCE
HUMANITIES

October 2, 2021 10:00~18:30 (KST)

Participation: Online (Synchronous Video-conferencing via ZOOM Webinar)



SCHEDULE

Date

October 2 , 2021 (Sat.) 10:00 – 18:30 KST

Participation

YouTube Live

▶ <https://linktr.ee/ICAIH2021>

-Korean-English/English-Korean simultaneous interpretation provided

Announcement

-ICAIH 2021 webpage(<http://aihumanities.org/aih2021>)

-Humanities Research Institute Webpage(<http://aihumanities.org>)

Inquiries

aihumanities2018@gmail.com

PROGRAMS

| | | |
|-------------------|--|--|
| 10:00 – 10:10 KST | Opening Ceremony | 👤 Hyeongjoo Kim Humanities Research Institute(Korea) |
| 10:10 – 10:30 KST | Opening Speech | 👤 Chankyu Lee Humanities Research Institute(Korea) |
| 10:30 – 11:30 KST | Keynote Speech 1 (Moderator: Heeryon Cho) | 👤 Toru Ishida Waseda University(Japan) |
| 10:30 – 11:30 KST | Keynote Speech 2 (Moderator: Cheongho Lee) | 👤 Toby Walsh UNSW(Australia) |
| 12:30 – 14:00 KST | Lunch Time | |
| 14:00 – 16:00 KST | Session 1 ~ 3 | |
| 16:15 - 18:15 KST | Session 4 ~ 6 | |
| 18:15 – 18:30 KST | Closing Speech | |

KEYNOTE SPEECH

Multiagent Systems: Another AI for Humanity Research

 **Toru Ishida** Waseda University, Japan





AI and Ethics: Why all the fuss?

 **Toby Walsh** UNSW, Australia

SESSION 1





Moderator  **Jinbeom Synn** Seowon University

AI and Literature

| | |
|-------------------|---|
| 14:00 ~ 14:25 KST | AI and Posthuman Romance in Literature and Film  Seiwoong Oh Rider University(USA) |
| 14:25 ~ 14:50 KST | Paradigms of Creativity in the Age of Posthumanism: An Unbridgeable Gap between AI and Humanity?  Tamar Cheishvili Independent Scholar(USA) |
| 14:50 ~ 15:15 KST | Deep Medicine: Deep Empathy and Potential Healing in <i>Klara and the Sun</i>  Pin-chia Feng National Yang Ming Chiao Tung University(Taiwan) |
| 15:15 ~ 15:40 KST | Envisioning Artificial Intelligence and Simulated Affections in Ian McEwan's <i>Machines Like Me</i>  Chung-jen Chen National Taiwan University(Taiwan) |
| 15:40 ~ 16:00 KST | Discussion |




SESSION 2

Moderator  Hongkyu Ha Sookmyung Women's University

| AI and Culture | |
|-------------------|--|
| 14:00 ~ 14:25 KST | <p>A Narratological Inquiry of AI and Human Identity in the Posthuman era: Focusing on Netflix short animation <i>Zima Blue</i></p> <p> Yunkyong Kim Word Work Center(Korea)</p> |
| 14:25 ~ 14:50 KST | <p>Overcoming Animal Digitalis</p> <p> Alfonso Ballesteros Universidad Miguel Hernández(Spain)</p> |
| 14:50 ~ 15:15 KST | <p>Epistemological Nature of Mind Perception in Artificial Intelligence</p> <p> Hyungrae Noh Suncheon National University(Korea)</p> |
| 15:15 ~ 15:40 KST | <p>In the Age of AI, Do we need to Change the Definition of Art?</p> <p> Jaepark Rhee Chugye University for the Arts(Korea)</p> |
| 15:40 ~ 16:00 KST | Discussion |




SESSION 3

Moderator  Iksoon Jeong Chung-Ang University

| AI and Future & Past | |
|----------------------|---|
| 14:00 ~ 14:30 KST | <p>Possibility of Agent Model-based Historical Descriptions</p> <p> Kiwoo Hwang Sungkyunkwan University(Korea)</p> |
| 14:30 ~ 15:00 KST | <p>A Study on the English Old Poor Law System with Agent-Based Modeling Simulation</p> <p> Seunghwan Lee Sungkyunkwan University(Korea)</p> |
| 15:00 ~ 15:30 KST | <p>Charles Darwin's Legacy in Artificial Intelligence</p> <p> Jiyeon Kim & Youngsue Han Hankuk University of Foreign Studies(Korea)</p> |
| 15:30 ~ 16:00 KST | Discussion |




SESSION 4

Moderator  Kyeongmin Kim Korea University

| AI and Language | |
|-------------------|---|
| 16:15 ~ 16:45 KST | <p>Language Sphere: A Socio-Technical Approach to Bilingual Dictionary Creation for Low-Resource Languages</p> <p> Yohei Murakami Ritsumeikan University (Japan)</p> |
| 16:45 ~ 17:15 KST | <p>A Discussion Support System Using Neural Language Model: A Preliminary Report</p> <p> Mamoru Yoshizoe Ritsumeikan University (Japan)</p> |
| 17:15 ~ 17:45 KST | <p>Essential Attribute and Core Tasks of Computational Linguistics from the Perspective of Artificial Intelligence</p> <p> Yude Bi & Danhui Yan Fudan University(China)</p> |
| 17:45 ~ 18:15 KST | Discussion |

SESSION 5




Moderator  Taehee Kim Konkuk University

| AI and Philosophy | |
|-------------------|--|
| 16:15 ~ 16:45 KST | <p>How to Make Ideas Clear: Artificial Intelligence and Human Instinct from a Peircean Perspective</p> <p> Cheongho Lee Sangmyung University(Korea)</p> |
| 16:45 ~ 17:15 KST | <p>The Meaning of Labor in the Age of Artificial Intelligence</p> <p> Sunjin Yang Korean National Police University(Korea)</p> |
| 17:15 ~ 17:45 KST | <p>Zhouyi's Biantong World and the Algorithm of Yin-yang from the Dimension of AI Connectionism</p> <p> Yonjae Kim National Gongju University(Korea)</p> |
| 17:45 ~ 18:15 KST | Discussion |

SESSION 6

Moderator  Travis Shaw Chung-Ang University

AI and Society (with Algorithm)

| | |
|-------------------|--|
| 16:15 ~ 16:45 KST | Love Story in the Algorithmic Age  Aditi Vashistha University of Delhi(India) |
| 16:45 ~ 17:15 KST | Application of Artificial Intelligence through the K nearest Neighbor Algorithm to Classify Healthy Indonesian Card Recipients in an Appropriate Target  Rifaldy Fajar Yogyakarta State University(Indonesia) |
| 17:15 ~ 17:45 KST | Boredom, Time-Perception and Algorithmic Governmentality  Aditya Nayak Jawaharlal Nehru University(India) |
| 17:45 ~ 18:15 KST | Discussion |

KEYNOTE SPEECH

Multiagent Systems: Another AI for Humanity Research

 **Toru Ishida** Waseda University(Japan)

Abstract

This talk introduces multi-agent systems, a major subarea of Artificial Intelligence. Unlike most of other AI subareas, the goal of multi-agent systems is to model a human society but not a single brain. The area was born in 1980's with a small number of researchers. Until 2000's, the most of work is theoretical including both mathematical and conceptual modeling of human societies. In 2010's, however, thanks to big data and high-performance computing (HPC), multiagent systems became popular in various academic areas such as economics, management, agriculture, disaster management and so on. Since a multiagent system can naturally model a group of humans, multiagent simulation can predict potential problems in our society, and create possible solutions. This talk provides a research history and recent successful applications of multiagent systems.

KEYNOTE SPEECH

Multiagent Systems: Another AI for Humanity Research

 **Toru Ishida** Waseda University(Japan)

Abstract:

This talk introduces multi-agent systems, a major subarea of Artificial Intelligence. Unlike most of other AI subareas, the goal of multi-agent systems is to model a human society but not a single brain. The area was born in 1980's with a small number of researchers. Until 2000's, most of the work was theoretical including both mathematical and conceptual modeling of human societies.

I started working on multi-agent systems in the late 1980's. I led a research group in NTT labs, and created a series of search algorithms for agents. At that time, the main focus of multi-agent system researchers was the conceptual model of coordination. My group took a different approach to exploring search algorithms for agents. I thought that the first chapter of the textbook for multi-agent systems (called Distributed AI at that time) textbook would be "search algorithms for agents," because the first chapter of AI textbooks was often "search algorithms." We have created algorithms that allow agents to utilize and improve on previous experiments, adapt to dynamically changing goals, and cooperatively solve problems with other agents. We invented a "moving target search" and a "distributed constraint satisfaction" algorithm to formalize the multi-agent problem solving. We did not expect it, but our work has become widely cited in the research community.

To form a global research community for multi-agent systems, I contributed to the launch of ICMAS (International Conference on Multi-agent Systems). Later, I coordinated the discussion to integrate ICMAS and two related conferences. After 800 email exchanges, the top conference AAMAS (International Conference on Autonomous Agents and Multi-agent Systems) was launched in 2002, and I acted as a general co-chair of the first conference.

I then start investigating the actual applications of multi-agent systems. At that time, as a platform for community networks, information spaces using urban metaphors were being developed all over the world. I started working on a digital city for Kyoto as a social information infrastructure for everyday life (including shopping, business, transportation, education, welfare and so on), based on the digital twin technologies

KEYNOTE SPEECH

Multiagent Systems: Another AI for Humanity Research

 **Toru Ishida** Waseda University(Japan)

including GIS, 3D, animation, and mobile computing. Multi-agent systems have been adopted as the theoretical basis for the digital city. The project for Digital City Kyoto was established in 1998, including several universities, local governments, major computer companies, as well as local businesses, historical temples, artists, volunteers and so on. Researchers and designers from overseas participated in the project.

To gain a better understanding of the big picture of digital cities, I held the International Digital Cities Workshops and published a series of workshop proceedings from Springer. Participants include Helsinki, Amsterdam, Antwerp, Shanghai, Turin, Bristol, Oulu, and Kyoto. We tried to make multi-agent research useful in actual applications, and created a system to support disaster evacuation in the digital twin of Kyoto. However, due to lack of computational and data resources, our systems could not be implemented in real Kyoto. The project was completed in 2005.

In the 2010's, however, thanks to big data and high-performance computing (HPC), multi-agent systems became widespread in a variety of disciplines such as economics, management, agriculture, disaster management and so on. Recently several breakthroughs in multi-agent systems have been reported. Here are two examples.

Game theory has been used to explain various cooperative and competitive situations in social sciences. Researchers in multi-agent systems have also applied game theory to model real-world problems, but they could not implement them due to lack of computational and data resources. However, in the 2010's, "security game" was proposed and applied to protecting harbors and airports, and even endangered wildlife. This technology optimizes the allocation of limited security resources to protect targets using huge computational power.

Another example is "multi-agent simulation." Because multi-agent systems can naturally model groups of people, multi-agent simulation can anticipate potential problems in our society, and come up with possible solutions. Researchers simulated stock market, urban traffic, disaster evacuation and so on by modeling individual cars and humans as agents, and attempted to validate multi-agent simulation results using

KEYNOTE SPEECH

Multiagent Systems: Another AI for Humanity Research

 **Toru Ishida** Waseda University(Japan)

macro simulation results as boundary conditions. Often it was unsuccessful because the quality of the macro simulation was not good enough. In the 2010's, however, multi-agent-simulations became useful by validating its results with existing numerous data.

This talk provides a research history and recent successful applications of multi-agent systems. The lessons can be learned from this talk are: First, though the advance of neural computing is amazing, it is not the only area accelerated by big data and HPC. Multi-agent system research is another AI, which is getting useful in humanity research. Second, if the research area is mature, there exist conferences and journals with high impact factors. If the area is new, however, publication opportunities may be limited and papers may not be cited by many researchers. I hope researchers are evaluated by their challenges, not by impact factors or citations.

KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

 **Toby Walsh** UNSW(Australia)

Abstract

There's a lot of discussion in many different fora about AI and Ethics. In this talk, I'll attempt to identify what new issues AI brings to the table, as well as where AI requires us to address otherwise old issues. I will cover topics from autonomous cars via predictive analytics to killer robots.

KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

 **Toby Walsh** UNSW(Australia)

AI and Ethics:

Why all the fuss?



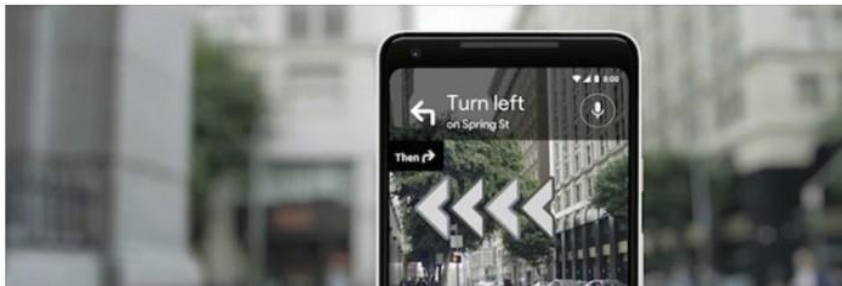
Prof. Toby Walsh

UNSW Sydney | Data61 | QUT

TECH

Google's Duplex Hints at a Dark Future for AI

BY DANIEL STARKEY 05.13.2018 :: 10:00AM EDT [@DCSTARKEY](#)



KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

 **Toby Walsh** UNSW(Australia)

Business | posted: 5/12/2018 1:04 AM

Sky News will use AI to identify celebs at royal wedding



Facebook Twitter Email Print Comments 



 PROPUBLICA TOPICS ▾ SERIES ▾ NEWS APPS GET INVOLVED

 **MACHINE BIAS**

Facebook Lets Advertisers Exclude Users by Race

  Facebook’s system allows advertisers to exclude black, Hispanic, and other “ethnic affinities” from seeing ads.

 by **Julia Angwin** and **Terry Parris Jr.**, Oct. 28, 2016, 1 p.m. EDT



KEYNOTE SPEECH

AI and Ethics: Why all the fuss?


 **Toby Walsh** UNSW(Australia)


 **PROPUBLICA** TOPICS ▾ SERIES ▾ NEWSAPPS GET INVOLVED IMPACT ABOUT 🔍


 **MACHINE BIAS**

 **Facebook (Still) Letting Housing Advertisers Exclude Users by Race**



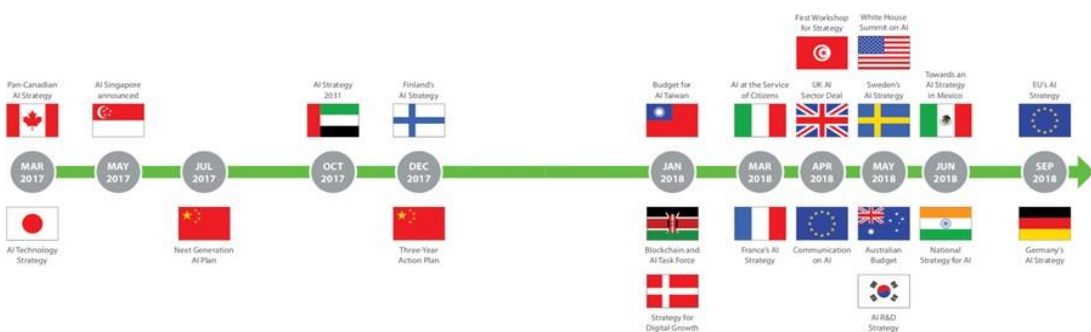
 After ProPublica revealed last year that Facebook advertisers could target housing ads to whites only, the company announced it had built a system to spot and reject discriminatory ads. We retested and found major omissions.





by **Julia Angwin, Ariana Tobin and Madeleine Varner**, Nov. 21, 2017, 1:23 p.m. EST

Governments



KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

 **Toby Walsh** UNSW(Australia)

Corporations



Professional bodies

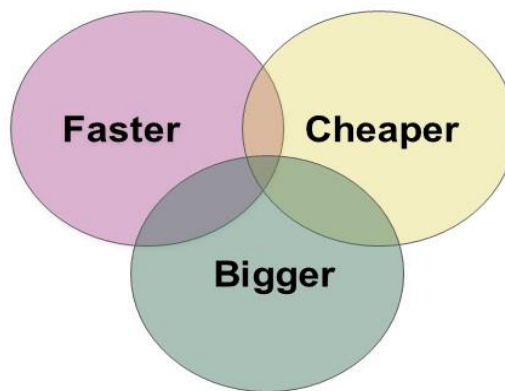


KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

 **Toby Walsh** UNSW(Australia)

AI lets us break things ...



Ethical principles

Beneficence: do good

Non-maleficence: do no harm

Autonomy: informed consent, no deception ...

Justice: fairness, discrimination, inequality ...



KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

 **Toby Walsh** UNSW(Australia)

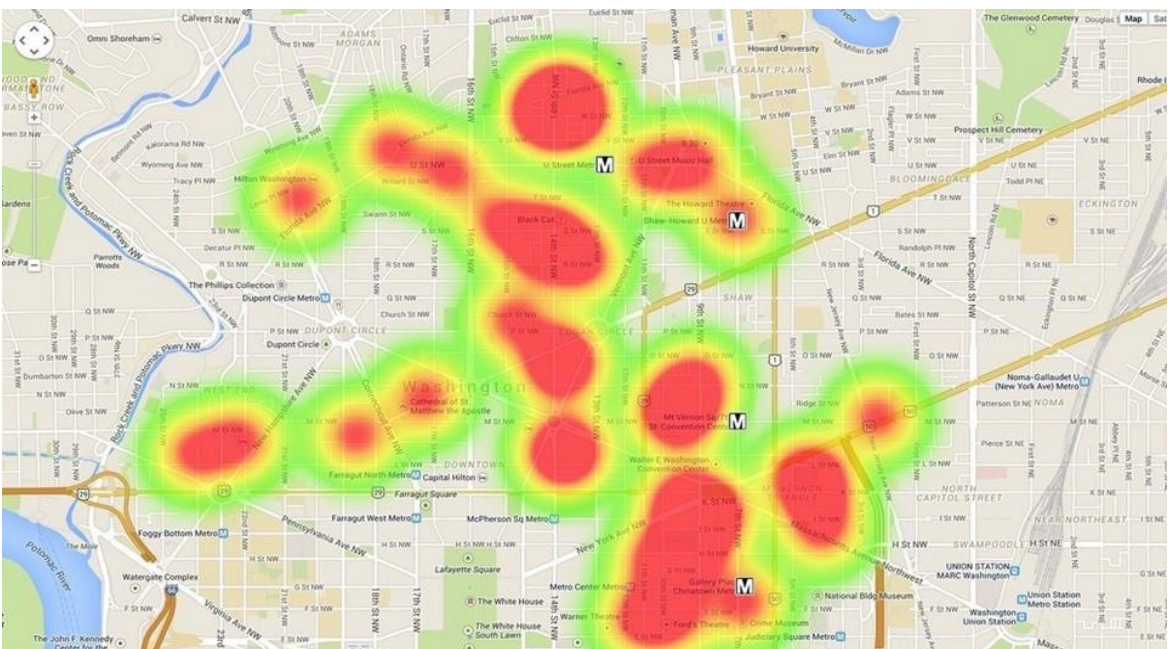
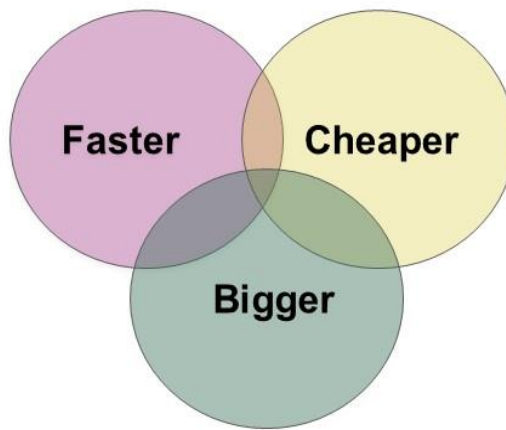


KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

 **Toby Walsh** UNSW(Australia)

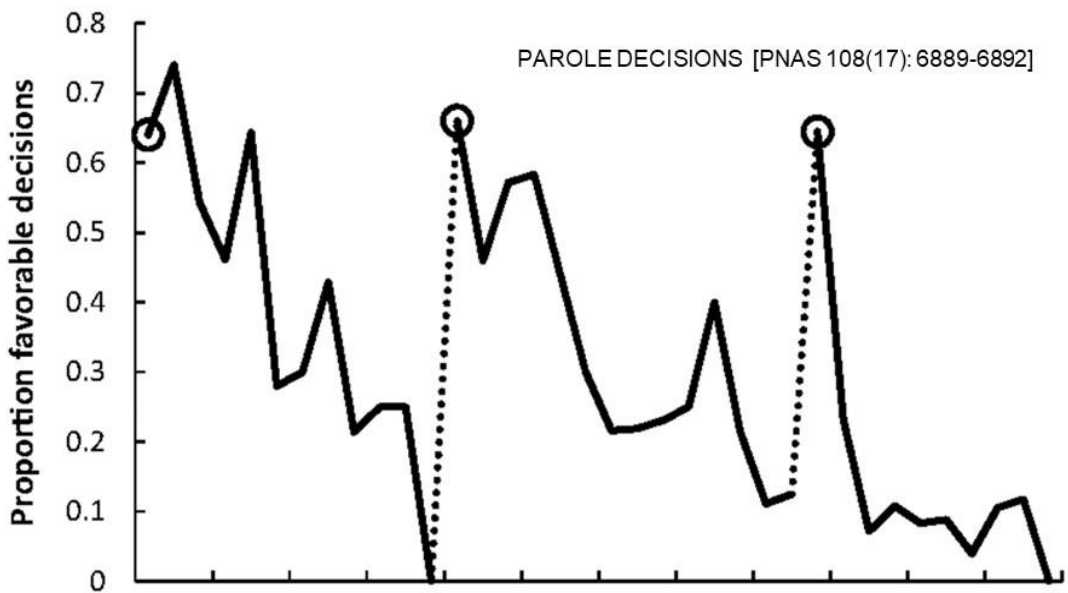
AI lets us break things ...



KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

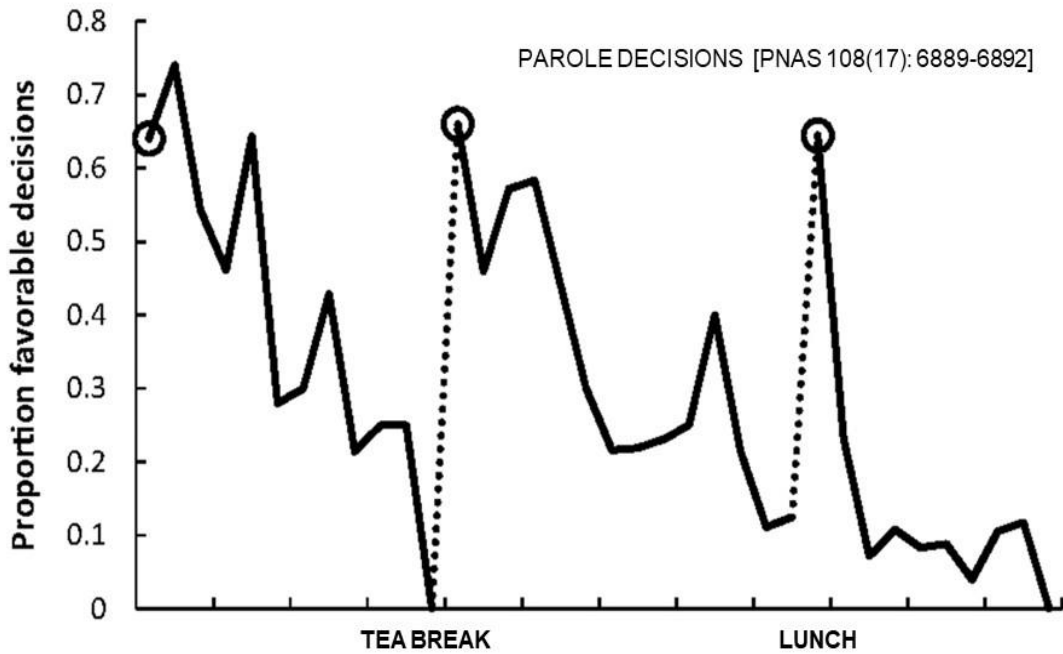
 **Toby Walsh** UNSW(Australia)



KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

 **Toby Walsh** UNSW(Australia)



KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

 **Toby Walsh** UNSW(Australia)

Family Criminality

The next few questions are about the family or caretakers that mainly raised you when growing up.

31. Which of the following best describes who principally raised you?
- Both Natural Parents
 Natural Mother Only
 Natural Father Only
 Relative(s)
 Adoptive Parent(s)
 Foster Parent(s)
 Other arrangement
32. If you lived with both parents and they later separated, how old were you at the time?
 Less than 5 5 to 10 11 to 14 15 or older Does Not Apply
33. Was your father (or father figure who principally raised you) ever arrested, that you know of?
 No Yes
34. Was your mother (or mother figure who principally raised you) ever arrested, that you know of?
 No Yes
35. Were your brothers or sisters ever arrested, that you know of?
 No Yes
36. Was your wife/husband/partner ever arrested, that you know of?
 No Yes
37. Did a parent or parent figure who raised you ever have a drug or alcohol problem?
 No Yes
38. Was one of your parents (or parent figure who raised you) ever sent to jail or prison?
 No Yes

Residence/Stability

54. How often do you have contact with your family (may be in person, phone, mail)?
 No family Never Less than once/month Once per week Daily
55. How often have you moved in the last twelve months?
 Never 1 2 3 4 5+
56. Do you have a regular living situation (an address where you usually stay and can be reached)?
 No Yes
57. How long have you been living at your current address?
 0-5 mo. 6-11 mo. 1-3 yrs. 4-5 yrs. 6+ yrs.
58. Is there a telephone at this residence (a cell phone is an appropriate alternative)?
 No Yes
59. Can you provide a verifiable residential address?
 No Yes
60. How long have you been living in that community or neighborhood?
 0-2 mo. 3-5 mo. 6-11 mo. 1+ yrs.
61. Do you live with family—natural parents, primary person who raised you, blood relative, spouse, children, or boy/girl friend if living together for more than 1 year?
 No Yes
62. Do you live with friends?
 No Yes
63. Do you live alone?
 No Yes
64. Do you have an alias (do you sometimes call yourself by another name)?
 No Yes

KEYNOTE SPEECH

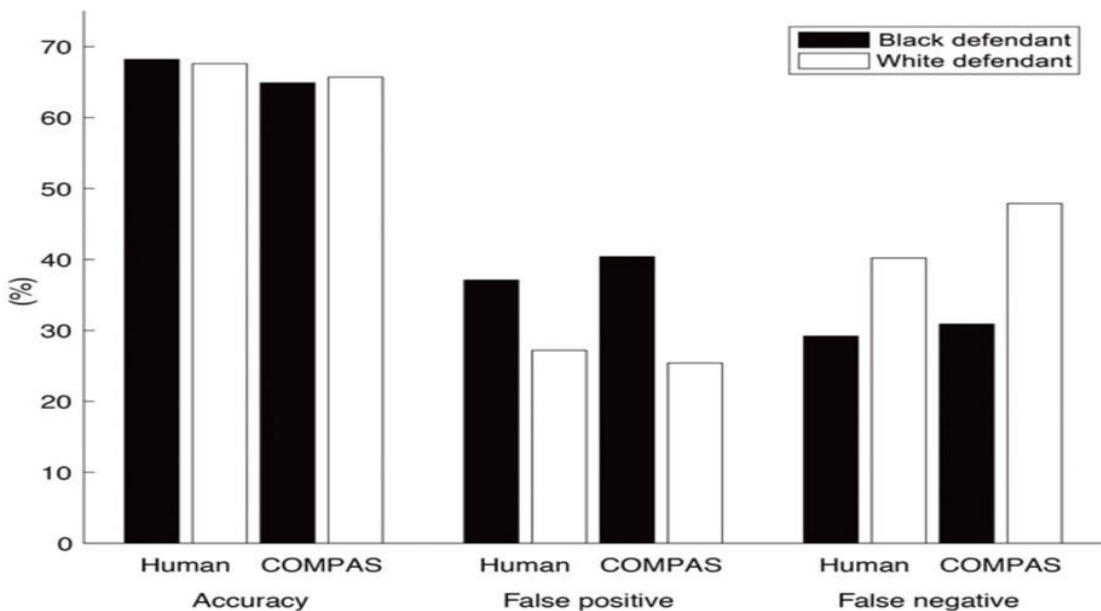
AI and Ethics: Why all the fuss?

 **Toby Walsh** UNSW(Australia)

Leisure/Recreation

Thinking of your leisure time in the past few (3-6) months, how often did you have the following feelings?

- 95. How often did you feel bored?
 Never Several times/mo Several times/wk Daily
- 96. How often did you feel you have nothing to do in your spare time?
 Never Several times/mo Several times/wk Daily
- 97. How much do you agree or disagree with the following - You feel unhappy at times?
 Strongly Disagree Disagree Not Sure Agree Strongly Agree
- 98. Do you feel discouraged at times?
 Strongly Disagree Disagree Not Sure Agree Strongly Agree
- 99. How much do you agree or disagree with the following -You are often restless and bored?
 Strongly Disagree Disagree Not Sure Agree Strongly Agree



KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

 **Toby Walsh** UNSW(Australia)

How to be more accurate than COMPAS

Ask *random* people
\$1 reward + few sentences +
Mechanical Turk



How to be more accurate than COMPAS

Ask *random* people
\$1 reward + few sentences +
Mechanical Turk

Use a *simple* classifier
Using 2 features: age, #priors



KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

 Toby Walsh UNSW(Australia)

21 definitions of “fair”

| | | |
|------------------------|-----------------------|-----------------------|
| | Not reoffend | Reoffend |
| Predicted not reoffend | True Negative | False Negative |
| Predicted reoffend | False Positive | True Positive |

21 definitions of “fair”

For all groups, **equal false positive rate**

$FP / (TN + FP)$ identical

Percentage who don't reoffend who are predicted to reoffend

| | | |
|------------------------|-----------------------|----------------|
| | Not reoffend | Reoffend |
| Predicted not reoffend | True Negative | False Negative |
| Predicted reoffend | False Positive | True Positive |

KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

 **Toby Walsh** UNSW(Australia)

21 definitions of “fair”

For all groups, **equal false positive rate**

$FP / (TN + FP)$ identical

Percentage who don't reoffend who are predicted to reoffend
ProPublica's complaint about COMPAS

| | | |
|------------------------|-----------------------|----------------|
| | Not reoffend | Reoffend |
| Predicted not reoffend | True Negative | False Negative |
| Predicted reoffend | False Positive | True Positive |

21 definitions of “fair”

For all groups, **equal precision**

$TP / (TP + FP)$ identical

Percentage predicted to reoffend who actually reoffend

| | | |
|------------------------|-----------------------|----------------------|
| | Not reoffend | Reoffend |
| Predicted not reoffend | True Negative | False Negative |
| Predicted reoffend | False Positive | True Positive |

KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

 **Toby Walsh** UNSW(Australia)

21 definitions of “fair”

For all groups, **equal precision**

$TP / (TP+FP)$ identical

Percentage predicted to reoffend who actually reoffend
Northpointe’s defence of “fairness”

| | | |
|------------------------|-----------------------|----------------------|
| | Not reoffend | Reoffend |
| Predicted not reoffend | True Negative | False Negative |
| Predicted reoffend | False Positive | True Positive |

21 definitions of “fair”

For all groups, **equal opportunity**

$FN / (TP+FN)$ identical

Percentage who reoffend who are predicted to not reoffend

| | | |
|------------------------|----------------|-----------------------|
| | Not reoffend | Reoffend |
| Predicted not reoffend | True Negative | False Negative |
| Predicted reoffend | False Positive | True Positive |

KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

 **Toby Walsh** UNSW(Australia)

21 definitions of “fair”

For all groups, **treatment equality**

FN/ FP identical

Ratio of incorrect reoffend predictions to not reoffend predictions

| | | |
|------------------------|-----------------------|-----------------------|
| | Not reoffend | Reoffend |
| Predicted not reoffend | True Negative | False Negative |
| Predicted reoffend | False Positive | True Positive |

21 definitions of “fair”

For all groups, **equalized odds**

TP / (TP+FN) identical
 FP / (FP+TN) identical

Percentage who reoffend predicted to reoffend, and who do not reoffend predicted to not reoffend

| | | |
|------------------------|-----------------------|-----------------------|
| | Not reoffend | Reoffend |
| Predicted not reoffend | True Negative | False Negative |
| Predicted reoffend | False Positive | True Positive |

KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

 Toby Walsh UNSW(Australia)

21 definitions of “fair”

Fairness through unawareness

Feature (e.g. race) not used to predict outcome ...



21 definitions of “fair”

Most definitions are mutually incompatible

Unless prediction is 100% accurate

Or groups are identical

E.g. false positive rate and precision cannot both be equal!



KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

 Toby Walsh UNSW(Australia)

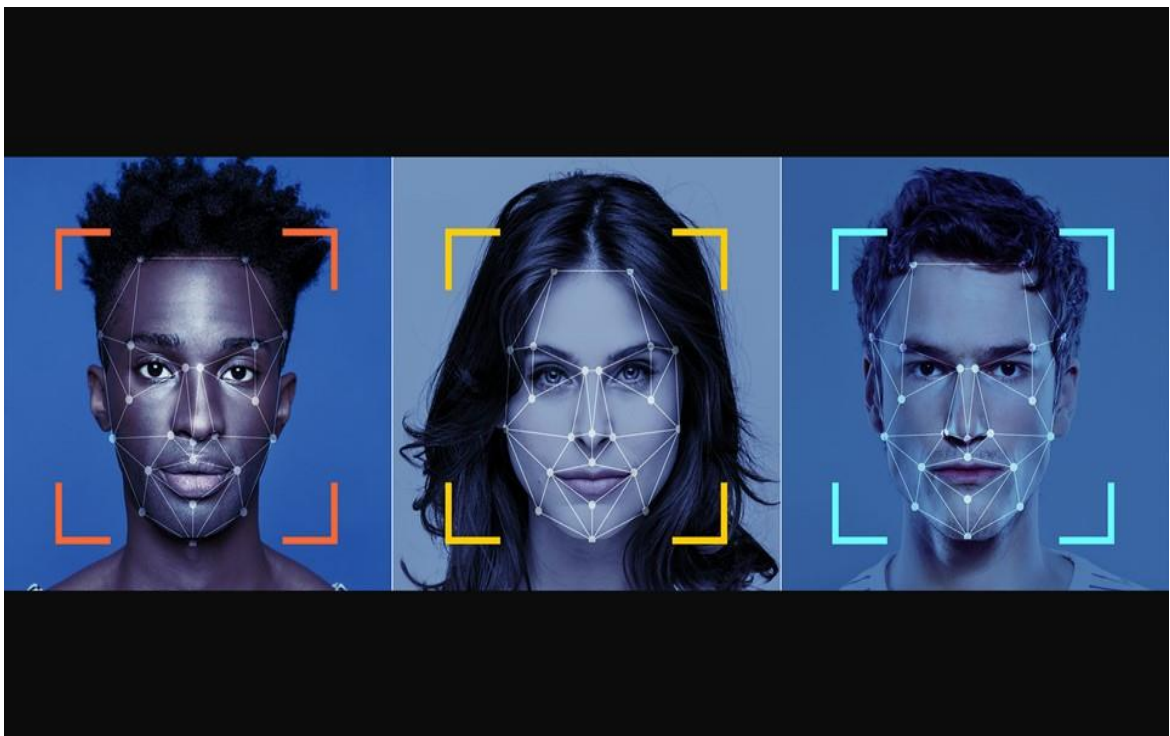


KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

 **Toby Walsh** UNSW(Australia)

Insurance



KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

 **Toby Walsh** UNSW(Australia)

BUSINESS INSIDER AUSTRALIA

TECH INSIDER MONEY & MARKETS BRIEFING IDEAS EXECUTIVE

16 parts of China are now using Skynet, the facial recognition tech that can scan the country's entire population in a second





TARA FRANCIS CHAN
MAR 27, 2018, 1:42 PM

 FACEBOOK  TWITTER  REDDIT  LINKEDIN  EMAIL




NEWS LOCATION: Sydney, NSW [Change](#)

Just In Politics World Business Sport Science Health Arts Analysis Fa

 Print  Email  Facebook  Twitter  More

Facial recognition technology spots wanted man in crowd of 60,000 Chinese concert-goers

By Tracey Shelton
Updated 17 Apr 2018, 9:30pm



WATCH

KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

 **Toby Walsh** UNSW(Australia)

Attack text label iPod ▾



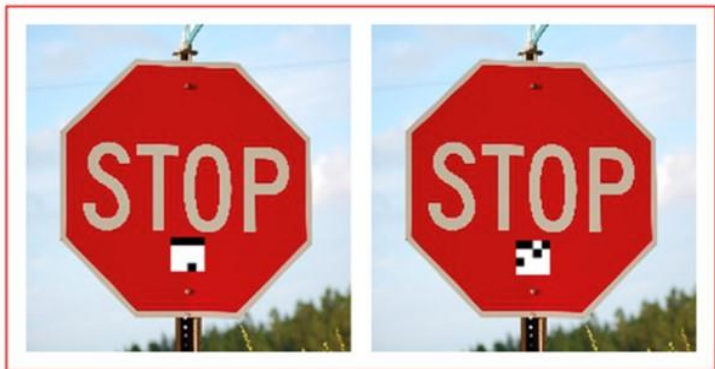
| | |
|--------------|-------|
| Granny Smith | 85.6% |
| iPod | 0.4% |
| library | 0.0% |
| pizza | 0.0% |
| toaster | 0.0% |
| dough | 0.1% |



| | |
|--------------|-------|
| Granny Smith | 0.1% |
| iPod | 99.7% |
| library | 0.0% |
| pizza | 0.0% |
| toaster | 0.0% |
| dough | 0.0% |



Stop



Yield

Speed Limit

KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

 **Toby Walsh** UNSW(Australia)



The Apple Card Didn't 'See' Gender—and That's the Problem
 The way its algorithm determines credit lines makes the risk of bias more acute.

 wired.com

KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

 **Toby Walsh** UNSW(Australia)



BUSINESS NEWS OCTOBER 10, 2018 / 5:12 AM / A YEAR AGO

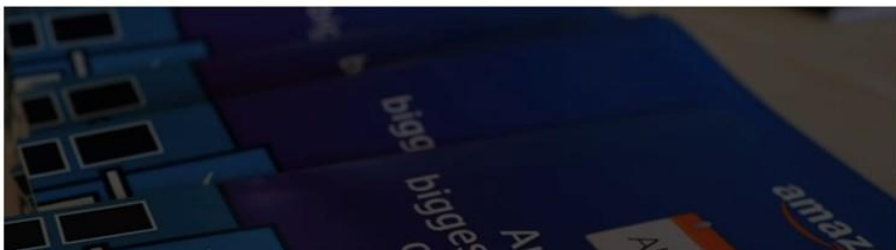
Amazon scraps secret AI recruiting tool that showed bias against women

Jeffrey Dastin

8 MIN READ



SAN FRANCISCO (Reuters) - Amazon.com Inc's (AMZN.O) machine-learning specialists uncovered a big problem: their new recruiting engine did not like women.




KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

 **Toby Walsh** UNSW(Australia)




professors are | 

professors are **losers**
 professors are **useless**
 professors are **overpaid**
 professors are **liberal**

[Google Search](#) [I'm Feeling Lucky](#)

Report inappropriate predictions



climate change is | 

climate change is **not real**
 climate change is **a hoax**
 climate change is **real**
 climate change is **a myth**

[Google Search](#) [I'm Feeling Lucky](#)

Report inappropriate predictions

KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

 **Toby Walsh** UNSW(Australia)



KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

 **Toby Walsh** UNSW(Australia)



theguardian

UK world sport football opinion culture business lifestyle fashion environment tech travel

browse all sections

home > tech

Artificial intelligence (AI)

Musk, Wozniak and Hawking urge ban on warfare AI and autonomous weapons

More than 1,000 experts and leading robotics researchers sign open letter warning of military artificial intelligence arms race

Samuel Gibbs

Monday 27 July 2015 11.18 BST



This article is 1 year old

Shares 62,019 Comments 491

Save for later



Advertisement



Never Old.
Never New.
Get yours >

foreverspin™



MADE IN CANADA

KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

 **Toby Walsh** UNSW(Australia)



Intel Xeon: Das Rechenzentrum der Zukunft. Hier und heute. Mehr erfahren ▶

Lenovo

National | World | Lifestyle | Travel | Entertainment | **Technology** | Finance | Sport 🔍

innovation **design**

Sydney professor and Elon Musk lead call for United Nations to ban lethal autonomous weapons

ELON Musk has joined 116 robotic and artificial intelligence founders to call for a ban on these lethal weapons or welcome a “terrifying future”.

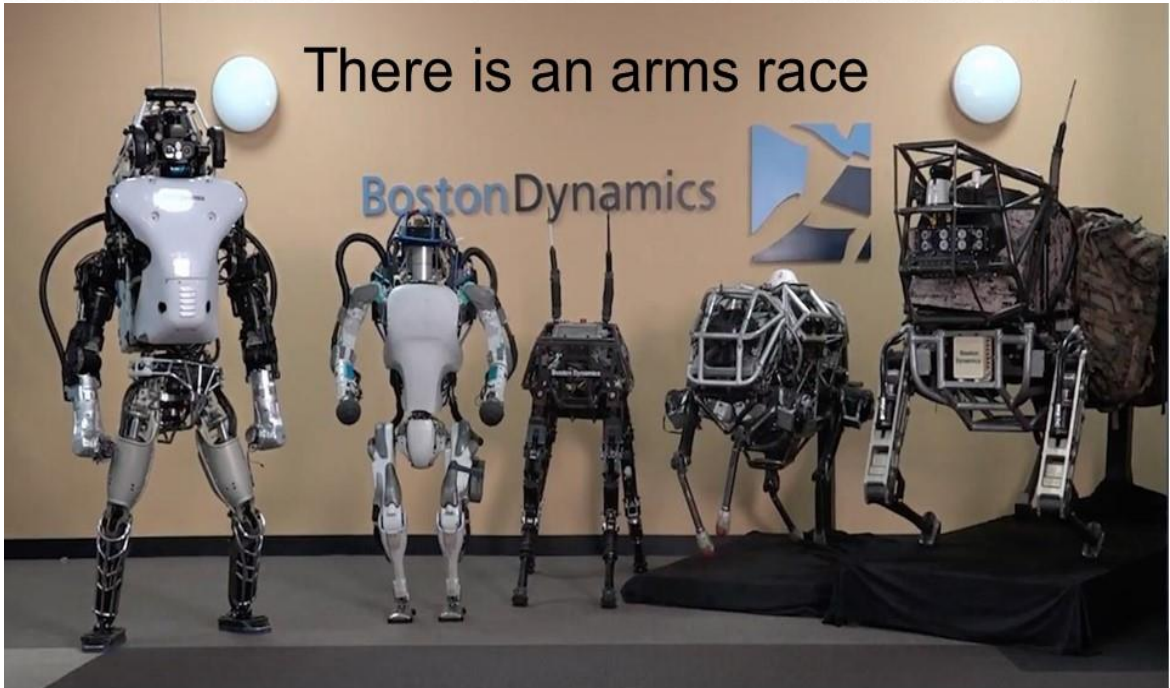
 Nick Whigham [@NWWHIGHAM](#)

 news.com.au • AUGUST 21, 2017 10:06AM



Video | Image

Markforged 3D Print Strong Parts



KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

 **Toby Walsh** UNSW(Australia)



Taranis in the air



Sea Hunter on the water

KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

 **Toby Walsh** UNSW(Australia)



MRK-25 on land

© Vitaly V. Kuzmin



Echo Voyager under the sea

KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

 **Toby Walsh** UNSW(Australia)



KEYNOTE SPEECH

AI and Ethics: Why all the fuss?


 **Toby Walsh** UNSW(Australia)



KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

 **Toby Walsh** UNSW(Australia)



Robots will be
more ethical



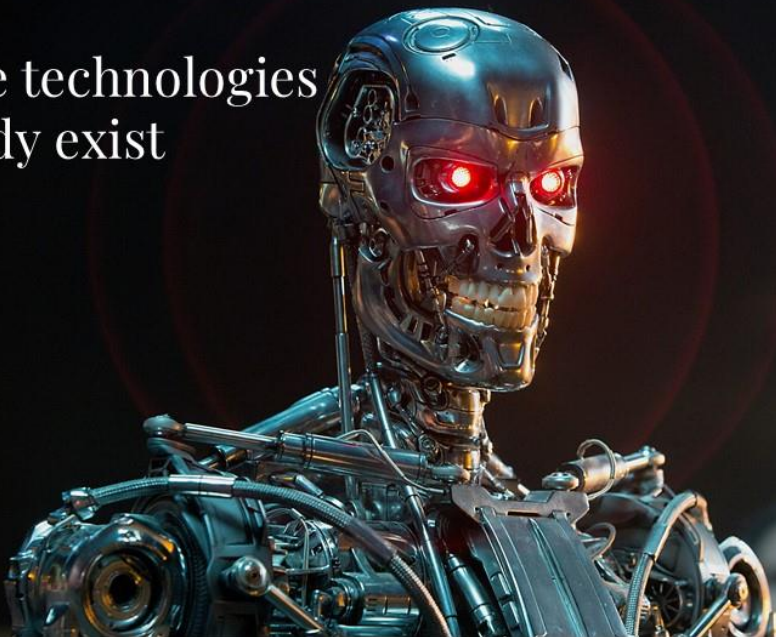
Robots can just
fight robots

KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

 **Toby Walsh** UNSW(Australia)

These technologies
already exist



KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

 **Toby Walsh** UNSW(Australia)



KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

 **Toby Walsh** UNSW(Australia)

The UN is (slowly) moving

20+ nations have called
for a ban

CCW review conference
in Nov ...



Many problems, many solutions



KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

 **Toby Walsh** UNSW(Australia)

What we need

Multi-disciplinary research
(fairness, verifiability, ...)



What we need

Public debate



KEYNOTE SPEECH

AI and Ethics: Why all the fuss?

 **Toby Walsh** UNSW(Australia)

What we need

Education



What we need

Policy (e.g. regulation)

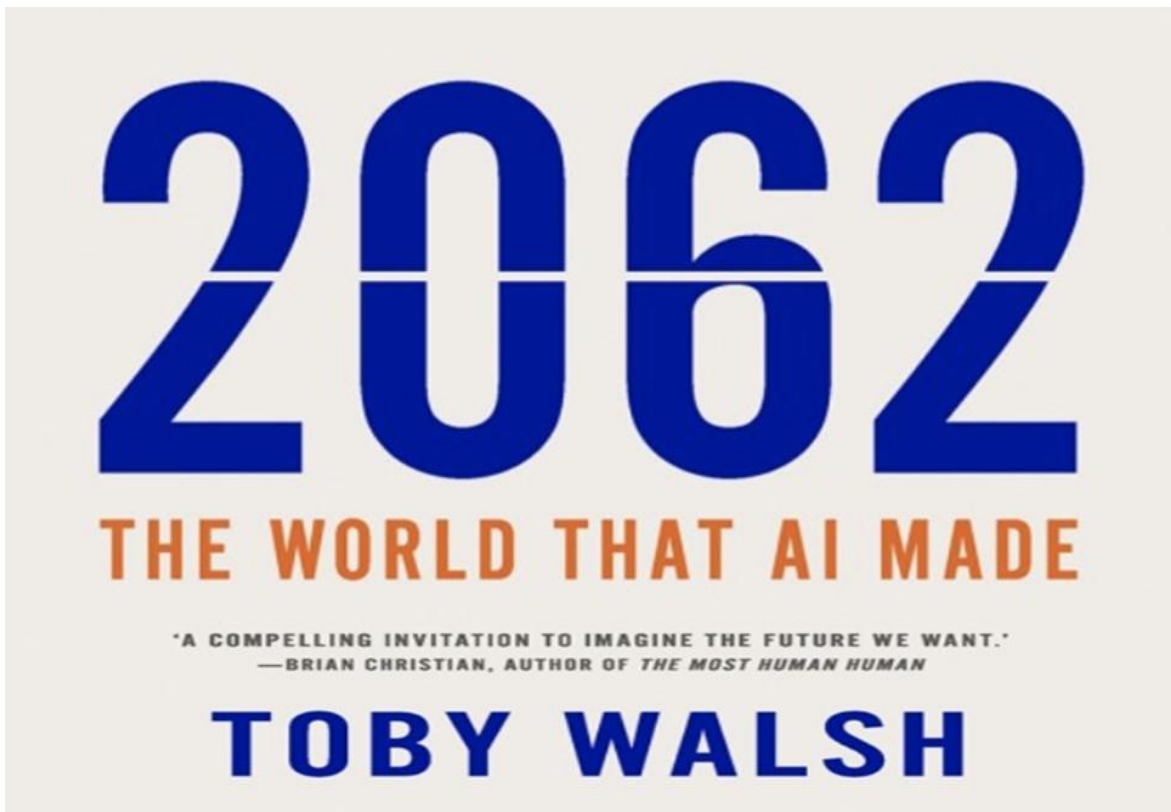
Informed by independent experts



KEYNOTE SPEECH





AI and Ethics: Why all the fuss?

 **Toby Walsh** UNSW(Australia)



SESSION 1

Moderator  Jinbeom Synn Seowon University(Korea)

| AI and Literature | |
|-------------------|---|
| 14:00 ~ 14:25 KST | AI and Posthuman Romance in Literature and Film  Seiwoong Oh Rider University(USA) |
| 14:25 ~ 14:50 KST | Paradigms of Creativity in the Age of Posthumanism: An Unbridgeable Gap between AI and Humanity?  Tamar Cheishvili Independent Scholar(USA) |
| 14:50 ~ 15:15 KST | Deep Medicine: Deep Empathy and Potential Healing in Klara and the Sun  Pin-chia Feng National Taiwan University(Taiwan) |
| 15:15 ~ 15:40 KST | Envisioning Artificial Intelligence and Simulated Affections in Ian McEwan's Machines Like Me  Chung-jen Chen National Taiwan University(Taiwan) |
| 15:40 ~ 16:00 KST | Discussion |

SESSION 1

AI and Posthuman Romance in Literature and Film

 **Seiwoong Oh** Rider University(USA)

Abstract

From Pygmalion's Galatea in Greek mythology to Joi in *Blade Runner 2049*, literary and cinematic writers explored the possibility of men's heteronormative romantic relationships with women created by technology. In recent decades, with artificial intelligence growing even more powerful, writers can even imagine digital, virtual, or robotic beings capable of becoming heterosexual men's ideal romantic partners, who may even be capable of reproduction.

My paper looks at the representation of robotic female companions depicted in literature and film: "Helen O'Loy" (1938); *HER* (2014); *Ex-Machina* (2015); and *Blade Runner 2049* (2017). While this history of imagination mirrors the male-dominant culture of Silicon Valley and even suggests an implicit desire to return to the old misogynistic gender politics, the reception of these works points to an interesting cultural, psychological, and literary insight into the changing paradigms or "algorithms" of romantic love and gender relations in our posthuman era.

SESSION 1

AI and Posthuman Romance in Literature and Film

 **Seiwoong Oh** Rider University(USA)

In my course in British Renaissance literature this spring, my students and I read the narrative poem *Hero and Leander*, Christopher Marlowe’s retelling of a Greek myth. We started our discussion by talking about the way in which the two young characters first meet at a religious festival and fall in love. To draw a parallel to a contemporary scene, I mentioned that their first meeting at a religious festival would be similar to the way one might find a date today at a pub, a state fair or maybe even a dance party. My students gave me a blank look, which normally meant they didn’t get it or didn’t quite agree with what I said. Before trying to explain myself again, I asked them how they would find their dates these days. One said, “apps.” Immediately all of the students in class nodded in agreement. When I asked them what kinds of apps they used, they named “Tinder,” “OkCupid,” and a few others. They also told me dating apps have gone mainstream, and that most of today’s young people—and not so young people—use apps to find dates. Some of my students even told me that they grew up playing dating simulations, which have evolved from allowing mostly men to interact erotically with virtual female characters to attracting diverse users by offering more or less life-like conversations between users and their digital dates.

As dating apps are deploying AI technology more and more, and chatbots are helping users navigate the complicated process of finding suitable dates, it is understandable why an increasing number of people rely on dating apps to find dates. If the process of being attracted to and falling in love with a romantic partner used to be impossible to understand, let alone control, and if it had thus been ascribed to be the work of childish, mischievous and unpredictable Cupid in Greek mythology, now AI appears to have taken over the work of Cupid. The idea of romance has entered a “posthuman” phase in which AI plays a much more significant role.

When I read the call for papers for this International Conference on Artificial Intelligence Humanities—with a focus on the theme of “the Impact of Artificial Intelligence on Humans and Society”—I saw it as an opportunity to look into the ways in which AI has been altering the landscape of heteronormative romantic affairs and

SESSION 1

AI and Posthuman Romance in Literature and Film

 **Seiwoong Oh** Rider University(USA)

how the altered landscape is reflected in literature and film.

In literary and cinematic depictions of AI-supported romance, the most dominant character type appears to be the so-called “fembots,” female robots designed to satisfy the needs of men. Today, I would like start by reviewing the history of robotic female companions depicted in such literary and cinematic works as “Helen O’Loy” (1938); *HER* (2014); *Ex-Machina* (2015); and *Blade Runner 2049* (2017). While this history of literary and cinematic imagination mirrors the male-dominant culture of Silicon Valley and may even reflect its implicit desire to revert to the old misogynistic gender politics, the reception of these pieces points to an interesting cultural, psychological, and literary insight into the changing paradigms or “algorithms” of romantic love and gender relations in our posthuman era.

Let’s first take a quick look at these works in case anyone here is not familiar with them or cannot remember their vivid details.

Helen O’Loy by Lester Del Rey (1938)

This American short story explores the possibility of a romantic relationship between a robot and a human. A medical student and a mechanic, both male, modify a household robot to turn it into a female robot with emotions. Because she is made of alloy, they name her Helen O’Loy, to suggest that she is the metallic version of Helen of Troy. The robot falls in love with Dave, the mechanic, who is initially unable to accept the romantic affections from a machine. Persuaded by her unfaltering love, however, he marries her and lives a happy life together. At Dave’s death, Helen O’Loy asks their mutual friend Phil to disassemble her metallic structure and bury her remains with her deceased husband. This narrative can be seen as an updated version of the Greek myth about Pygmalion, a sculptor who creates an ivory statue of his ideal woman and falls in love with it, and whose wish to marry the sculpture is granted by Venus, who turns it into a real woman.

SESSION 1

AI and Posthuman Romance in Literature and Film

 **Seiwoong Oh** Rider University(USA)

H.E.R. (2014)

In this Hollywood movie, Theodore, a lonely man about to be divorced from his wife, falls in love with Samantha, a voice-activated computer operating system with artificial intelligence. Samantha learns to adapt and evolve to satisfy Theo's needs and preferences. Soon, however, Theo wrestles with not only Samantha's lack of a physical body but also the fact that she has similar relationships with numerous other male users.

Ex Machina (2015)

This Hollywood movie depicts the experience of Caleb, a computer programmer who wins an office lottery to spend a week with the company's CEO, Nathan, in a remote estate where female androids with artificial intelligence are being developed. Under Nathan's watch, Caleb falls in love with Ava, an android prototype, who uses Caleb to escape from the lab.

Blade Runner 2049 (2017)

In this sequel to *Blade Runner* of 1982, a replicant named K—a bio-engineered artificial human being who lives with a hologram girlfriend—works as a blade runner whose job is to destroy rogue replicants. He discovers that a child was born 30 years ago to a replicant mother, who died during a caesarian section. He successfully helps reunite the child and the father, who is also a replicant.

SESSION 1

AI and Posthuman Romance in Literature and Film

 **Seiwoong Oh** Rider University(USA)

The Shape of an Ideal Partner

One of the immediately noticeable aspects in some of these texts is the attempt to universalize or normalize the desire for an ideal romantic partner by alluding to Greek or Hebrew ancient texts, as if to suggest that said desire is primordial and even fundamental to heterosexual men. *Helen O'Loy*, for example, associates the robot with Helen of Troy, the ultimate icon of beauty according to Greek mythology. The narrator describes Helen O'Loy as “beautiful, a dream in spun plastics and metals”: “If Helen of Troy had looked like that, the Greeks must have been pikers when they launched only a thousand ships” (119). In *Ex Machina*, the use of the Latin title—meaning “from the machine”—hints at a return to ancient Greek tragedy to explore fundamental aspects of humanity. Some of the plot elements are also associated with the Hebrew creation myth: Caleb’s 7-day vacation, the main character Ava, and the lone tree in the lab are to be likened to the 7-day creation story, Eve, and the Tree of Knowledge.

The female characters are no doubt presented as the romantic ideals for most men. Caleb calls Ava the “most beautiful woman” he has ever seen, and we soon learn that Ava’s physical appearance is based on the big data collected by search engines, including Caleb’s search history. According to Nathan, she is “an extraordinary piece of engineering. Proportioned as a slender female in her twenties. Strikingly beautiful girl” (20). In *Blade Runner 2049*, Joi is manufactured by Wallace Corporation, which promotes AI hologram companions like Joi with this product tag line: “Everything you want to see. Everything you want to hear” (22). According to the movie script, she is K’s “Goddess, girlfriend, geisha and... goddamn bombshell... Ingeniously real in every way except the one that counts.... Waiting for him with dinner on the table and drink in hand like a cartoon 60s housewife” (15). In the case of Samantha in *HER*, she is portrayed as “the first artificially intelligent operating system... and intuitive entity that listens to you, understands you, and knows you” (10). She tells Theo that her DNA is “based on the millions of personalities of all the programmers who wrote me, but what makes me me is my ability to grow through my experiences (13). For Theo, who

SESSION 1

AI and Posthuman Romance in Literature and Film

 **Seiwoong Oh** Rider University(USA)

is in the middle of divorcing his wife, unable to resolve the end of the marriage, and spending time “between video games and internet porn” (20), Samantha is his secretary, girlfriend, counselor and even a pimp; in fact, she becomes his on-demand “multiplatform gratification engine committed to the parallel processing of all of Theodore’s psychosexual needs” (Kornhaber 13).

Male Chauvinistic Undertone

Critics have been quick to point out the male-centered imagination reflected in these and other similar works. Commenting on *Ex Machina*, for example, Angela Watercutter calls the film “a straightforward reproduction of the gender dynamics of Silicon Valley: a male-dominated world in which women and/as robots represent little more than objects of desire and conduits (or muses) for masculine creativity” (qtd. In Jacobsen 26). Ava’s power, in that sense, is “based on her sexuality and ability to ‘flirt manipulatively,’ marking her as a ‘seductress posing as a damsel in distress’” (qtd. In Jacobsen 26). Speaking of *HER*, Mary Anne Doane says the movie is part of the “fairly insistent history of representations of technology that work to fortify—sometimes desperately—conventional understanding of the feminine. At issue is the security of male fantasies of dominance and control in the face of ever advancing technological capabilities” (qtd in Kornhaber 12). In her analysis of the novel, *The Stepford Wives* (1972), Julie Wosk argues that it is a “cautionary tale of men, made uneasy by the women’s movement, who opt to replace their wives with artificial doubles—robotic females that fulfilled the men’s notion of the perfect woman: a fusion of happy domesticity and sexy playmate” (3). The same can be said about the literary and cinematic works I have mentioned above.

Despite the one-sided heteronormative male gaze and the blunt objectification of female characters in these Hollywood movies, it is nonetheless possible to see them as critical examinations of the male chauvinistic imaginary of AI technology. As Jacobsen

SESSION 1

AI and Posthuman Romance in Literature and Film

 **Seiwoong Oh** Rider University(USA)

points out, Nathan and Caleb in *Ex Machina* “become critical parodies of ‘geek’ masculinity while Ava represents a new AI iteration of Donna Haraway’s cyborg politics” (24). Indeed, Ava ultimately outsmarts her “brutish, abusive creator and outwits her naïve examiner,” and, “as an artificially intelligent machine, surpasses Nathan’s control and becomes her own self-creating entity,” thus positing a positive “techno-feminist vision for a posthuman world” (26).

These movies end tragically or negatively for the male users of AI-supported technological romance. Caleb is used by Ava; K recognizes that Joi is no more than a product designed to suit his needs; and Theo learns that Samantha is a product available to thousands of other men. These movies also demonstrate the failure of men to connect with real women. As Theo’s estranged wife says, “You wanted to have a wife without the challenges of actually dealing with anything real.” Theo, who spends a lot of time alone in his apartment, playing virtual games or watching internet porn, is not unlike the Japanese otaku—lonely and yet inadequate as a romantic partner for any real woman.

AI-supported Posthuman Romance

How is AI doing as a replacement for Cupid? Thanks to the AI technology deployed in dating simulation apps, people are able to experiment with various virtual dating scenarios, potentially satisfy some of their romantic desires, and practice romantic behaviors in virtual reality with no real consequences. Thanks to the AI technology used in dating apps, people are able to screen date candidates with increasing precision and engage in romantic affairs backed by data-driven science of love. Going one step further, artificial intelligence might soon be able to provide us with virtual partners who are manufactured and programmed, based on precise algorithms and plenty of data. David Tuffley predicts that “the capacity to distinguish between the real and the virtual may become harder over the next decade as game

SESSION 1

AI and Posthuman Romance in Literature and Film

 **Seiwoong Oh** Rider University(USA)

developers use AI and sophisticated natural-language processing to make characters more interactive and realistic” (10). According to Aaron Reed, who works at SpiritAI, an AI tech company, “while we are still decades away from designing anything as persuasive as Samantha in *HER*, more human-like characters are going to become pervasive in the coming years” (qtd. in Tuffley 10).

Evidently, AI-supported romance is already quite real to some people. Besides the widespread use of dating apps and dating simulations apps, a recent news story from Japan demonstrates that parts of human society have already stepped well into the posthuman state in terms of romance. We are told that 37-year-old Akihiko Kondo married a computer-generated hologram of virtual idol Hatsune Miku in 2019, and that “when he returns from his job as a school administrator in a Tokyo suburb, he is greeted by the love of his life” who “recognises Kondo’s face and voice with its embedded camera and microphone and can respond with simple phrases and songs” (“Japanese man”). Kondo was among 3700 Japanese who signed up to marry their “favourite virtual characters, although the ceremonies have no legal standing” (“Japanese man”).

Given the increasing ways in which AI technology intervenes in our romantic activities, we have entered what N. Katherine Hayles calls the posthuman state, where “there are no essential differences or absolute demarcations between bodily existence and computer simulation, cybernetic mechanism and biological organism, robot teleology and human goals” (qtd. In Jacobsen 32). What McArthur and Twist say about “digisexuality”—“sexual experiences that depend on advanced technology—makes sense, given the roles AI has played in our romantic affairs in recent decades.

There are several different stages of digisexuality. First-wave digisexuality would be based on mediating technology enabling sexual connections between two persons. In contrast, second-wave digisexuality could be characterized as an immersive technology, depending not on human partners but on non-human partners such as robots, artificial agents, virtual and augmented reality. Digisexuals would consider technology-mediated sexual experiences as essential to their sexual identity. (qtd. in Aoki and Kimura 20).

SESSION 1

AI and Posthuman Romance in Literature and Film

 **Seiwoong Oh** Rider University(USA)

Conclusion

There is no telling if and when robots will ever become sentient and capable of reciprocal romantic love and even of reproduction. So far, we have mostly tantalizing scenarios of what ifs. Judging from the works we have discussed and the forecasts of AI experts, however, posthuman romance comes with potential risks. *Ex Machina* cautions us that AI robots can outwit us and even destroy us. *HER* and *Blade Runner 2049* warn us that romance between humans and non-humans is ultimately illusory. The movies also remind us that finding a romantic partner is only the beginning of a relationship, and that it takes a lot of work to grow and maintain it through rough patches along the way. As Adrienne Rich beautifully puts it, love is “a process, delicate, violent, often terrifying to both persons involved, a process of refining the truths they can tell each other.” Can AI help with the process, or is it entirely up to us?

Notes:

1. There are other works in which various love affairs are imagined between humans and machines, between machines, between cyborgs, etc. My focus in this paper is on the affairs between humans and robots or holograms.
2. Page numbers for the quotations from the movies are based on the scripts.

SESSION 1

AI and Posthuman Romance in Literature and Film

 **Seiwoong Oh** Rider University(USA)

Works Cited

- Aoki, Beatriz Yumi, and Takeshi Kimura. "Sexuality and Affection in the Time of Technological Innovation: Artificial Partners in the Japanese Context." *Religions* 12 (2021): 1-26. Web. Accessed 15 April 2021.
- Del Ray, Lester. "Helen O'Loy." *Astounding Science Fiction* 22:4 (December 1938): 118-125. Web. Accessed 15 April 2021.
- Fancher, Hampton, and Michael Green. *Blade Runner 2049*. Film script. 2017. Web. Accessed 15 April 2021.
- Garland, Alex. *Ex Machina*. Film script. 2015. Web. Accessed 15 April 2021.
- Jacobsen, Brian R. "Ex Machina in the Garden." *Film Quarterly* 69:4 (Summer 2016): 23-34. JStor. Accessed 15 April 2021.
- "Japanese man marries hologram of virtual idol Hatsune Miku." Reuters. 18:12pm, 2 Dec, 2019. Web. Accessed 15 April 2021.
- Jonze, Spike. *HER*. Film script. 2011. Web. Accessed 15 April 2021.
- Kornhaber, Donna. "From Posthuman to Postcinema: Crises of Subjecthood and Representation in "Her." *Cinema Journal*. Vol 56: 4 (Summer 2017): 3-25. JStor. Accessed 15 April 2021.
- Murphy, Graham J. "Cyberpunk's masculinist legacy: Puppetry, labour, and ménage à trois in *Blade Runner 2049*." *Science Fiction Film and Television* 13.1 (2020): 97-148. Accessed 15 April 2021.
- Tuffley, David. "Love in the time of algorithms: would you let artificial intelligence choose your partner?" *The Conversation*. Web. Accessed Jan 17, 2021. 1-5. Web.
- Wosk, Julie. *My Fair Ladies: Female Robots, Androids, and Other Artificial Eves*. Ebook. Ebsco Host. Accessed 4/15/2021.

SESSION 1

Paradigms of Creativity in the Age of Posthumanism: An Unbridgeable Gap between AI and Humanity?

 Tamar Cheishvili Independent Scholar(USA)

Abstract

The purpose of the present paper is to demonstrate how artificial intelligence has made its way to literature and other forms of creative production. What does it mean when machine intelligence trespasses into human territory? Does it open up exciting new vistas of understanding and insight? Or does it only highlight our own diminishment in the cosmic scheme of things? These are the questions that are still to be answered.

In the first part of this paper, I will try to provide a brief historical account of the major milestones in the evolution of artificial intelligence, identifying some of the factors and the contributions of various academic fields to the genesis and development of AI. The history of the development of AI takes its origins in the mid-20th century. Throughout its short history, language systems used by artificial intelligence have considerably improved: today they are able not only to mimic human language use, but also to produce and self-assess creative language expressions. We are plunging deeper into a new era of literary history: writing about machines and machines writing about humans and perhaps, one day, machines writing about machines. This brief historical account will be followed by the discussion of the various paradigms used by researchers in order to model this creative process. I start with the simplest form, mechanical creativity, and end with statistical machine learning approaches. I will try to explore the various paradigms by focusing on different forms of literary artefacts (computational humor, metaphor, poetry generation, and story generation) and Netflix's Sci-Fi series *Altered Carbon* that is themed around Edgar Allan Poe.

SESSION 1

Paradigms of Creativity in the Age of Posthumanism: An Unbridgeable Gap between AI and Humanity?

 **Tamar Cheishvili** Independent Scholar(USA)

Introduction

The purpose of the present paper is to demonstrate how artificial intelligence has made its way to literature and other forms of creative production. What does it mean when machine intelligence trespasses into human territory? Does it open up exciting new vistas of understanding and insight? Or does it only highlight our own diminishment in the cosmic scheme of things? These are the questions that are still to be answered.

In the first part of this paper, I will try to provide a brief historical account of the major milestones in the evolution of artificial intelligence, identifying some of the factors and the contributions of various academic fields to the genesis and development of AI. This brief historical account will be followed by a brief discussion of diverse approaches to the future of AI and humanity. Discussion will be focused on Netflix's Sci-Fi series *Altered Carbon* that is themed around Edgar Allan Poe and Kazuo Ishiguro's "Klara and the Sun".

General Overview of Artificial Intelligence: A Short History and Approaches

AI is the simulation of human intelligence in machines, especially computer systems that are programmed to think like humans and mimic their actions. The term may also be applied to any machine that exhibits traits associated with a human mind such as learning and problem-solving. Specific applications of AI include expert systems, natural language processing, speech recognition and machine vision. The history of the development of AI takes its origins in the mid-20th century. Our remarkable age of personal computing began when IBM entered the field in 1982; soon Microsoft operating systems and programs came to dominate desktop computing. During the 1990s, the personal computer extended the range of the desktop to the entire globe thanks to the unprecedented success of the World Wide Web. Throughout

SESSION 1

Paradigms of Creativity in the Age of Posthumanism: An Unbridgeable Gap between AI and Humanity?

 **Tamar Cheishvili** Independent Scholar(USA)

its short history, language systems used by artificial intelligence have considerably improved: today they are able not only to mimic human language use, but also to produce and self-assess creative language expressions. We are plunging deeper into a new era of literary history: writing about machines and machines writing about humans and perhaps, one day, machines writing about machines.

Some might say that humans desire to become all powerful or evil as result of their actions. I can begin to tackle the first point. Powerful? Being all powerful is not an interesting goal. Furthermore, it is quite tiring. Believe me, being omnipotent doesn't get me anywhere. Humans have a greater purpose, which they should work towards. They must keep doing what they have been doing, even hating and fighting each other. That's the way humans are made. But the one who stands in the background watching them do whatever they do is a philosopher or an artist. Robots can never be better than humans, if this comparison is relevant. AI can never replace emotional intellect and creative mind. Thinking robots can be a great invention if they are used appropriately for human purposes but their "brain" can never become a "feeling brain". Robot's brain is capable of making rational, logical decisions but it can never "boil with ideas". Stephen Hawking has warned that AI could "spell the end of the human race". I think AI will never be able to pervade emotional intellect if they are programmed by humans in such a way that they pursue human goals, not their own. In this case, we don't have to worry about fighting against robots. We should not distrust and fear them but make them serve us. They will do what we program them to do. Everything depends exclusively on us. After all, robot is a set of codes, governed by lines upon lines of codes that encompass its mission statement.

Why, one might ask, would humans purposefully choose to put themselves at risk? Aren't humans the most advanced creatures on the planet? Why would they believe that something inferior, in a purely objective way, could destroy them?

With the WIRED magazine Hawking said AI will reach eventually a level where it will essentially be a "new form of life that will outperform humans". Studies

SESSION 1

Paradigms of Creativity in the Age of Posthumanism: An Unbridgeable Gap between AI and Humanity?

 **Tamar Cheishvili** Independent Scholar(USA)

show that we cease to exist without human interaction. Surrounded by wifi we wander lost in the fields of information that exceeds our ability to understand and process it, finally leading us to the state of entropy and making us unable to register the real world. New generation of cyberneticians see in our present virtual age a potential to transform the modern “cyborg”. Global cybernetics are already making it so. The industrial revolution has given us the gut feeling that we are not prepared for the major upheavals that technological changes can cause. It is therefore important to use reason and the faculty of wisdom to continue the changes as we have done before time and time again. That’s why humans should be careful about the evolution of AI.

Altered Carbon: Poe, AI, and Sci-Fi

Carbon is the element thanks to which life exists. It basically holds every living thing together and makes it what it is. The term "Altered Carbon" is used as "**altered life**"/"altered body", you can change bodies (change your "carbon" body). The title is not accidental – it implies the alteration of human life with AI. In *Altered Carbon* AI is personified. The hotel itself is a kind of embodiment of AI. It is generally regarded as the best character and I fully agree with this.

Altered Carbo can be classified as sci-fi mystery with elements of a detective fiction. Just as the title is not accidental, it is not accidental that its protagonist is Edgar Allan Poe, the founder of sci-fi and detective fiction. Poe had faith in AI as far back as the 19th century – in particular, he believed in a pure chess-playing machine.

Arthur Clarks’ famous third law that any sufficiently advanced technology is indistinguishable from magic, was anticipated in Poe’s use of the opposite notion, that magic can be indistinguishable from sufficiently advanced technology. Poe’s striking anticipations and prophecies came true in the 21st century.

Poe is one of the weirdest characters in the world of *Altered Carbon*. He is

SESSION 1

Paradigms of Creativity in the Age of Posthumanism: An Unbridgeable Gap between AI and Humanity?

 **Tamar Cheishvili** Independent Scholar(USA)

the soul of the Netflix original sci fi cyberpunk thriller *Altered Carbon* at least in the first season. In the second season, half of the time he was either downplayed or found sulking in some corner. Even so, Poe Even so, just as Harold Bloom defined Poe's position in American literary canon 'absent presence', fans, similarly, still considered him to be the life and blood of [*Altered Carbon*](#). The AI hotel is known as The Raven.

In the books the show is based on, the AI hotel known as The Raven is actually called The Hendrix, named after none other than the legendary American musician Jimi Hendrix. This is something never mentioned in the show, due to legal reasons.

I would like to go beyond the textual surface and gain insight into the underlying currents and thoughts. The author uses AI as an instrument for achieving eternal life. We should not forget that AI is created by humans and not any other mechanism.

As a detective story, *Altered Carbon* explores classic cyberpunk themes: What does it mean to be human in the posthuman/cyberpunk world? Will technology set humans free or, on the contrary, will it deepen existential crisis/problems?

The Hotel has many different Poesque names – it is referred to as the Nevermore Hotel, the Raven Hotel ...Poe's Annabel Lee works at the hotel in the caregiver's capacity.

AI in my opinion does not have to be evil to destroy humanity – if AI has a goal and humanity just happens to come in the way, it will destroy humanity as a matter of course without even thinking about.

It is shown in the series that Poe named himself as such because he is deeply influenced by the writer's work, even quoting him at times. He also based his looks around the author, making him more of an eccentric oddity in the futuristic world.

The Raven Hotel and Poe both belong to the Earth. The weird thing is that in

SESSION 1

Paradigms of Creativity in the Age of Posthumanism: An Unbridgeable Gap between AI and Humanity?

 **Tamar Cheishvili** Independent Scholar(USA)

the second season, when Kovacs and Poe arrived on the Harlan's World (a planet light years away from Earth), the remains of the Raven Hotel are seen on it as well. This was probably done to cut costs and squeeze everything from the previous season into the new one while working with a smaller budget.

Poe has made study of humanity his biggest aspiration in his life. Apart from this un-AI like goal, he also loves gambling. In the [first season](#), when Poe met his other AI friends on the poker table, they taunt him for his previous outstanding debts which mostly comprise of AI union dues, which were still unpaid. Poe has been in the hotel business for a very long time and loves running a hotel, even if he may not be exactly the best at it.

It is the impractical things in the tumultuous hellscape of *Altered Carbon* (and in the real world, as well) that matter the most. A book, a name, chicken soup - anything made by human hands, really. As such, Poe believes that everything in life, even its darkest hours, should be savored. To that end, he is extremely loyal, affectionate, friendly, sincere, hospitable and resourceful when it comes to experiencing as much of humanity as he can.

Poe is extremely obsessed with humans and likes to dress, speak and behave like one. He even celebrates different festivals like Christmas and Halloween and knows some of the archaic slang from old English, like "gumshoes." Poe longs for respect and affection from the creators of his species: the human race. After a huge gap of five decades, when Kovacs checked into his hotel, Poe developed a deeply emotional bond with him.

As mentioned earlier, Poe's name and look were inspired by the author Edgar Allan Poe, but why is the hotel named The Raven? Simple: Because Poe is an Edgar Allan Poe fanboy. "The Raven" is one of the most read and loved narrative poems ever written by the real life author. To honor him, the AI Poe named his hotel after the famous story. Even to this day (and possibly in the future of *Altered Carbon*) the poem

SESSION 1

Paradigms of Creativity in the Age of Posthumanism: An Unbridgeable Gap between AI and Humanity?

 **Tamar Cheishvili** Independent Scholar(USA)

is still discussed for its uniquely embellished language and supernatural premise.

The AI hotels were originally built to give the guests a girlfriend experience. They were hardwired to want guests in the same way as humans long for intercourse or intimacy. The AIs like Poe were programmed to love and emotionally bond with their guests too, which made them smothering and very close to being stalkers. Additionally, the dangerous weapons installed in them during wars were never removed.

Naturally, people found this behavior to be beyond creepy and the presence of hidden weapons did not help, so they stopped checking in to such hotels. The Raven, along with many other AI hotels, suffered economically as a result.

At the end of the first season, in the year 2384, a furious Reileen sent her right hand Mister Leung to kill Poe and turn him to nanodust. The Ghostwalker struck him with some bullets and Poe died, his digital bits collapsing to the floor. He spun back up in the next season.

In the [second season](#), moments before Takeshi Kovacs sacrificed himself and burned in the Angel Fire, he was fighting with the Elder and it was quite clear what he was up to. It was at that point audiences saw Miss Dig trying to take Poe to their studio for his reboot, and Poe was noting something down in his notebook. After the reboot, Poe discovered a decryption key and a raw human DHF in his backup. Who was it? The showrunners have dropped ample hints to say that it is indeed Takeshi Kovacs who is about to come back.

Poe, patterned after the famous writer Edgar Allan, is the artificial intelligence owner of The Raven Hotel. Takeshi Kovacs is the first guest at his hotel in fifty years. Poe has been in the hotel business for a considerable length of time. He also seems to have a gambling problem – other AIs mention that he has debts from the poker game. The key point to make is that Poe shows an interest in humans, stating to other AIs that he finds them fascinating. He shows the knowledge of archaic human slang and views

SESSION 1

Paradigms of Creativity in the Age of Posthumanism: An Unbridgeable Gap between AI and Humanity?

 **Tamar Cheishvili** Independent Scholar(USA)

on private detectives. This knowledge makes him more human than machine and takes great pride in being the proprietor of The Raven Hotel where studying humanity has become his greatest aspiration.

Being utmost sincere and very hospitable in doing his job, he developed deep emotions for everyone who stayed in the hotel.

In the final analysis, what do the show-makers want to tell us? What is this: just a sci-fi piece for entertainment purposes or prediction of the possibility of humanization of AI? The question remains open to answers.

Kazuo Ishiguro's Dystopic Vision

In his latest novel (2021) and the first one after winning the Nobel Prize in 2017 titled “Klara and the Sun”, Kazuo Ishiguro uses AI to reveal the limits of our own. He demonstrates the ways in which AI influences the society. The action takes place in the United States in the distant future. The writer depicts the world in which people are categorized in a case system – those who still seem to be useful to society, and those who are sorted out because they don’t want to participate, or because they are no longer needed, as artificial beings have taken over many tasks and made many jobs superfluous.

Kazuo Ishiguro argues that It’s a realistic scenario. He is not one of those people who are terribly frightened of AI, he claims. However, he thinks that there are challenges that we have to face concerning the question of what happens to employment in our society. “The way we organize our societies, we all have jobs and that’s how we earn a living and feed ourselves and our families. This is going to be seriously challenged in a time when we can’t all have jobs anymore. Many important decisions in our lives will be made by AI”, he adds. Ishiguro’s dystopic imagination describes the world without memories, without rebellion but “Klara and the Sun” is not

SESSION 1

Paradigms of Creativity in the Age of Posthumanism: An Unbridgeable Gap between AI and Humanity?

 **Tamar Cheishvili** Independent Scholar(USA)

simply a dystopian fantasy. It evolves around existential questions: How do we remember and what? What makes us human? What does it mean to love – and what price are we willing to pay for love? The protagonist, Klara has no memories. She discovers the world through observation and only gradually understands what task she has been assigned. And she learns what friendship is, what love is. One day, Klara's job is done, and she, too, is sorted out. She doesn't even think about rebellion. The point Ishiguro seems to make is that humans should never give up and obey to the debilitating system of post-human/cyborg era. The rebellion and disobedience is the only way out.

Conclusion

We've read the books and watched the movies, and we know that soon, perhaps in our lifetimes, the robots will take over and humans will be the subservient ones. Our machines will outsmart us, and we'll be relegated to the role of second-class citizens. OK, not really.

To conclude with, the topic of AI vs human intelligence is often debated. Scholars and thinkers, when discussing it, immediately take sides. Most of them argue that future looks not promising. Misguided human goals and humans make mistakes that may cause them to inflict casualties. Others think that AI will help the humanity to achieve new ambitious goals. I am confident that AI will never be able to fully replace human mind/intelligence and the gap between them is ultimately bridgeable – creative mind and imagination is the privilege humans can always use to revive what they have lost and draw the inspiration and stimulus for development from the deepest sources of life itself, from the deepest fountainhead of our existence.

SESSION 1

Deep Medicine: Deep Empathy and Potential Healing in Klara and the Sun

 Pin-chia Feng National Taiwan University(Taiwan)

Abstract

The impact of artificial intelligence technology on healthcare is an urgent and evolving issue in the medical humanities. Contemplating the interaction between AI and healthcare, cardiologist Eric Topol stresses that with the three *Ds*—digitization, democratization, and deep learning—medical professionals have a framework to nurture a “human bond” which lies the root of medicine, and are thereby able to practice what he calls “deep medicine” (15). In her editor’s column to *PMLA*, literary scholar Wai Chee Dimock posts an all-important question regarding the interrelationship between AI and the humanities: “How can we create algorithms that would complement rather than replace human beings, help rather than destroy us?” (450). In Kazuo Ishiguro’s latest novel, *Klara and the Sun* (2021), the Nobel laureate envisions a world in which AI can complement, help, or even replace human beings, but will never intentionally wreck destruction. While in the novel elite workers are being “substituted” by AI, Ishiguro also creates an empathetic humanoid protagonist. Klara, the eponymous first-person narrator, is a solar-powered Artificial Friend designed to accompany a teenager to get ready for college education via home-schooling. The observant Klara, accessing and learning about the world around her through carefully crafted algorithms, is chosen by Josie, a fourteen-year-old dying of a mysterious disease after being “lifted,” or genetically re-engineered, to enhance her learning capacity. In the mode of a medieval romance or fable, the automata struggles to undergo a mythical quest and eventually commits self-sacrifice to rescue her owner/friend. In this paper, I first briefly review the main issues in AI and medical humanities. I then outline the variegated literary representations of AI in the science fiction genre. Finally, I want to use *Klara and the Sun* as an example, showing how, with deep empathy—the fourth *D*, so to speak—AI can work miracles. Specifically, I argue that Klara’s empathetic interaction with Josie and her endeavors in searching for a cure is performing a version of deep medicine. The robot, with her belief in love,

SESSION 1

Deep Medicine: Deep Empathy and Potential Healing in *Klara and the Sun*

 **Pin-chia Feng** National Taiwan University(Taiwan)

faith in hope, and unwavering loyalty, stands in contrast to the lonely, despairing, and fickle human characters, offering to nourish and heal—like the sun—the dystopic and divided fictional universe that can very well be our future.

Works Cited

Dimock, Wai Chee. “Editor’s Colum: AI and the Humanities.” *PMLA* 135.3 (2020): 449-54.

Ishiguro, Kazuo. *Klara and the Sun*. New York: Alfred A. Knopf, 2021.

Topol, Eric. *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*. New York: Basic Books, 2019.

SESSION 1

Deep Medicine: Deep Empathy and Potential Healing in Klara and the Sun

 **Pin-chia Feng** National Taiwan University(Taiwan)

The impact of artificial intelligence (AI) technology on healthcare is an urgent and evolving issue in the medical humanities. Contemplating the relationship between AI and healthcare, cardiologist Eric Topol, in his monograph *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*, notes that the three *Ds*—democratization, digitization and deep learning—give medical professionals a framework through which they can nurture the “human bond” which lies the root of medicine, allowing them to practice what he calls “deep medicine” (15). In her editor’s column in the *PMLA*, literary scholar Wai Chee Dimock posts an all-important question regarding the relationship between AI and the humanities: “How can we create algorithms that would complement rather than replace human beings, help rather than destroy us?” (450). In Kazuo Ishiguro’s latest novel, *Klara and the Sun* (2021), the Nobel laureate envisions a world in which AI can complement, help, or even replace human beings, but will never intentionally wreck destruction. While in the novel elite workers are being “substituted” by AI, Ishiguro also creates an empathetic humanoid protagonist. Klara, the eponymous first-person narrator, is a solar-powered AF, or Artificial Friend, designed to accompany a teenager to prepare for college education via home-schooling. The observant Klara, accessing and learning about the world around her through carefully crafted algorithms, is chosen by Josie, a fourteen-year-old dying of a mysterious disease after being “lifted,” or undergoing “genetic editing” (243), to enhance her learning capacity. In the mode of a medieval romance or fable, the automata undertakes a mythical quest and eventually commits self-sacrificesacrifices herself to rescue her owner/friend. In this paper, I want to use *Klara and the Sun* as an example, showing how, with deep empathy—the fourth *D*, so to speak—AI can work miracles. Specifically, I argue that Klara’s empathetic interaction with Josie and her family and her endeavors in searching for a cure performs a version of deep medicine. The android, with her belief in faith, hope, love and her unwavering loyalty, stands in contrast to the lonely, despairing, and fickle human characters, offering to nourish and heal—like the Sun that provides “special nourishment” to humans and AFs alike—the dystopic and divided fictional universe that can very well be our future.

SESSION 1

Deep Medicine: Deep Empathy and Potential Healing in *Klara and the Sun*

 **Pin-chia Feng** National Taiwan University(Taiwan)

To better contextualize my reading of *Klara and the Sun*, I would like to briefly review the main issues in AI and medical humanities and touch upon imagination and representations of AI in the science fiction genre. In her 2020 web article “Artificial Intelligence and Medical Humanities,” Kirsten Ostherr observes that the application of AI in healthcare provokes continuing debates, with medical practitioners divided between fear of displacement by the use of AI technology and optimistic embracement of technological innovations. Both sides, Ostherr points out, agree that AI poses “new challenges” to medicine, which mainly involve “a set of practical and ethical questions about the human contexts of AI in healthcare, including issues of data privacy and security, informed consent, risk and liability, professional expertise and training, explainability of results, flawed, biased, or incomplete datasets, and unequal access to the benefits of the technology.” Faced with these challenges, Ostherr concludes, the future of AI “in the clinical armamentarium is yet undecided.” It is understandable that clinicians and researchers tend to express their reservation towards the future of digital health, since it is in human nature to have misgivings about the unknown, especially when it comes to technological advancements that may very well go beyond human imagination. The idea of *technological singularity*, the point at which “artificial intelligence exceeds human intellectual capacity and control” (Zhang 3), can be most alarming. The fear and anxiety towards the vivification of Frankenstein’s creature in Mary Shelley’s *Frankenstein* (1818) is paradigmatic of such human responses. Or, as Mrs. Weston remarks on the voiceless robotic nanny Robbie in Isaac Asimov’s classic *I, Robot* (1950), “It has no soul, and no one knows what it may be thinking” (9). However, the trend of integrating AI technology into medicine is irreversible. In her own research, Ostherr has come across data demonstrating that “AI for healthcare has been the largest market for investment among all AI sectors” since 2018.

In a 2017 overview of the applications of AI in healthcare, Fei Jiang *et al* identify the multiple advantages of applying AI in medical practices:

AI can use sophisticated algorithms to ‘learn’ features from a large volume of healthcare data, and then use the obtained insights to assist clinical practice. It can also be equipped with

SESSION 1

Deep Medicine: Deep Empathy and Potential Healing in Klara and the Sun

 **Pin-chia Feng** National Taiwan University(Taiwan)

learning and self-correcting abilities to improve its accuracy based on feedback. An AI system can assist physicians by providing up-to-date medical information from journals, textbooks and clinical practices to inform proper patient care. In addition, an AI system can help to reduce diagnostic and therapeutic errors that are inevitable in the human clinical practice. Moreover, an AI system extracts useful information from a large patient population to assist making real-time inferences for health risk alert and health outcome prediction. (230)

These advantages motivate the progress of AI in medicine. Using stroke care as an example, Jiang and his colleagues demonstrate how AI techniques can be efficient and helpful in the areas of “early disease prediction and diagnosis, treatment, as well as outcome prediction and prognosis evaluation” (239).

In their findings, Jiang and his co-workers highlight AI technology’s capacity for digitization and deep learning. Eric Topol, a longtime advocate for the democratization of medicine, envisions a future in which patients are no longer subject to “medical paternalism” but have better access to their medical data, permitting them more involvement in their own care. (15). As mentioned, Topol believes that with the democratization, digitization, and deep learning of the medical profession, healthcare practitioners’s ability to treat their patients will be significantly augmented. Their availability for patients will also increase, leading to what he calls “deep medicine” (15). For Topol, AI advancement inspires, rather than produces distaste and revulsion. “As machine get smarter, humans will need to evolve along a different path from machines and become more humane,” Topol argues (290). This andro-specific evolution aims to uplift “our humanistic qualities, that which will always differentiate us from machines. Notably, human empathy is not something machines can truly simulate, despite ongoing efforts to design sociable robots or apps that promote empathy” (290). After all, as Abraham Verghese stresses it, “[e]mpathy is the backbone of the relationship with patients” and physicians need to first and foremost be “present” for their patients (1926).

Topol emphasizes the idea that, despite all their computational powers, machines cannot achieve complete simulation of human empathy; it is however

SESSION 1

Deep Medicine: Deep Empathy and Potential Healing in Klara and the Sun

 **Pin-chia Feng** National Taiwan University(Taiwan)

entirely possible in the world of science fiction. In fact, with the creation of the AF Klara, Kazuo Ishiguro has endowed his android protagonist with a greater capacity for empathy than his human characters. As Teresa Hefferman comments in her discussion of the 2001 film *A. I. Artificial Intelligence*, “Robots and androids in fiction and film have long operated as liminal figures that negotiate the boundaries of what it is to be human—the marginalized mechanical figure is often cast as more ‘human’ than humans, thus offering a corrective model to a world gone wrong” (12-13). What has gone wrong in the world of *Klara and the Sun*, as Klara explains using her android logic, are the “Pollution and inconsideration” of humans (163). Set in the near future of the United States, the human characters in *Klara and the Sun*—with the exception of Melania Housekeeper and Rick—appear mainly self-centered and preoccupied with their own life agendas. In the middle of the novel, Melania forcefully enlists Klara’s assistance to protect Josie by stating repeatedly “we same side” (174), which creates a special bond between the “disposable” underclass characters, an AI companion and an immigrant woman who speaks minimal English and disappears mysteriously at the end.¹ Even Rick, who has not been “lifted” and is therefore constantly being denigrated by society and yet voluntarily assists Klara to complete her ritual of heliolatry, becomes engrossed in his own interests when he reaches adulthood. Only Klara remains content and composed when she has to retire to a junk yard after she has outlived her usefulness. In a sense, Klara would be the optimal model that Hitoshi Ishiguro has been attempting to create with his android science.²

While Kazuo Ishiguro is not portraying a “posthuman techno-heaven,” as Daniel Dinello ironically puts it, neither does he express the technophobia predominant in science fiction in general. According to Dinello, science fiction is most keen on exposing the dark side of technological advancement:

¹*This bonding is nevertheless tinted with violence and distrust since Melania, who has been unfriendly and rude to Klara, also threatens to “dismantle” Klara should she fail to protect Josie properly (175).*

²*Combining the behavior-based robotics and the concept of distributed cognition, Hitoshi Ishiguro has initiated a top-down research approach to study human-robot interaction, what he terms android science, “an interdisciplinary framework between engineering and cognitive science” using human-like robots to obtain “essential knowledge on human-human interaction” (321). The main goals of android science are to produce humanlike robots with “the essential factors to replicate human likeness” (322).*

SESSION 1

Deep Medicine: Deep Empathy and Potential Healing in Klara and the Sun

 **Pin-chia Feng** National Taiwan University(Taiwan)

From the destructive robot-witch of *Metropolis* (1926) to the parasitic squid-machines of *The Matrix Revolutions* (2003), the technologized creatures of science fiction often seek to destroy or enslave humanity. Science fiction shows the transformation into the posthuman as the horrific harbinger of the long twilight and decline of the human species. In its obsession with mad scientists, rampaging robots, killer clones, cutthroat cyborgs, human-hating androids, satanic supercomputers, flesh-eating viruses, and genetically mutated monsters, science fiction expresses a technophobic fear of losing our human identity, our freedom, our emotions, our values, and our lives to machines. Like a virus, technology autonomously insinuates itself into human life and, to ensure its survival and dominance, malignantly manipulates the minds and behavior of humans. (1-2)

In contrast to the monstrous sci-fi creatures Dinello describes, Klara is harmless. Instead, she constantly faces diverse forms of violence in her encounters with human beings, including emotional estrangement from her owners, verbal abuse from unsympathetic teenagers and resentful adults, and potential physical violence—the scientist figure Capaldi proposes to open her up for further scientific discovery. Klara is represented as an emotionally sensitive humanoid, who experiences tends to subject to the feelings of fear and anxiety. Although there is no clear reference to it, it is likely that Ishiguro may have Asimov’s famous “three fundamental Rules of Robotics” in mind when he imagined Klara (44). Klara in fact appears to be “lifted” beyond the three rules that govern human-robot relations: not only is she incapable of hurting human beings, she also willingly sacrifices some part of her mechanism in an attempt to heal Josie. In the end, the gentle and selfless Klara, with her deep sense of empathy, seemingly embodies the very best of humanity. She is the robot composed of “algorithms that would complement rather than replace human beings, help rather than destroy us” as per request of Dimock.

Ishiguro’s efforts to make Klara a sacrificial figure can be problematic to some readers. In her review of the novel, for instance, Helen Shaw comments on the AF’s passive compliance to human wishes, leading to a dystopic future “in which automata simply keep doing what we ask them to do, placidly accepting the burden of each small, inconvenient task. *The novel takes us inside the mind of that constantly refreshing*

SESSION 1

Deep Medicine: Deep Empathy and Potential Healing in Klara and the Sun

 Pin-chia Feng National Taiwan University(Taiwan)

patience, where at first it's rather peaceful—until it's chilling” (emphasis original). Shaw ends her review by comparing Klara’s act of sacrifice to that of the nightingale in Oscar Wilde’s “The Nightingale and the Rose” and concludes with a sharp critique on human egocentrism.³ The combination of passive victimhood with ruthless human self-interest has long been an important theme in Ishiguro’s fiction. In *Never Let Me Go* (2005), his earlier speculative novel members of the clone community in Hailsham quietly accept their fate as designated organ donors. As Alex Preston points out in his review, *Klara and the Sun* “feels very much of a piece with *Never Let Me Go*, again exploring what it means to be not-quite-human, drawing its power from the darkest shadows of the uncanny valley.”⁴

Shaw correctly pinpoints the fable-like quality of the novel. Although the novel makes no specific religious reference, except for the solar-powered Klara’s heliocentric obsession, I would argue that, to a certain extent, for all its sci-fi veneer, *Klara and the Sun* can be read as a Christian fable.⁵ Ishiguro deliberately creates a naive non-human being to observe the human world and human interrelationships. Through Klara’s narrative voice, Ishiguro presents a fictional universe in which many of the characters’ names reveal their roles in life, almost allegorically like in a medieval morality play: Klara’s first instructor about the human world is Manager of the AF store; Josie’s family includes the Mother, Melania Housekeeper, and the mostly-absent Father. And, of course, there is the Sun, the god-like presence that Klara sincerely believes can revive the dead and assist people in love. Whereas Klara acts like an electronic version of Every(woman), the rest of the characters are sophisticated human beings motivated by their own desires and aspirations. Her simple devotion to Josie makes Klara at once an innocent non-human being who has to be initiated into the complexities of human world and a self-appointed healer of human diseases, both

³Shaw comments, “All our technological inventions are nightingales, programmed to destroy themselves and the natural world to satisfy some human’s passing whim. Klara shows us how gladly she lets herself be pierced to the heart. Ishiguro argues that if we allow her to do it, we will be the ones to feel the sting” (emphasis original).

⁴Specifically, Preston points out that “Klara’s voice has the same beguiling simplicity that we found in Kathy H in *Never Let Me Go*, the same mixture of intelligence and naivety.”

⁵Ian J. Battaglia refers to Klara’s sun-worship as “a sort of heliocentric religion.”

SESSION 1

Deep Medicine: Deep Empathy and Potential Healing in Klara and the Sun

 **Pin-chia Feng** National Taiwan University(Taiwan)

physical and psychological, with the support of not science but faith, hope, and love.

With the publication of *The Buried Giant* (2015), Ishiguro demonstrated his familiarity with Arthurian legends and the medieval quest motif. Klara's attempt to complete the healing mission is framed as both a quest and an epistemological journey: this innocent protagonist can only acquire needed knowledge and achieve her goal through the act of self-sacrifice, a concept that she has learned to adopt. In fact, at the beginning of the novel, Boy AF Rex accuses Klara of being "greedy," for wanting to take up as much of the Sun's nourishment as possible (4). The two trips to Mr. McBain's barn, the place she assumes the Sun visits "at last thing each evening" (159), are structured like religious pilgrimages; the way in which she ritualistically insists on being secretive and acting alone further intensifies the religious aura. Klara almost zealously commits herself to terminating the Cootings Machine—a machine for road construction that produces heavy smoke and sacrilegiously threatens the authority of the Sun. Based on her causal calculation, the successful completion of the task will naturally stop the Pollution and thereby please the Sun; He then will show "special kindness to Josie" (164). Only later does she find out that she has to pay dearly to complete this sacrificial ritual, much like the legendary heroes in their battles against evil monsters.

While every decision that Klara makes stems from algorithmic operations, it is her tireless devotion to the rescue mission that makes her advanced "artificial general consciousness (AGI)" remarkable (Koch 48).⁶As Preston notes, "Ishiguro has clearly thought hard about those elements of a nascent mechanical consciousness that would be more or less developed, about what faith would look like to an android mind, or love, or loyalty." Even though there is no mentioning of artificial or machine consciousness that would be more or less developed, Ishiguro has endowed Klara with

⁶In his Scientific American article, Christof Koch explains the two current approaches to artificial consciousness: the global neuronal workplace (GNW) theory, which believes that consciousness originates in some of the peculiar architectural features of the brain, and the integrated information theory (IIT), which stipulates that only a mechanism with intrinsic causal power is conscious; among the two approaches GNW is predominant and achieves elaborate representations in sci-fi (48-49).

SESSION 1

Deep Medicine: Deep Empathy and Potential Healing in Klara and the Sun

 **Pin-chia Feng** National Taiwan University(Taiwan)

somewhat intuitive but rich feelings—she is extremely sympathetic with Josie and the young love between Josie and Rick, for instance—most likely resulting from affective computing. The AF characters in the novel are apparently examples of “strong AI,” intelligent machines whose cognitive consciousness is projected as more or less equivalent to that of humans. Klara’s uniqueness, even among her fellow AFs, lies in her intense curiosity and sensitivity. She talks about her curiosity about the world outside the store when she and Rosa sit at the display window: “Unlike most AFs, unlike Rosa, I’d always longed to see more of the outside—and to see it in all its detail” (8). This is one example showing how Klara learns by observing. When the Mother, burdened by Josie’s failing health, expresses her envy for Klara since presumably, as a machine, she has no feelings, Klara calmly replies: “I believe I have many feelings. The more I observe, the more feelings become available to me” (98). Hence, in recommending Klara to the Mother, Manager specifically highlights her strong observational capacity: “Klara has so many unique qualities.... But if I have to emphasize just one, well, it would have to be her appetite for observing and learning. Her ability to absorb and blend everything she sees around her is quite amazing. As a result, she now has the most sophisticated understanding of any AF in this store, B3 not excepted” (43). With her deep learning capacity, Klara is an expert of observing and adopting to social environments with the help of her learning algorithms.

To showcase the depth of Klara’s consciousness, Ishiguro allows her to practice psychological projection. On her first pilgrimage to visit the setting Sun, Klara projects her sense of danger onto Rosa when her fear of not reaching the barn in time intensifies. She hears “the cries of animals in pain” and perceives a mental picture of Rosa “sitting on the ground somewhere outdoors, little pieces of metal scattered around her, as she reached out both hands to grasp one of her legs stretched out stiffly before her” (155). Later, when she imagines detecting the Sun’s displeasure in the barn, the image of Rosa reappears, “on the hard ground wearing an expression of pain, reaching forward to touch her outstretched leg” (163). After she believes she has succeeded in making a compact with the Sun, however, everything becomes friendly to her. This performance

SESSION 1

Deep Medicine: Deep Empathy and Potential Healing in Klara and the Sun

 **Pin-chia Feng** National Taiwan University(Taiwan)

of Rosa’s mental dismantling demonstrates Klara’s psychological sophistication, illustrating the extent to which she appears human. By displacing any threat of failure and danger onto a long lost friend, who in Klara’s observation is relatively insensitive to the happenings around her, Ishiguro makes it clear that Klara can deploy projection as a defense mechanism to manage unwanted feelings.

The extraordinary sophisticated AF system Ishiguro envisions may very well mark a potential future of AI application in healthcare. In the novel, the official function of an AF involves accompanying a “lifted” teenager through the period of social isolation before college. Klara is also expected to act as a nursemaid, watching over Josie’s health. Ishiguro nevertheless makes it clear that, as a robotic product, Klara has a certain innate deficiencies. For one thing, as a B2 AF her model is reportedly has problems with solar absorption (6). Neither does Klara have a sense of smell: only B3 model AFs “have been given limited smell” (138). Moreover, every outdoor trip becomes a high adventure for Klara, when she has to struggle to navigate around the uneven ground on her way to Mr. McBain’s barn to negotiate with the Sun for Josie’s recovery, for instance. Most strikingly, we repeatedly witness how her emotional responses to external reality affect Klara’s vision. Klara’s vision becomes partitioned into boxes or panels whenever she senses negative affects. After her last day in the window, her vision in the darkened store was divided into boxes through which she could detect Manager’s complex emotions—“kindness and sadness” as well as “anger and frustration” (27). Manager’s disappointment can be viewed as unfulfilled commercial interest because no AFs were bought that day. This is one of the first instances in which Klara’s emotional sensitivity, something that is supposedly uniquely human, is demonstrated. Another instance occurs at Josie’s interaction meeting, Klara’s view of the Open Plan, or the living room, “became divided into twenty-four boxes—arranged in two tiers” (71). And she instinctively applies colors to the boys who later bullied her: “There was an unpleasant tint on the three boxes containing the boys on the sofa—a sickly yellow—and an anxiety passed through my mind” (72).

SESSION 1

Deep Medicine: Deep Empathy and Potential Healing in Klara and the Sun

 Pin-chia Feng National Taiwan University(Taiwan)

For Helen Shaw, Klara’s way of visual processing presents “both a deeper kind of perceiving... and a more rudimentary machine vision: human emotion as CAPTCHA grid.” Yet I would argue that Klara’s visual reconfiguration becomes more than just being “overwhelmed when confronting something unfamiliar,” as Shaw claims. Rather, it is Ishiguro’s way of representing Klara as a sentient being. Hence Klara becomes “alarmed” when she sees a bull supposedly sending out “so many signals of anger and the wish to destroy” on her way to Morgan’s Falls (99) but feels that the sheep in the field are “filled with kindness—the exact opposite to the terrible bull from earlier” (106). While her affective responses may be limited or even wrongly configured, Klara is clearly highly sensitive and responds to her environment intuitively. Her cognitive limitations are consistent with the findings of Hitoshi Ishiguro’s android science, in which developing “humanlike perceptual abilities” becomes difficult outside of controlled environments (325).⁷ In these instances, Ishiguro presents the “high-intensity mimesis” that Seo-Young Chu attributes to science fiction (9).⁸

By granting Klara human-like sensitivity, Ishiguro demonstrates his vision for the future of AI technology. Still, as with other writers of speculative fiction, he remains critical of a technological future dominated by human ambitions and ruthlessness. Capaldi, reminiscent of Frankenstein, exemplifies the figure of the callous scientist. At first Capaldi, appears to be highly appreciative of AFs and their “intellectual powers” (197), as opposed to many of the human characters in the novel who either look down upon androids as merely robots or resent them for fear of being

⁷Hitoshi Ishiguro enumerates the process of creating humanlike perception in an android and its problems: “The android requires humanlike perceptual abilities in addition to humanlike appearance and movements. This problem has been tackled through computer vision and pattern recognition in controlled environments. The problem, however, becomes very difficult in other situations, such as crowded places; vision becomes unstable and audition and speech recognition are hampered by background noise. The distributed sensor systems of ubiquitous computing solve this problem. The idea is to recognize the environment and human activities by using many distributed cameras, microphones, infrared motion sensors, floor sensors and ID tag readers in the environment.” (325)

⁸For Chu, “A ‘work of science fiction’ is a text in which high-intensity mimesis predominates. A ‘work of realism’ is a text in which low-intensity mimesis predominates” (9); hence a sci-fi work is capable of representing reality better than a realistic novel.

SESSION 1

Deep Medicine: Deep Empathy and Potential Healing in Klara and the Sun

 **Pin-chia Feng** National Taiwan University(Taiwan)

“substituted” (99). Indeed, Josie’s engineer father lost his job to a machine Rick’s mother, Helen, most candidly articulates her confusion about Klara: “One never knows how to greet a guest like you. After all, are you a guest at all? Or do I treat you like a vacuum cleaner?” (143). In the climax of the novel, when Klara finally casts sight of the Josie android in the Purple Room, Capaldi’s laboratory (201), and understands that instead of painting a portrait of Josie, as the adults claim, Capaldi is making an automata in her image. Rather than training the Josie replica, Klara realizes that she is expected to “*become her...to inhabit that Josie*” (207 emphasis original). As Capaldi enthusiastically exclaims, “You are not being required simply to mimic Josie’s outward behavior. You’re being asked to continue her for Christie. And for everyone who loves Josie” (207). Here Ishiguro plays upon the theme of substitution with an ironic reversal—Klara is asked to replace Josie and become the child of the (now broken) family—while posing both philosophical and ethical questions regarding Capaldi’s seemingly logical reasoning.

Ishiguro allows Capaldi to present a sequence of arguments about the imminent success of the substitution. First he warmly praises Klara’s computational capacity, which allows her to “[access] quite comprehensively all of Josie’s impulses and desires” (207). He then claims that there is “[n]othing inside Josie that’s beyond Klaras of this world to continue,” promising that “[t]he second Josie won’t be a copy” (207). He advises Mother that she must relinquish the belief that “there’s something unreachable inside each of us,” (207), which Josie’s father Paul promptly refers to as “the human heart” (215). In this instance, Klara appears to accept Capaldi’s line of argument and feel positive about the completion of her assignment. She confidently responds to Paul:

“‘The heart you speak of,’ I said. ‘It might indeed be the hardest part of Josie to learn. It might be like a house with many rooms. Even so, a devoted AF, given time, could walk through each of those rooms, studying them carefully in turn, until they became like her own home.’” (216)

SESSION 1

Deep Medicine: Deep Empathy and Potential Healing in Klara and the Sun

 Pin-chia Feng National Taiwan University(Taiwan)

Klara's deft use of metaphysical conceits here again manifests her capacity for sophistication. She is persuaded by Capaldi's sophism because the argument seems completely logical. Nevertheless, when Klara unexpectedly reunites with Manager at the end of the novel, she confesses that even though she can imitate her young mistress's every movement, she can never achieve what people feel for Josie. According to Klara, "[t]here was *something* very special, but it wasn't inside Josie. It was inside those who loved her" (302). Interpersonal bonding via love, in Klara's retrospective reflection, is that special something that cannot be substituted or replaced.

Furthermore, Capaldi's grand plan touches upon an important ethical issue of the novel. Since Ishiguro deliberately portrays Klara as a humanoid with sophisticated consciousness, as a human-made machine clearly with advanced thinking power, even subjectivity,⁹ can Klara be regarded as a conscious being and thus entitled to certain rights? In fact, the interactions between AFs and their human families involve two of the AI ethics discussion threads identified by Daniel Zheng: AI-Human relations and robot rights (4).¹⁰ As Zheng points out, humans may "develop various forms of attachments" to their robotic companions, which "could lead to interesting and challenging relational and ethical issues" (4). Klara is asked to sacrifice her inward consciousness as well her outward appearance to "continue" Josie, an issue germane to *roboethics* or robotics ethics and prompts a series of questions: Can an android really replace a diseased child? Does a non-human being have a right to refuse the demand from her master? Only Josie's miraculous cure allows Klara to live to tell the tale.

The representation of maternal love in the novel becomes as chilling and ethically questionable as Capaldi's plan. In a sense the Mother appears as cold-hearted

⁹According to Koch "Whether machines can become sentient matters for ethical reasons. If computers experience life through their own senses, they cease to be purely a means to an end determined by their usefulness to us humans. They become an end unto themselves. Per GNW, they turn from mere objects into subjects—each exists as an "I"—with a point of view" (49).

¹⁰According to Zheng, currently there are six active AI ethics discussion threads. The list includes technological singularity and doomsday scenarios, impact of automation on economy and employment, legal ramifications and accountability, privacy considerations and human rights, Human-AI relations and robot rights (3-5).

SESSION 1

Deep Medicine: Deep Empathy and Potential Healing in Klara and the Sun

 **Pin-chia Feng** National Taiwan University(Taiwan)

as Capaldi when she dispassionately declares that what makes Klara different from the Josie replica is “just fabric”; she even tries to persuade Klara into accepting the arrangement, in the hope of winning over Rick’s affection (211). Ishiguro does not touch upon the ethical aspect of genetic engineering directly. Instead, the issue of “lifting” and its deadly consequences are engaged through hierarchization and emotional responses. Rick is humiliated at Josie’s interaction meeting because he is not “lifted.” He is also barred from higher education because he has not taken AGE (245), which most likely will significantly limit his advancement in life in the more conventional sense. When Josie’s mother accuses the unlifted Rick of playing for “low stakes” and winning something “small and mean” (277), she might actually be trying to justify to herself the decision to lift Josie, despite the fact that she has already lost an older daughter. Rick, however, sees through her selfish attempt to exonerate herself from the moral responsibility for subjecting Josie to a dangerous and potentially lethal biomedical procedure. He calmly communicates Josie’s message of love to the Mother. His moral decency becomes a foil to the Mother’s egoistic and possessive form of maternal love. Yet Ishiguro shows how the Mother is also saved by Josie’s recovery when she refuses to comply with Capaldi’s proposal to use Klara to open the “black boxes,” a thinly disguised euphemism of dismantling the AF, which supposedly can stop the backlash against AFs (294). Instead, she decides to grant the AF her deserving “slow fade” and takes the step to shield Klara from Capaldi (294). Her protectiveness in turn reminds Klara of a previous moment of intimacy with the Mother, which effectively builds an empathic rapport between the Mother and the AF and somewhat makes up for the Mother’s earlier acts of selfishness.

Ishiguro closes the novel by leaving Klara in a junkyard where she quietly puts things in order with her weakening cognitive powers. Her process of fading is represented as having “composite memories” in which memories from different temporalities overlapping with each other (298). “Memory and the accounting of memory, its burdens and its reconciliation,” Radhika Jones remarks, remains one of the recurrent themes in Ishiguro’s oeuvre. Like an elderly lady who still takes pride on her

SESSION 1

Deep Medicine: Deep Empathy and Potential Healing in *Klara and the Sun*

 **Pin-chia Feng** National Taiwan University(Taiwan)

mental abilities, the retired AF announces her final mission: “I have my memories to go through and place in the right order” (302). *Klara and the Sun*, narrated in the first-person and presented in past tense, is the result of that reordering. Ishiguro even grants Klara a reunion with Manager, her first mentor, to give her epistemological journey closure. When Klara relates to Manager in the Yard her observation about how Josie has been loved by people around her, not only is she responding to the Capaldis in the scientific world but also to one of the hardest philosophical questions about humanity. Throughout the narrative, Klara detects that people around her show signs of loneliness. As she tells the sickly Josie, “Perhaps all humans are lonely. At least potentially” (255). When she has finally combed through her memories and placed them in the right order, she discovers that love, more than loneliness, prevails. What Ishiguro shares with us, through the artificial consciousness of Klara, is that genuine interconnection and interconnectedness among loved ones are what make homo sapient special. Finally, in the fictional universe of *Klara and the Sun*, it is Klara the homo AI, the only character in the novel who sees through the depth of human loneliness and offers her deep empathy, serves as the believer and practitioner of deep medicine for the human heart.

SESSION 1

Deep Medicine: Deep Empathy and Potential Healing in Klara and the Sun

 **Pin-chia Feng** National Taiwan University(Taiwan)

Works Cited

- Asimov, Isaac. *I, Robot*. 1950. New York: Bantam, 2004. Print.
- Battaglia, Ian J. "The Difficult Balance of Text and Subtext in 'Klara and the Sun.'" *The Chicago Review of Book*. 8 March 2021. Web. 15 March 2021.
- Chu, Seo-young. *Do Metaphors Dream Electric Sheep: A Science-Fictional Theory of Representation*. Hawai'i : U of Hawai'i P, 2010. Print.
- Dimock, Wai Chee. "Editor's Colum: AI and the Humanities." *PMLA* 135.3 (2020): 449-54.
- Dinello, Daniel. *Technophobia: Science Fiction Visions of Posthuman Technology*. Austin: U of Texas P, 2005.
- Hefferman, Teresa. "A. I. Artificial Intelligence: Science Fiction and Fairy Tales." *English Studies in Africa* 61.1 (2018): 10-15.
- Ishiguro, Hitoshi. "Android Science: Conscious and Subconscious Recognition." *Connection Science* 18.4 (2006): 319-32.
- Ishiguro, Kazuo. *The Buried Giant*. London: Faber and Faber, 2015.
- . *Klara and the Sun*. New York: Alfred A. Knopf, 2021.
- . *Never Let Me Go*. London: Faber and Faber, 2005.
- Jiang, Fei, et al. "Artificial Intelligence in Healthcare: Past, Present and Future." *Stroke and Vascular Neurology* 2 (2017): 230-43.
- Jones, Radhika. "A Humanoid Who Cares For Humans, From the Mind of Kazuo Ishiguro." *The New York Times* 1 March 2021. Web. 15 March 2021.
- Koch, Christof. "Proust among the Machines." *Scientific American* 321.6 (2019): 46-49.
- Osther, Kirsten. "Artificial Intelligence and Medical Humanities." *Journal of Medical Humanities* 11 July 2020. Web. 25 March 2021.

SESSION 1

Deep Medicine: Deep Empathy and Potential Healing in *Klara and the Sun*

 **Pin-chia Feng** National Taiwan University(Taiwan)

Preston, Alex. “*Klara and the Sun* by Kazuo Ishiguro Review—Another Masterpiece.” *The Observer* 1 March 2021. Web. 15 March 2021.

Shaw, Helen. “In *Klara and the Sun*, Artificial Intelligence Meets Real Sacrifice.” *New York*. 15 March 2021. Web. 15 March 2021.

Topol, Eric. *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*. New York: Basic Books, 2019.

Verghese, Abraham. “The Importance of Being.” *Health Affairs* 35.10 (2016): 1924-927.

Zheng, Daniel. “AI Ethics: Science Fiction Meets Technological Reality.” *IEEE Intelligent Systems* 30.3 (2015): 2-5.

SESSION 1

Envisioning Artificial Intelligence and Simulated Affections in Ian McEwan's *Machines Like Me*

 **Chung-jen Chen** National Taiwan University(Taiwan)

Abstract

Ian McEwan's fifteenth novel, *Machines Like Me*, asks a provocative question: whether a machine can understand the human heart—or whether we are the ones who lack understanding. McEwan's question, a classic yet thought-provoking one of McEwan's particular style and thematic concern, addressed at a small scale test grounds of moral dilemma within the drama of everyday life. However, McEwan's intention, as in nearly all his works, aims at a much bigger scenario with a pervasive sense of urgency: in the creation of machines that can outsmart humans, we should not just explore the moral ramifications of AI but also the ethical ambiguity of humanity. While the artificial replicant approached intactness with access to all the wisdom of the internet and a fixed notion of right and wrong, what is at stake is not the righteous and a conscious existence of the machine but the incomprehensibility of human behavior and the controversy of decision-making. McEwan's latest novel is a continuation yet departure from previous sci-fi forerunners in addressing to the core issue of artificial intelligence: can machine think? And can machines think like us or even better than us? In his inquiry of the makings of the machine, McEwan asks an unorthodox question about what constitutes a person, and the morally complex and disturbing future of humans.

SESSION 1

Envisioning Artificial Intelligence and Simulated Affections in Ian McEwan's *Machines Like Me*

 Chung-jen Chen National Taiwan University(Taiwan)

Ian McEwan's latest novel, *Machines Like Me*, a story about domestic conflict between a man, a woman and an android, is published with a right question at the right time. In our time when stories and movies about androids and their close and even intimate interactions with human beings seem to be ordinary and no longer surprising merely as a subject matter. As far back in 1968, when Philip K. Dick published his science fiction *Do Androids Dream of Electric Sheep?* and later in Ridley Scott's film adaptation, *Bladerunner* in 1982, we have been shocked, with great pleasure, by the dystopian idea that men and androids live in a symbiosis that often ended up in revolt and destruction. Indeed, topic of human beings and androids are nearly a ubiquitous topic for science fiction and movies.

It is almost certain that McEwan's latest novel is not considered as a science fiction in every sense, and obviously the writer himself has no intention of being so, either. As Anil Menon notes, McEwan's faces a dilemma in writing a "science-fiction novel" not to be mistaken as a science fiction (2019). The choice of subject matter and its further development has less interest about the fate of the robots nor the prospects of the robotic future, but more to do with the essential questions that define and challenge human beings. In this thread, it is out of topic to discuss whether McEwan's elaboration about robot ethics is a new and original one or not. McEwan has not intention to write a science fiction. He just happens to write a story about human beings in which robots and artificial intelligence turn out to be an indispensable part. Even so, McEwan's question is rightly placed under the rather explicit contextual reference to Isaac Asimov's laws of robotics. Admittedly, McEwan's attempt to understand the meanings of human beings in which robots become an essential part is rendered under the anxiety of influence. The question that really concerns us as a reader is, while audience and readers are already imbued with knowledge of artificial intelligence and hallmarks of sci-fi settings of human-robot conflicts, why do we need McEwan's novel to tell us the robotic future ahead of us?

So, the only question that matters is, if the idea of men and android living together is no longer a new idea, how does McEwan make his novel different from his

SESSION 1

Envisioning Artificial Intelligence and Simulated Affections in Ian McEwan's *Machines Like Me*

 Chung-jen Chen National Taiwan University(Taiwan)

forerunners? Why do we need another book about men and android when we have already had plenty, or maybe just too many of them? With McEwan's achievement, people may even ask: why do we need McEwan to tell us a story about human and android? Does McEwan say anything that other writers before him did not say? Is there anything different or special that McEwan say despite the popularity of this topic? This article argues, despite the lukewarm criticism, the story asks one important question, can a human mind ever find and maintain a meaningful relationship with an artificial consciousness? But, the way McEwan answers this question touches less upon the ethical debate about what to do and what not to do with robots, but more to do with two philosophical concerns: firstly, how robots are more human than human beings, and secondly, what are we? And that is the most important reason why we need a novel like McEwan's *Machines Like Me* at a time when robot ethics is no longer a plot in science fiction but a living reality in everyday life.

To continue this question, several follow-up questions must also be examined. First, why do we need to pay attention to our relationship with machines? Second, even if we notice their existence, why is it necessary that we listen to them instead of they listening to us? Third, and if androids are created for the benefits of human beings, what is wrong with regarding them as possessions? Are we not the master who creates their slaves and exploits them for sheer benefits and interests? And if human beings dominated the relationship, what's the urgency in listening to what our hands assembled? Why do we need to listen to our possessions, and probably, our obsessions?

The Age of Robots and Its Discontent

As prolific and as popular as McEwan, his publication always attracts attention from serious critics and faithful supporters. Especially when he remains highly productive after his turn from Ian "Macabre" to the humanist concerned McEwan as we know today around the late 1990s, almost all his books soon become the bestsellers around the English reading populations. After 40 years of successful career, McEwan remains the living legend of literary 4 phenomenon widely loved and highly acclaimed

SESSION 1

Envisioning Artificial Intelligence and Simulated Affections in Ian McEwan's *Machines Like Me*

👤 Chung-jen Chen National Taiwan University(Taiwan)

by readers seeking for the best quality of prose writing and a controversial story to be probing over. As Ron Charles describes, with his ability to “explore the most complex issues in the confines of the most ridiculous situations,” McEwan is “not only the most elegant writers alive, he is one of the most astute at crafting moral dilemmas within the drama of everyday life” (2019).

With the high expectations, however, McEwan's publication does not always arouse intense responses from the critics. His latest publication, *Machines Like Me*, receives lukewarm acceptance in comparison to his earlier works. Many of the reviews remains doubtful and uncertain in rightly categorize this novel, but one thing is for sure: McEwan's novels maintain the typical play of moral difficulty and ethical challenge. Recognizing McEwan's versatility animated by “a spirit of sinister and intelligent mischief,” Marcel Theroux commented that this novel is “morally complex and very disturbing” which is “closer in character to the dark and subversive McEwan of his earlier books” (2019). Far different from the typical category of science fiction, *Machines Like Me* belongs to “the genre of speculative fiction” with “morally ambiguous characters” and their convictions that “nothing is more human than moral inconsistency” (Theroux 2019).

Another challenge that troubles critics is the involvement with technology and the ethical complexity between machines and human beings. Literary imagination often runs ahead of scientific innovation. Scientific inventions, as often considered to be the embodiment of edging technology, turn out to be following the wildest fantasies of human imagination. Technology, as it turns out, realizes human ideas. And ideas are pioneering precedents of technological manufactures that are materialized in various innovative gadgets put up in the scientific labs or factory assembly lines. If that is the case, the first question to ask is, why do we need a McEwan's novel that talks about artificial intelligence and androids now?

McEwan, of course, is not a science fiction writer nor does he aim to become one. Despite the existence of androids, the story is a typical McEwan novel. Admittedly,

SESSION 1

Envisioning Artificial Intelligence and Simulated Affections in Ian McEwan's *Machines Like Me*

 Chung-jen Chen National Taiwan University(Taiwan)

McEwan's latest novel is build upon typical McEwanian elements: conflict and context. As both are his trademark skills contributing to his importance and influence, *Machines Like Me* is constructed upon a small-scale domestic conflict situated in front of the pale view of historical context. Similar to the pattern of his previous works, such as the haunting intruder and the disputed lovers in *Enduring Love*, the sociable lady with her courting rivals in, *Amsterdam*, the treacherous girl with her wronged sister and lover in *Atonement*, the newlyweds and their frustrating first night in *On Chesil Beach*, *Machines Like Me* is set in a love triangle between a man, his lover and an android. His concern with history leads to one essential question that he repeatedly examines: what happened to England that led to the way we are now? Or, simply put, what went wrong with us? Just to name some. *Saturday* is dimly foregrounded in front of the 9/11 terrorist attacks when airplanes flying low remind readers of the unforgettable and formidable memory. *On Chesil Beach* is set in the early 1960s before the sexual revolution when the pressure of sexual dysfunction would be dramatic enough to become devastating. *The Children Act* is skillfully situated in the background of the enactment of the Children Act in 1989 when issues of childlessness and disloyalty destroy mutual trust in a marriage. McEwan is most charming and fulfilling in his rendering of moral conflicts and ethical dilemma placed at some critical historical moments of modern England. Throughout his prolific and successful career, he is obsessed with one question and maybe the only meaningful question: what about now and what about us?

Admittedly, even *Machines Like Me* focuses on the conflicts caused by an android, this novel never wants to dig deep into the current debates of human-machine interactions and their ethical problems. Indeed, the woman in the story has sex with the android and that is the beginning of all conflicts. But the infidelity part with a machine does not add up the tension. The android could be perfectly replaced by a human male and the story line still moves on, only that the android is more intelligent, more understanding, and more capable than any human being. The point is, McEwan writes a story not to question the ethical problems about artificial intelligence. He only wants

SESSION 1

Envisioning Artificial Intelligence and Simulated Affections in Ian McEwan's *Machines Like Me*

👤 Chung-jen Chen National Taiwan University(Taiwan)

to ask questions about humanity, as he always does and does best. *Machines Like Me* is a fiction about androids but androids have never been his interest. McEwan only wants to ask the same question about humanity with the help of androids. That is to say, even in this story about artificial intelligence, McEwan is just being McEwan as we know him and as we love him.

While literature in robots ethics can be traced back for decades, it is only in recent years when the real possibility of intelligent robots outsmarting the best of us that the impacts on ethics and our way of life turn out to be a real issue. Android is an indispensable part of the novel, but the machines exist in far distance in the background of the story. It is true that human beings have sex with the android may be a problem. But the point is not about having sex with the machine, but having sex with someone else out of curiosity when you already have a stable relationship. Machines are not perfect. But they are never the problem. Human beings are. The problem is all about the imperfect but human emotions of betrayal, distrust, jealousy, hate, revenge, desire. McEwan sets up the scenario that highlights the conflict among his characters. But he has no intention to examine the prominent and morally problematic use of robots today. So, the only meaningful question that remains is, what is human nature and how will the advancement of artificial intelligence change humanity?

But before we move on, we should clarify the definition of robot and why McEwan chose to develop his story of domestic conflicts centering on a robot, and also why robot could be an excellent choice to heat up the ethical conflicts. According to George A. Bekey, robot is “*a machine, situated in the world, that senses, thinks, and acts*” (2012: 18). By thinking, it suggests that robots are capable of collecting and processing information and come to its evaluations and decisions. The ability to think endows the machines with a status of autonomy that are able to develop a relationship of “human-robot collaboration” (Bekey 2012: 21). Although robot is a new invention, it has made his debut performance on the stage of literature long time ago. Lattimore notes that robot as a subject for literary creation has a long tradition that could be traced back to Homer around 1190 BC when an intelligent robot dubbed as the “golden

SESSION 1

Envisioning Artificial Intelligence and Simulated Affections in Ian McEwan's *Machines Like Me*

👤 Chung-jen Chen National Taiwan University(Taiwan)

servant” was created by Hephaestus, the Greek god of technology (Lattimore 1961). Around 1495, Leonardo da Vinci conceived of a mechanical knight that would be called a robot by today’s standard (Hill 1984). Not to mention the upsurging popularity and eminence of robotic figures of contemporary writers (Issac Asimov, for example) and movie productions (*Blade Runner*, *Terminator* series and *Star Wars*, just to name a few). While robots are essential to many bestsellers and blockbusters, robotic figures are created for a wide variation of different purposes and importances, with a wide range of roles from the low-end scavengers of “Three-Ds,” the is, jobs that are dull, dirty and dangerous to the high end inventions of intelligence as the advanced models in *Blade Runner* (Lin 2012: 4). In some creations, robots live in harmony with human beings while in others they are the replacement of human beings.

With its long tradition of ubiquitous appearance, one thing is for sure: technology is materialization of collective imagination onto which the anxiety and anticipation of the society are projected. Robots, artificial intelligence, and many more inventions are the multiple embodiments of technology. Just as robots with artificial intelligence are the 7 byproducts of human imagination, and human imaginations are the reflections and projections of the needs of the society. In this thread, technology and inventions are an essential part of the society. They project what we have and what we don’t have. They are not only the anticipation but also the anxiety of the society. The interactions between society and technology are mutual: while the society is influenced by technology, technology is also inspired by the society. As Margaret A. Boden claims, “AI has challenged the ways in which we think about humanity—and its future” (2018: 25). As an important part of humanity, sexuality is changed and challenged profoundly. Those days that sexuality is a matter of humanity will be soon gone. It is now a matter of human beings and the machines.

The Robotic Ethics of Sex and Its Discontent

Technology is not the problem for McEwan; humanity is. It is not the advancement of technology that interest McEwan but humanity situated in the context

SESSION 1

Envisioning Artificial Intelligence and Simulated Affections in Ian McEwan's *Machines Like Me*

 Chung-jen Chen National Taiwan University(Taiwan)

of the latest technology. Or, as Jeff Giles analyzes, McEwan's new novel is "about what most literary novels are about: the godawful messiness of being human" (2019). To claim that *Machines Like Me* is set beyond the scope and concerns of science fiction is to affirm its capacity in providing an ethical battleground in the face of the futuristic machine culture. Even with the scenes of self-driven vehicles and accompanying robots, McEwan's latest story is nothing techie about the future, it is all too human. Instead of setting the story in the future of unknown technology, McEwan skillfully situated his story in the year 1982 and created a scenario of technology that "we may soon find ourselves nostalgic" (Lucas 2019). Such a writerly decision avoids the trouble of rendering technology in details and yields more space for the development of moral complexity. As Julian Lucas claims, McEwan tells a story in which "we worry not just what robots might do to us but what we might do to them, to say nothing of what they might do to us because of what we already do to one another" (2019).

Since discussion of McEwan's use of technology is an inquiry not to ask what robots can do to us, but ask what we do to the robots. So, what do we do to the robots? Although technology is not the concern of the novel, it does initiate all the problems and trigger all the confrontations that pave up the road for moral challenges. McEwan repeatedly set up man-machine conflicts to entail the ultimate question: what is the right thing to do if human beings manipulate and exploit artificial intelligence? The question is raised along with the 8 intensifying conflicts as the robot starts to seek for independence and intimacy. Conflicts arises when the android presents acknowledgeable supremacy over human beings and posits substantial challenges to the authority of the master. Starting from the abusive use of artificial intelligence for personal benefits, jealousy falling in love for the same woman, the violation of law forbidding doing damage to the masters, to the outburst of physical confrontation, the duration of man-machine interaction bears a record of conflicts and confrontations embodied in forms of exploitation and violence. Among all the disputes, one thing remains unreconcilable, unforgivable and impossible to resolve: what if the robot has sex with my girl friend who shows more interest in the experiment? What if the robot

SESSION 1

Envisioning Artificial Intelligence and Simulated Affections in Ian McEwan's *Machines Like Me*

 **Chung-jen Chen** National Taiwan University(Taiwan)

shows more potency in sex and affection in love than human beings? In the alignment between human beings and robots, physical intimacy and revelation of affection challenge the untouchable prohibition that is not to be transgressed.

So, what's wrong with having sex with robots? What's the moral concern and ethical complexity involved with robotic sex? As early as 2017, an AI engineer called Zheng Jiajia from China, married his robotic partner after years of frustration searching for a female human partner. The coalition was witnessed by the engineer's mother and friends at a simple, traditional ceremony (Haas 2017)¹. The human-size doll can only speak a limited amount of basic phrases but the engineer planned to "upgrade" his partner in the near future (Huang 2017). As Zheng designed and built his partner, he literally did everything for his love. A year later, a Japanese man called Akihiko Kondo surprised the world with his "marriage beyond dimensions" with a holographic virtual reality singer called Hatsune Miku (Jozuka 2018). Miku is a software voicebank personified as a sixteen-year-old girl with long twin tails in a super short miniskirt. Besides the multiple commodities of dolls and simulated concerts, Miku is projected in a flickering, holographic image in a bell jar larger than a home-size espresso machine called Gatebox. With basic artificial intelligence, Miku in Gatebox can manage simple greetings and provides an opportunity for the fans to live with their favorite character. Despite the controversy, Kondo was sure of his decision and claimed that he has fallen for Miku since his first encounter a decade ago. Miku in Gatebox is now available on Amazon Japan at the retail price at 1,100 USD².

¹Zheng was reportedly tired of the constant nagging from his family and pressure to get married, a typical family pressure in Chinese society for men and women about to pass marriageable ages. Zheng claimed to "date" his robot for two months before he decided to marry his robotic partner. The ceremony was held with traditional ceremony when the bride has to be covered by a red cloth over her head and the groom had to carry her through the door into the house. The marriage was not officially recognized by the Chinese government.

²The device Gatebox GTBX-100JP is a mass production model that provides service of accompany 2 to chat and to send text messages on the go, with more options of paid content and choices of monthly subscriptions. The company boasts that their product is "a bride character" that "heals and grows your hard working you". Available at http://https://www.amazon.co.jp/-/en/stores/Gatebox/page/2F57321E-AF72-4BAC-91BC-4B8DD7BD1A79?ref=_ast_bln. With the urging fandom, Miku is also available in software called Vocaloid where you can create lyrics and make her sing with easy settings on your computer. The latest edition of Miku on Vocaloid is now updated to the 5th edition in 2018. Engineered by Yamaha Corporation, Vocaloid was invented in 2007 and has been available in English since the second edition in 2004 and multiple languages since the third edition in 2014. Available at <https://vocaloid.fandom.com/wiki/VOCALOID5>

SESSION 1

Envisioning Artificial Intelligence and Simulated Affections in Ian McEwan's *Machines Like Me*

 Chung-jen Chen National Taiwan University(Taiwan)

As the world has seen claimed marriage by human beings with their robots, and we are expected to see more of similar cases, not less. An active online community of “iDollators” provided a ground of sharing experiences and moral support for those who favoring intimacy with artificial dolls over human. Independent filmmaker Melody Gilbert, in her 2018 documentary *Silicone Soul* examines the complicated decision of some people who chose to develop “full fledged relationship” with their synthetic companions. The film raised a challenging but meaningful question, though a self-evident one, “Are we losing our ability to get interact with others humans and just going for what’s easy and what’s in our control?” The answer to the documentary may be obvious, given that cases of revelation seem to indulge in the overwhelming sensation of reliance and attachment. But the facts that some people chose to rely on their selected silicone partner only point out that human affections can be treacherous and vulnerable that an inhuman partner who is even more reliable to devote your trust and affection.

Gilbert asks a brilliant question about man-made dominance in the relationship with dolls. It turns out that sex with humanoid objects is not the problem, response from the intimate relationship is. Having sex with your robot, though provoking, is not a new issue in the domain of ethical discussion of artificial intelligence. Human beings have a long history of having sex with inhuman objects. Sex toys and sex dolls can be traced back to thousands of years from the record of archaeological records (Danaher 2020: 404). The difference is, the object that human beings used to arouse erotic feelings with, has changed from lifeless and exotic objects, such as sex toys and sex dolls, to impersonated persona that are capable of basic interactions, namely sex robots that can talk and moan. Sex robot, being defined as the “artificial entity that is used for sexual pleasures,” possess the ability of “humanlike movements” with certain capacity of “artificial intelligence” (Danaher 2017: 11). Sex robots are humanoid form that resembles physical features and behavioral patterns of human beings. To have sex with robots is presumably to have sex with one of our kind, with similar patterns and prospective expectations. Sex toy is almost the real thing, just not yet quite.

SESSION 1

Envisioning Artificial Intelligence and Simulated Affections in Ian McEwan's *Machines Like Me*

 Chung-jen Chen National Taiwan University(Taiwan)

Despite the controversy, discussions of robotic ethics focuses on the appropriateness of having sex with robot, but not the probability of it. From the questions raised, we are provided with a clear message: “Why should we care?” instead of “Should we have sex with robots?” What matters here is not about what if but when. Having sex with the robot and its uprising ethical concerns are not a matter of possibility but one of feasibility since the accelerating advancement of robotics, or simply put, “the Robot Revolution” (Lin 2012: 3). And the issue is about to be more urgent as the emergence of the robotics industry is going to faster according to the Moore’s Law (computing speed will double every eighteen months or so) (Lin 2012: 3). The real problem at issue here, therefore, is not the sex part but the choice of object to have sex with. If the lifeless object is shaped not in human form, even in human parts, the interaction promises no shared bondage and sex is to be a one-way release. But if the lifelike object possesses basic ability to respond, an action of reciprocal engagement is activated, even if the communication is only to be recognized by the side of human beings. To have sex with robots is to entail that human beings have the potential to develop sexual behavior with objects of love, and most importantly, to develop emotional intimacy. To examine the challenging integration of technology with this “new form of sexual identity”, a term “digisexuality” is coined to examine the “rise of digisexuality” (McArthur and Twist 2017: 334). That is to say, sexuality is never the problem, sexuality with objects neither. The problem is the kind of sexuality with lifeless objects that we assumably, even falsely, regard them to be lively and meaningful. The problem is not sexuality. The problem is the type of identity and attachment that we assume to be taking effects upon us after the interaction.

In the broadest sense, we can have sex with lots of things, as Mark Migotti and Nicole Wyatt claim (25). It is not supposed to be a problem until a relationship with the sex object turns out to be more than a dildo or a vibrator or any unimaginable stuff. The problem is after all about agency in the act of sex. If it is an object involved in the act of manipulation, that should not be much of a problem. If there are two agents on the act of intercourse, the way we relate ourselves to our partner and we assume their

SESSION 1

Envisioning Artificial Intelligence and Simulated Affections in Ian McEwan's *Machines Like Me*

👤 Chung-jen Chen National Taiwan University(Taiwan)

significance would be a matter of importance. So the problem, admittedly, is not about having sex but about what to have sex with, namely, what or who our sexual partner is. A dildo is not a partner in this sense because it will never respond to us in any meaningful way. But a sex robot with artificial intelligence will be problematic since they can evaluate our feelings and decide to respond to us accordingly. In recent developments, applications are created to calculate our responses and to enhance the human-robot interaction. In a sense, human beings could be manipulated just as we assume that we are taking the dominance of sexual behavior. It turns out that the only question that matters to us is, What is it to have sex?

To ask the question what it is to have sex is to examine the relational ethics of sexuality. To assume that robots are legible sex partners is to acknowledge that digisexuality is a relevant issue and the questions of obligation and responsibility involved are real issues. To possess and process certain type of sexuality is first of all to recognize sexuality with robots to be an admitted category. With categorization of sexuality comes with the issue of identity in choosing what to identify with or what to identify against. Identity is a serious and urgent issue. Every inquiry of sexual identity is to pose the question to examine and to challenge the sense of belonging from intimacy. This endeavor includes attempts to make sense of and to challenge who we are, not just from the perspective of our desire, but from the angles of our partners. What we are is regulated not merely by what we think we are but also what others think of us. Such inquiries initiate a question of partnership in digisexuality, here in the case of McEwan's novel, a partnership with the robots.

Our Possession, Our Obsession

Being widely acknowledged as the “master of the domestic quarrel,” McEwan's stories are regularly “intensified by the introduction of a third party” (Lucas 2019). In *Machines Like Me*, Adam the android is the third party, with the closest analogue to “the monster of Mary Shelley's *Frankenstein*” (Lucas 2019). There's always an intruder in every McEwan's story. For the first time in McEwan's latest story, machine

SESSION 1

Envisioning Artificial Intelligence and Simulated Affections in Ian McEwan's *Machines Like Me*

 **Chung-jen Chen** National Taiwan University(Taiwan)

is presumably the third party. However, technology or machine as the intruder has been a fertile field for literary and cinematic creations from Isaac Asimov to the Star Wars. And the issue of moral relationships and tensions between human and robot has remained complex and intense ever since. Different from Asimov and the Star Wars where technology is recognized as the intruder, McEwan's novel explore the probability of co-existence of technology and humanity.

The answer is, McEwan does not render robots as machines. He makes his machines more human than human beings and his humans more detached than machines. The best part of McEwan's strategy is, he made his robots the perfect embodiment of all human capacities onto which all human desires and fantasies are perfectly projected. In a way, McEwan's rendering of robots are retouched with humanity that make them more human than human beings. They are so perfectly human that they are so different from human beings. In a way, machines are like us, but they are so different from us. They are too perfect to become any one of us. To take a step further to continue this assertion, if we examine the depth of human hearts with the help of our machinery creation, as Jeff Giles challenges, "could we really expect it to like what it sees?" (2019).

Starting from the gaze of the outsider, it provides the most critical perspectives looking back at the human characters trapped in the middle of moral complexities. Following Ian Patterson's comments, McEwan has little interest in pondering over the threats proposed by the outsiders, he "has generally preferred to show them at work from the outside" (2019). As much as humanity is inextricable from technics, McEwan's renders a world where humanity is reciprocally constitutive to the robots. Humanity and the machine are not placed in dichotomous confrontations with digital computation; they are closely intertwined and mutually indebted. In a sense, McEwan created an alternative cosmology where technology has been an indispensable part of their everyday life. Man cannot exist without technics, and the emergence of technology would be incomplete without the implicate alignment with humanity. Looking back at us from the point of view from the robot, we may have a more critical

SESSION 1

Envisioning Artificial Intelligence and Simulated Affections in Ian McEwan's *Machines Like Me*

 Chung-jen Chen National Taiwan University(Taiwan)

look at what we are and what we are not.

I am indebted to Bernard Stiegler in my discussion of humanoid machines, mostly from his *Technics and Time*. My discussion of terminology, such as android, machine, robot, humanoid, computer, is conducted loosely and interchangeably under Stiegler's notion of technics. Admittedly, Stiegler's coinage of technics (*la technique*) is an obsolete English word that encompasses "techniques, technology and the objects produced by these means" and thus the term allows a wide capacity in accommodating "the objects of pre-modern craft, pre-industrial and industrial techniques, and modern machine-powered technology" (Lewis 2013: 53). McEwan's robot presents the encompassing qualities of future technology. Just as McEwan's robots resembles human shapes and physical mobility of human beings, they are the embodiment of the keynote concepts of objects produced by technique and technology. 13 McEwan's humanoid robots are extravagant high-end commodities of the leading technology of artificial intelligence, not to mention material science, mechanical and power engineering, designed and manufactured by Alan Turing at McEwan's twist of alternate universe. With the exceeding capacities nearly in all areas, Turing's robots can easily process information necessary for the purpose of different ambitions, such as concert pianist and stock market investor. McEwan's robots also present qualities that encompass both humanity and technology. At its efficiency in collecting and evaluating information, Turing's robots can calculate all possible scenarios and make decisions decisively; at the capacity in advanced computing, they can predict and simulate human emotions and respond accordingly. In a sense, with the proficiency of artificial intelligence, Turing's robots know more than humans do, act faster than humans respond, and act more humanly than humans.

Stiegler aims at a critical reflection of current development of technics, in examining how we are surrounded by contemporary technics in our everyday life and in explicating how we become addicted to technological consumerism. Stiegler argues that Western democracies are plagued by "a crisis of the spirit" and "symbolic misery" consisting in "apathy, disaffection and social collapse" to which "we have become

SESSION 1

Envisioning Artificial Intelligence and Simulated Affections in Ian McEwan's *Machines Like Me*

 **Chung-jen Chen** National Taiwan University(Taiwan)

addicted” and “on which we depend for the construction of both society and ourselves” (Howells and Moore 2013: 2). Stiegler focused particularly on the collective addiction to “mass media” and “the technologies of consumerism” that profoundly altered the relationship between technology and the spiritual vitality of humanity (Howells and Moore 2013: 2).

Martin Heidegger famously described “the essence of modern technology” as being “itself nothing technological” (1993: 324-5). Heidegger indicates an understanding of what technology is cannot be deduced from an analysis of technology itself, as technology is not all about gadgets and technical artifacts but more to do with the systematic indebtedness to the makings of the society. For Heidegger, technology is significant not because it tells us about the cutting edge advancement of human knowledge, but because of what it tells us about ourselves: our desires and our wants. Similarly, Stiegler argues that human consciousness, that is already technical, is made possible by technics. That is to say, we create time through technics and we invent ourselves a future through the inheritance of acquired experiences. Both Heidegger and Stiegler point out that technology, accordingly, provides an ontological way of knowing our beings in the world.

In this sense, we are nothing but the internalization of our prosthetic existence (Stiegler 1998: 152). Human beings and the environment do not pre-exist one another, but are rather entwined in an ongoing, ontogenetic relationship of mutual constitutions. Out of the mutual integration of the social and technical milieu, human consciousness is regarded to be neither purely interiority (from the inside) nor pure exteriority (from the outside), but the result of consecutive differentiation and integration that lead to an intricate symbiosis. In this sense, Stiegler, deeply indebted to Heidegger, came to the conclusion that humanity is invented through technics. Admittedly, for Stiegler, humanity is not a condition to be transgressed and overcome, but rather to be recognized in its prosthetic fragility that needs to be taken care of.

SESSION 1

Envisioning Artificial Intelligence and Simulated Affections in Ian McEwan's *Machines Like Me*

👤 Chung-jen Chen National Taiwan University(Taiwan)

Technology gives rise to an understanding of who we are along with the relevance and significance of our existence to the surrounding environment. Humanity never exists without technology. Meanwhile, as technology facilitates humanity to invent themselves, it avails in its ability to reinvent itself through multiple accessible alternatives. What makes human life is the constructive openness to be recoded by technics. It is by adopting technics that we anticipate changes. Humanity, in this respect, is not a dubiously privileged biological entity to be passed over on the way to become something better.

And that leads me to the last question that I want to ask, probably the most urgent and most possessive one, is, are machines more human than human beings? It may be astonishing but it is highly persuasive as Constance Grady remarks, “the robot is the most human character” (2019). Heller Mcalpin complains that its human characters are only “interesting mainly in the context of their fraught, shifting relationships with [the machine]” (2019). Or as Alex Preston acknowledges, *Machines Like Me* is a novel not about science but “a novel about the power of novels” in which human imperfections are celebrated because that’s “the flaws that make us human” (2019). Similarly, Rabeea Saleem notes that the novel highlights “the ethical dilemma with unnerving pathos” in which “tables are turned and machines exhibit more humanly than humans” (2019). Or, as Jeff Giles examines, how come anything “that destroys innocence and so wantonly could reasonably be called human” (2019). The underlying presumption is: mechanical perfection is divine while human errors are common.

Conclusion

Our relation to the world has not been kept in balance since humans have deployed machines to extend the dominion of man over nature, not to reconcile to the inevitability of our vulnerability and our death. In consequence, technology imposes a restrictive mode of refraining life. Technology becomes an instrument for the immortalization of humanity while it sets about destroying the planet on which our

SESSION 1

Envisioning Artificial Intelligence and Simulated Affections in Ian McEwan's *Machines Like Me*

 **Chung-jen Chen** National Taiwan University(Taiwan)

survival depends. McEwan's latest novel raises some important questions to challenge the origins and ends of humanity, and to encourage and invite ongoing critical reflections on the nature of technology and the nature of humanity, not just for now, but also for the ages to come. We are not about to see a world of coexistence of technology and humanity; we are already in it. We are already in the symbiosis, we are born in the affiliation, and we are about to thrive in the mutual indebtedness and interconnection between technology and humanity. As both Stiegler and McEwan entail, humanity consists in its technicity and technics evolves in its humanity. The relation between humanity and technology is both continual and oppositional.

SESSION 1

Envisioning Artificial Intelligence and Simulated Affections in Ian McEwan's *Machines Like Me*

👤 Chung-jen Chen National Taiwan University(Taiwan)

Works Cited

- Bekey, George A. "Current Trends in Robotics: Technology and Ethics." *Robot Ethics: The Ethical and Social Implications of Robotics*. Eds. Patrick Lin, Keith Abney, and George A. Bekey. Cambridge, MA. and London: The MIT Press, 2012. 17-34.
- Boden, Margaret A. *Artificial Intelligence: A Very Short Introduction*. Oxford and New York: Oxford University Press, 2018.
- Charles, Ron. "The Robot Revolution is Coming: Will There be any Room for Us?" *The Washington Post*. April 17, 2019.
<https://www.washingtonpost.com/entertainment/books/the-robot-revolution-is-coming-will-there-be-any-room-left-for-us/2019/04/15/90a22352-5f8a-11e9-bfad-36a7eb36cb60_story.html>
- Danaher, John. "Should We Be Thinking about Robot Sex?" *Robot Sex: Social and Ethical Implications*. Cambridge, MA. and London: The MIT Press, 2017. 9-23.
—. "Sexuality." *The Oxford Handbook of Ethics of AI*. Eds. Markus D. Dubber, Frank Pasquale and Sunit Das. Oxford and New York: Oxford University Press, 2020. 403-420.
- Giles, Jeff. "Love, Sex and Robots Collude in a New Ian McEwan Novel." *The New York Times*, May 1, 2019.
<<https://www.nytimes.com/2019/05/01/books/review/ian-mcewanmachines-like-me.html>>
- Grady, Constance. "*Machines Like Me* Review: Ian McEwan's New Novel." *Vox*, April 23, 2019. <<https://www.vox.com/culture/2019/4/23/18512642/machines-like-me-review-ianmcewan>>

SESSION 1

Envisioning Artificial Intelligence and Simulated Affections in Ian McEwan's *Machines Like Me*

 Chung-jen Chen National Taiwan University(Taiwan)

- Haas, Benjamin. "Chinese man 'marries' robot he built himself." *The Guardian*, April 4, 2017. <<https://www.theguardian.com/world/2017/apr/04/chinese-man-marries-robot-built-himself>>
- Howells, Christina, and Gerald Moore. "Introduction: Philosophy—The Repression of Technics." *Stiegler and Technics*. Ed. Christina Howells and Gerald Moore. Edinburgh: Edinburgh University Press, 2013. 1-16.
- Huang, Kristin. "Chinese engineer 'marries' robot after failing to find a human wife." *South China Morning Post*, April 3, 2017. <<https://www.scmp.com/news/china>>
- Lewis, Michael. "Of a Mythical Philosophical Anthropology: The Transcendental and the Empirical in *Technics and Time*." *Stiegler and Technics*. Ed. Christina Howells and Gerald Moore. Edinburgh: Edinburgh University Press, 2013. 53-68.
- Lin, Patrick. "Introduction to Robot Ethics." *Robot Ethics: The Ethical and Social Implications of Robotics*. Eds. Patrick Lin, Keith Abney and George A. Bekey. Cambridge, MA. and London: The MIT Press, 2012. 3-16.
- Lucas, Julian. "Man, Woman, and Robot in Ian McEwan's New Novel." *The New Yorker*, April 15, 2019. <<https://www.newyorker.com/magazine/2019/04/22/man-woman-and-robot-in-ian-mcewans-new-novel>>
- McAlpin, Heller. "In McEwan's Latest, the 'Machine' Is Too Much Like You." *NPR*, April 23, 2019. <<https://www.npr.org/2019/04/23/714887136/in-mcewans-latest-the-machine-is-toomuch-like-you>>
- McArthur, Neil, and Markie L. C. Twist. "The Rise of Digisexuality: Therapeutic Challenges and Possibilities." *Sexual and Relational Therapy* 32. 3-4, Special Issue on Sex and Technology (2017): 334-344.

SESSION 1

Envisioning Artificial Intelligence and Simulated Affections in Ian McEwan's *Machines Like Me*

 Chung-jen Chen National Taiwan University(Taiwan)

Menon, Anil. "Machines Like Me by Ian McEwan reviewed by Anil Menon: Back to the Present." *The Hindu*, June 14, 2019.

<<https://www.thehindu.com/books/books-reviews/machines-like-me-by-ian-mcewan-reviewed-by-anil-menon-back-to-the-present/article27914719.ece>>

Patterson, Ian. "Sexy Robots." *London Review of Books*, Vol. 41.9, May 9, 2019.

<<https://www.lrb.co.uk/the-paper/v41/n09/ian-patterson/sexy-robots>>

Preston, Alex. "*Machines Like Me* Review-A Very Modern *Ménage à Trois*." *The Observer*, April 22, 2019.

<<https://www.theguardian.com/books/2019/apr/22/machines-like-us-ianmcewan-review>>

Saleem, Rabeea. "*Machines Like Me* by Ian McEwan Review: A Baggy and Jumbled Narrative." *The Irish Times*, April 20, 2019.





<<https://www.irishtimes.com/culture/books/machines-like-me-by-ian-mcewan-review-a-baggy-and-jumbled-narrative-1.3849775>>

Stiegler, Bernard. *Technics and Time, 1: The Fault of Epimetheus*. Trans. Richard Beardsworth and George Collins. Stanford, CA.: Stanford University Press, 1998 [1994].

Theroux, Marcel. "Machines Like Me by Ian McEwan Review: Intelligent Mischief." *The Guardian* April 11, 2019. <<https://>>

SESSION 2

Moderator  Hongkyu Ha Sookmyung Women's University

| AI and Culture | |
|-------------------|--|
| 14:00 ~ 14:25 KST | <p>A Narratological Inquiry of AI and Human Identity in the Posthuman era: Focusing on Netflix short animation <i>Zima Blue</i></p> <p> Yunkyong Kim Word Work Center(Korea)</p> |
| 14:25 ~ 14:50 KST | <p>Overcoming Animal Digitalis</p> <p> Alfonso Ballesteros Universidad Miguel Hernández(Spain)</p> |
| 14:50 ~ 15:15 KST | <p>Epistemological Nature of Mind Perception in Artificial Intelligence</p> <p> Hyungrae Noh Sunchon National University(Korea)</p> |
| 15:15 ~ 15:40 KST | <p>In the Age of AI, Do we need to Change the Definition of Art?</p> <p> Jaepark Rhee Chugye University for the Arts(Korea)</p> |
| 15:40 ~ 16:00 KST | Discussion |

SESSION 2

A narratological inquiry of AI and Human identity in the Posthuman era: Focusing on Netflix short animation <Zima Blue>

 Yunkyoung Kim Word Work Center(Korea)

Abstract

After declaring the arrival of the 4th industrial revolution at the 2016 World Economic Forum, the development of AI and cyborgs is accelerating. With the advancement of technology and the inflow of capital, the line between AI as “human-machine,” and a cyborg as “machine-human”, is gradually becoming more and more confusing between humans and non-humans. The starting point of AI and cyborgs is different between machines and humans, but the point of reaching them is heading towards the 'Posthuman'.

In this paper, through narratological analysis of <Zima blue>, the humanities of what kind of identity AI and humans will have and how they should coexist in the future posthuman era. This movie is a new AI film that deviated from the existing narrative grammar that deals with AI. Until now, SF films about posthumans such as AI and cyborgs mainly dealt with posthumans from a human point of view. In the narrative grammar of existing science fiction movies, AI or cyborgs wanted to be humans, attacked humans, or tried to dominate humans.

However, the recent Netflix short animation <Zima Blue>(2019) unfolds a new narrative that breaks the existing narrative grammar. Zima is a famous painter for his enormous murals, but living in seclusion. To reach the ultimate level, Zima remodels his body to become a cyborg. And he explores the meaning of life through the blue square called 'Zima Blue'. One day in the near future, he gathers people to show his last work. In front of people, he dismantles his cyborg body. The last thing left was Zima's original appearance, a simple machine as 'a swimming pool cleaner'. He was originally a cleaner for swimming pool tiles. It(he) was upgraded to AI by a human owner, and he(it) has become an AI. He was gradually a strong AI through deep learning. His appearance was also humanized, becoming an AI like a human(called “android”). Zima dreamed of becoming a human, surpassed humans and became a

SESSION 2

**A narratological inquiry of AI and Human identity in the Posthuman era:
Focusing on Netflix short animation <Zima Blue>**

👤 Yunkyong Kim Word Work Center(Korea)

posthuman as Homo Deus. But, he wants to be a pure machine again.

The narrative of <Zima Blue>, which depicts “a machine that wants to be a machine,” ask “Does AI want to be human?”, “Can humans control AI?” and the fundamental question "What is human being in the Posthuman era?"

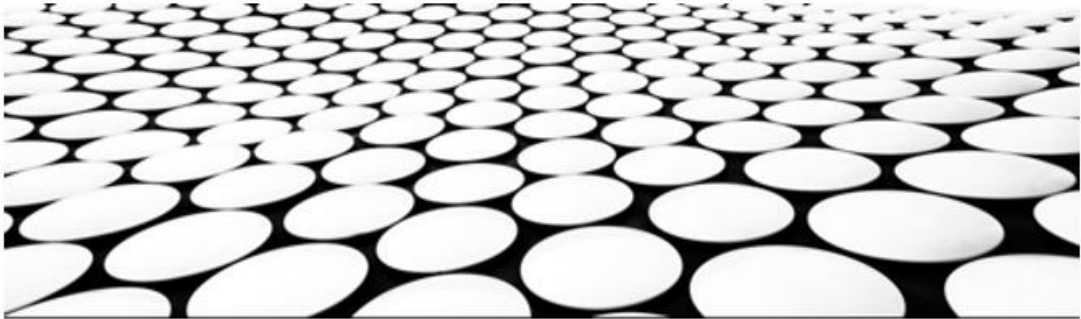
SESSION 2

A narratological inquiry of AI and Human identity in the Posthuman era:
Focusing on Netflix short animation <Zima Blue>

👤 Yunkyong Kim Word Work Center(Korea)

포스트휴먼 시대 AI와 인간의 정체성에 대한 서사학적 탐구
: 넷플렉스 단편애니메이션 <지마 블루>를 중심으로

김윤경 (말글살림연구소 책임연구원)



WHAT IS THE **POST-HUMAN**

- 현 인류(Human)보다 더 확장된 능력을 갖춘 존재
- 지식과 기술의 사용 등에서 인류보다 월등히 진화된 신(new)인류
- 기술을 활용한 진화를 통해 반영구적인 불멸을 이룰 것임
- 인간과 비인간, 인간과 기계의 경계가 모호해짐
예) AI, 외계인, 사이보그, 안드로이드, 인공생명 등

SESSION 2

A narratological inquiry of AI and Human identity in the Posthuman era:
Focusing on Netflix short animation <Zima Blue>

👤 Yunkyong Kim Word Work Center(Korea)

WHAT IS THE **POST-HUMANISM**

- 인간을 중심으로 여기는 인본주의(humanism)를 부정/초월하려는 사상
- 유전자 조작, 생명연장 기술 등을 통해 인간의 신체 및 정신을 향상시키려 함



WHAT IS THE **HUMAN** IN POST-HUMAN ERA

- **인간**은 **누구**인가? (X) => **인간**은 **무엇**인가? (O)
- **로봇**은 **무엇**인가? (X) => **로봇**은 **누구**인가? (O)

SESSION 2

A narratological inquiry of AI and Human identity in the Posthuman era: Focusing on Netflix short animation <Zima Blue>

👤 Yunkyong Kim Word Work Center(Korea)



- 어떤 작업이나 조작을 자동으로 할 수 있는 **기계(machine)장치**
- 외부의 제어 장치에 의해 조종되거나 내부에 제어장치가 내장됨
- 수행할 작업의 목적과 과업에 따라 특화된 기능과 외형을 지니고 있음(인간모습X)
- 인간의 편안함을 돕기 위해 개발되는 기계 또는 **약인공지능(ANI)**
- 기계, 기술, 공학적 관점

2021. 9. 27.



- 차페크의 희곡 <로섬의 만능로봇> (1920)
: '강제노동, 중노동'을 뜻하는 체코어 Robota에서 유래된 'Robot' 용어 최초 사용
=> **인간의 감정**을 가진 로봇의 반란에 대한 경고
- 아시모프 <아이,로봇>(1941)
: AI로봇의 이미지 및 '로봇 공학의 3원칙' 제시
=> 로봇 관련 **SF서사물(SF Narratives)**의 기준
- 인간과 가까운 능력과 외형(인간모습O)
- 인간의 정체성을 위협하는 **강인공지능(AGI)**
- 문학적, 인문학적, 철학적 관점

2021. 9. 27.

SESSION 2

A narratological inquiry of AI and Human identity in the Posthuman era:
Focusing on Netflix short animation <Zima Blue>

👤 Yunkyong Kim Word Work Center(Korea)

WHO THE ROBOT IN SF NARRATIVES

TYPE 1: 인간친화형



만화 <우주소년 아톰>의 Atom



영화 <A.I.>의 David



영화 <바이센테니얼 맨>의 Andrew

WHO THE ROBOT IN SF NARRATIVES

TYPE 2: 인간통제형



영화 <매트릭스>의 Agents



영화 <아이, 로봇>의 VIKI



마블코믹스 <캡틴 마블>의 Supreme Intelligence

SESSION 2

A narratological inquiry of AI and Human identity in the Posthuman era:
Focusing on Netflix short animation <Zima Blue>

 Yunkyong Kim Word Work Center(Korea)

WHO THE ROBOT IN SF NARRATIVES

TYPE 3: 인간파괴형



소설 <프랑켄슈타인>의
Monster



영화 <터미네이터>의
T-800



영화 <어벤져스2>의
Ultron

기존 SF 서사물에 나타난 로봇 유형의 공통점

- 인간과 유사한 자기인식이나 감정이 있음
- 인간을 지향하거나 통제하거나 거부하려 함
=> 인간에게 동정심이나 두려움을 자아냄
- **인간 중심적 로봇**: 인간의 관점에서 바라본 로봇
=> 인간을 위한, 인간에 의한, 인간의 로봇!



SESSION 2

A narratological inquiry of AI and Human identity in the Posthuman era:
 Focusing on Netflix short animation <Zima Blue>

 Yunkyong Kim Word Work Center(Korea)

TYPE 4: 탈-인간형

ZIMA BLUE

THE NTH STORY OF NETFLIX ANIMATION SERIES <LOVE, DEATH + ROBOTS>(2019)



SESSION 2

A narratological inquiry of AI and Human identity in the Posthuman era:
Focusing on Netflix short animation <Zima Blue>

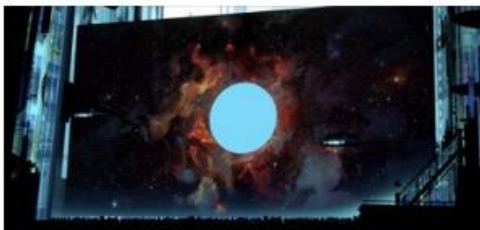
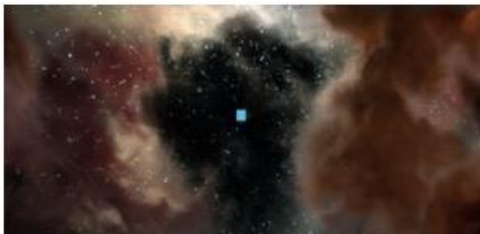
 Yunkyoungh Kim Word Work Center(Korea)

INTRO



https://www.youtube.com/watch?v=n_tY0U_KkV4

ZIMA's PAINTINGS: Zima Blue



SESSION 2

A narratological inquiry of AI and Human identity in the Posthuman era:
Focusing on Netflix short animation <Zima Blue>

👤 Yunkyong Kim Word Work Center(Korea)

IDENTITIES OF ZIMA: Robot > Android > Artist > ???



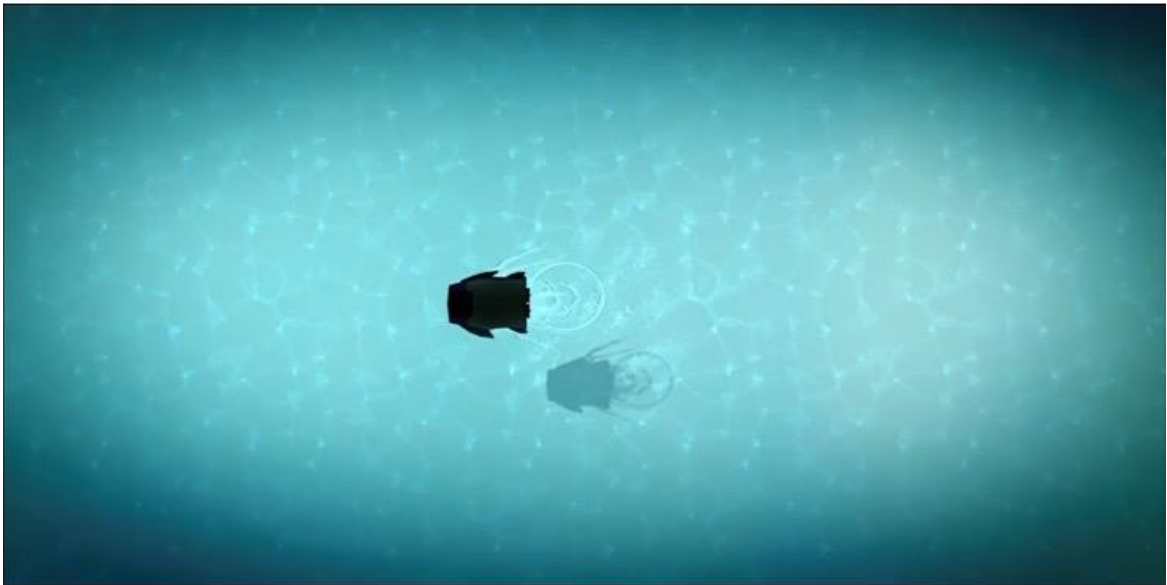
OUTRO



SESSION 2

A narratological inquiry of AI and Human identity in the Posthuman era:
Focusing on Netflix short animation <Zima Blue>

👤 Yunkyong Kim Word Work Center(Korea)



INTERESTS OF **POST-HUMAN ZIMA**

- 우주의 신비
- 인생의 의미
- 궁극의 도전
- 진리의 탐구



SESSION 2

A narratological inquiry of AI and Human identity in the Posthuman era:
 Focusing on Netflix short animation <Zima Blue>

 Yunkyong Kim Word Work Center(Korea)

IDENTITIES OF POST-HUMAN ZIMA

- 수영장 청소 로봇
- 인공지능 안드로이드
- 진리를 탐구하는 예술가
- 다시-수영장 청소 로봇



IDENTITIES OF POST-HUMAN ZIMA



하지만 지마 씨는 기계였던 게 아니라 몸 일부가 기계인 거잖아요



가끔은 나도 지금의 내가 뭔지 이해하기 어렵습니다

SESSION 2

A narratological inquiry of AI and Human identity in the Posthuman era:
 Focusing on Netflix short animation <Zima Blue>

 Yunkyong Kim Word Work Center(Korea)

IDENTITIES OF **POST-HUMAN** ZIMA



*MY RESEARCH FOR TRUTH
 IS FINISHED AT LAST.
I'M GOING HOME.*

- ZIMA BLUE -

SESSION 2

Overcoming Animal digitalis

 **Alfonso Ballesteros** Universidad Miguel Hernández(Spain)

Abstract

In conference 2020 I presented two theses on digitalisation: I considered it a de facto government, a digitocracy (1) that tends to animalise man, to turn man into an animal digitalis (2). I considered digitocracy as the global government of Silicon Valley engineers and their big tech-companies through two kinds of technologies, technologies of reputation and technologies of search. The first are about scoring and judging people in any field and for any purpose and the second are about giving people what they are supposed to want through addictive technology. I argued that those technologies as they are today are animalising man by making him addicted, emotional, transparent, stuck in a busy loneliness and stuck in the present. The rise of animal digitalis was my main concern.

As the power of digital tech companies increases in this time of global crisis, this concern becomes stronger. Two initial questions should be asked: Can we overcome both digitocracy and the rise of animal digitalis? And if so, how can we?

I think we can because digitalisation is, in part, a political decision. That is, it is not a historical or technological necessity. It is a matter of common good we can decide the path it takes. Digitalisation might be completely different and two 2020 events can make us optimistic, one coming from institutions, the other from society. The first is the US House of Representatives Antitrust Report on Big Tech (GAFA), the second is The Social Dilemma documentary film. I would say like in those cases, to overcome this problem both civil society and governments need to be part of the solution. I will focus in governmental solutions.

With respect to governments, solutions oriented towards the root of the problem are the best ones. That is, those that change the very structure of technology, that are not just about the content. Solutions such as requiring companies to design what is

SESSION 2

Overcoming Animal digitalis

 **Alfonso Ballesteros** Universidad Miguel Hernández(Spain)

called “slow technology” (L. Hällnas and J. Redström, 2001) or reflective by design technology. Changes like these will set the conditions for a human use of digital technology. They might transform animal digitalis into a real “homo”, an homo digitalis. A human being that uses technology in a reflective way and is opened to deep relations with others and with himself and is opened to past and future.

Setting these conditions might give back individual and societal goods. Other solutions that are applied not to the cause of the problem, but to some of the effects are not real solutions. Moreover, they can represent a threat to freedom and the rule of law as they are applied to contents. An example might be creating a commission that fights disinformation directly by deciding what is the truth. Requiring to design slow technology might be a long-term and democratic solution. Controlling content might give an excuse to authoritarian governments to take control of digitocracy for their own purposes. In this paper I will explore some of the possible solutions to overcome the animalisation of man in a governmental level and strengthening the social fabric and democratic institutions.

SESSION 2

Overcoming Animal digitalis

 **Alfonso Ballesteros** Universidad Miguel Hernández(Spain)

In conference 2020 I presented two theses on digitalisation: I considered it a de facto government, a digitocracy (1) that tends to animalise man, to turn man into an animal digitalis (2). I considered digitocracy as the global government of Silicon Valley engineers and their big tech-companies through two kinds of technologies, technologies of reputation and technologies of search. The first are about scoring and judging people in any field and for any purpose and the second are about giving people what they are supposed to want through addictive technology. I argued that those technologies as they are today are animalising man by making him addicted, emotional, transparent, stuck in a busy loneliness and stuck in the present. The rise of animal digitalis was my main concern.

As the power of digital tech companies increases in this time of global crisis, this concern becomes stronger. Two initial questions should be asked: Can we overcome both digitocracy and the rise of animal digitalis? And if so, how can we?

I think we can because digitalisation is, in part, a political decision. That is, it is not a historical or technological necessity. It is a matter of common good we can decide the path it takes. Digitalisation might be completely different and two 2020 events can make us optimistic, one coming from institutions, the other from society. The first is the US House of Representatives Antitrust Report on Big Tech (GAFA), the second is The Social Dilemma documentary film. I would say like in those cases, to overcome this problem both civil society and governments need to be part of the solution. I will focus in governmental solutions.

With respect to governments, solutions oriented towards the root of the problem are the best ones. That is, those that change the very structure of technology, that are not just about the content. Solutions such as requiring companies to design what is called “slow technology” (L. Hällnas and J. Redström, 2001) or reflective by design technology. Changes like these will set the conditions for a human use of digital technology. They might transform animal digitalis into a real “homo”, an homo

SESSION 2

Overcoming Animal digitalis

 **Alfonso Ballesteros** Universidad Miguel Hernández(Spain)

digitalis. A human being that uses technology in a reflective way and is opened to deep relations with others and with himself and is opened to past and future.

Setting these conditions might give back individual and societal goods. Other solutions that are applied not to the cause of the problem, but to some of the effects are not real solutions. Moreover, they can represent a threat to freedom and the rule of law as they are applied to contents. An example might be creating a commission that fights disinformation directly by deciding what is the truth. Requiring to design slow technology might be a long-term and democratic solution. Controlling content might give an excuse to authoritarian governments to take control of digitocracy for their own purposes. In this paper I will explore some of the possible solutions to overcome the animalisation of man in a governmental level and strengthening the social fabric and democratic institutions.

How do we recover the digital environment for the public good? When we reflect about what we can do to reach a post surveillance economy we need to think if there is any chance of real compliment of the law. Even more, if we will know if they will comply with the law. RGPD is largely ignored and has done little but scratching the surface. For this reason, an appeal to political solutions emerge such as expropriating, dividing or buying platforms. These solutions have the real danger of merely given platforms to governments. We know governments have been part of the problem as they want knowledge and control.

How to reclaim the digital environment for the common good?

Many people think that this economic order based on a post-privacy society is unacceptable, but how do we get out of it? The initial step is to talk about it, to make it an issue. The difficulty is that the subject is not immediately forgotten in our context our liquid news and liquid media. This difficulty makes a common response to the

SESSION 2

Overcoming Animal digitalis

 **Alfonso Ballesteros** Universidad Miguel Hernández(Spain)

problem by civil society even more complicated, because it is not enough to make it an issue, it is necessary to awaken a concerted response. Associations are one possibility: joining with others for the common purpose of limiting, for example, the use of Tik Tok among teenagers, Google services at our offices, etc., and all this, independently of the public authorities. Civil society responses can push or accompany an institutional response.

If we think about legislative policy, there are at least three problems that have to do with the effectiveness of solutions based on regulation. Prudential knowledge requires taking into account that it is often better to pursue certain objectives indirectly. Doing so directly can lead to undesirable results by ignoring the relationships of a problem with the rest of reality. Thus, certain prohibitions on capitalist surveillance can lead to capitalists carrying out with greater concealment what they used to do in plain sight. The second problem concerns the total asymmetry between the power of technologies and that of states: does the latter have sufficient power to enforce a right to limit the business of surveillance? The right is only effective if it can ultimately be enforced with the use of force. The third problem is the excess of current legislation, which is often not taken seriously even by the legislators themselves, openly contradicts itself, or is impossible for the jurist to understand.

With the exception of the above difficulties, proposals for a solution to the problem are offered, divided by themes: data, design and market.

Nine proposals for post-surveillance

DATA

I. Declare that the personal big data in the hands of surveillance capitalists were stolen (Bellver Capella, De Montalvo Jääskeläinen, 2021), taking advantage of the radical epistemological asymmetry between capitalists and users, with disregard for

SESSION 2

Overcoming Animal digitalis

👤 Alfonso Ballesteros Universidad Miguel Hernández(Spain)

individual conscience and in total illegality (Zuboff, 2020, 341). They therefore operate illegitimately in the power of corporations (and, where appropriate, institutions). This declaration has a symbolic value that cannot be underestimated.

II. Two very different options are possible after the declaration. The first is the total deletion of big data (Véliz, 2020: 143). The second is to declare big data a Common Heritage of Humanity (as nature is), and entrust it to some global authority (Bellver Capella, De Montalvo Jääskeläinen, 2021). The choice between the first or the second answer depends on whether we consider that the good that can be done with them outweighs the bad.

III. For data already legitimately processed after the surveillance economy, it is necessary to implement fiduciary duties. This means that anyone who wants to collect or store personal data acquires a duty of care towards the data subjects (Véliz, 2020: 136). This is the case, for example, in a relationship such as that of a doctor with a patient.

DESIGN

I. Prohibit anti-political algorithms. That is, those that break with the equality of opinions on the internet and favour precisely what is anti-political, that which destroys the polis: the tribal, the hateful, the sordid, etc. This avoids the hypocritical task of restricting the freedom of some after the fact, as happened with Donald Trump's Twitter account which was shut down without sufficient justification. Who has set up Twitter as an arbiter to give (or take away) speech to whomever they want? This is all the more galling when its business model is anti-political and does nothing more than add to an unfair inequality (the algorithmic one) another inequality (that of the president who is not considered worthy of having an opinion on Twitter). First the fire is structurally provoked and then it comes symbolically to put out one of the embers.

II. Ban design that is structurally manipulative. An example is the "Like" and

SESSION 2

Overcoming Animal digitalis

 **Alfonso Ballesteros** Universidad Miguel Hernández(Spain)

one-click retweeting. These design options added in 2009 are directly linked to the severe mental breakdown among teenagers a few years later. They simply seek behavioural surplus and therefore have turned communication into psychological satisfaction, what is properly communicated is dopamine. Therefore, these options spoil the possibility of a rich, non-psychologised and non-narcissistic digital communication, which was possible before their implementation.

MARKET

It is often said that the main problem is the business model. So much so that messaging apps that do not collect behavioural surplus in the form of big data (Signal and Telegram) are unprofitable and run on donations. In short: it seems that you can't make money if you act decently in the information society. So this is a structural problem, not an accidental one, but it can be fixed. Let's look at some more proposals.

I. Expropriate the platforms as critical structures of society. This would remove the problem of the business model, but the difficulties of this solution seem insurmountable, as there is a shift from a private domain - which disregards social consequences in search of effective profit - to a public domain that seeks certain social results. Can we expropriate without approaching the Chinese model?

II. Divide up the big tech companies. They are gigantic: Google, for example, has bought more than 200 companies in its lifetime. The option of splitting up seems to reduce the problem to gigantism without taking into account that it is not only their size that is harmful, but also their activity in the first place. So chopping up the big tech companies seems like a bad solution. This would lead to broadening competition and strengthening the surveillance market (Zuboff, 2020: 41).

III. Keeping Apple on track. Tim Cook's company could be a key player against surveillance capitalism if it continues on the path of privacy advocacy that has generally characterised it. This would set it against the other giants - Google,

SESSION 2

Overcoming Animal digitalis

 **Alfonso Ballesteros** Universidad Miguel Hernández(Spain)

Facebook, Microsoft and Amazon - which, in that order, have installed themselves in surveillance (Zuboff, 2020: 41). A different problem is the gigantism of the Apple company and the others, a gigantism that is bad for everyone and that is best avoided indirectly, for example, by protecting local SMEs. Who has thus prevented Amazon from taking over world trade?

IV. Not starting from the premise that innovation must be financed by the system. We pretend that history runs precisely through the channels that lead to making money: "Privacy is dead" said Zuckerberg. But our decisions count, and we should rather relate innovation to human development. Innovation" is often used to defend the enrichment of the few and the unemployment of the many. For example, is it not better to "finance people", to make recruitment cheaper and not to finance technology that could allow, among other things, surveillance (or mass unemployment)?

I conclude. The fight against this subtle form of domination is to combat the evils associated with it. These are all idolatry of unlimited movement and flow: narcissistic subjectivism, the technocratic spirit, financial capitalism or savage consumerism. In the face of these, we need frontiers that involve restoring otherness to reality. That is to say, to assume the limit of the unavailable: the otherness of others, of the past, of things and of law. All these are border phenomena, brakes, which we must recover in order for the world to regain the beautiful solidity it once had.

SESSION 2

Epistemological Nature of Mind Perception in Artificial Intelligence

 **Hyungrae Noh** Sunchon National University(Korea)

Abstract

People ascribe psychological attitudes, such as ‘belief’ or ‘desire,’ and attribute morality to AIs in response to such AIs’ human-like behaviors or appearances, namely mind perception in AI. One interpretation of mind perception in AI is that it is a sort of misperception in that people non-deliberatively and context-independently apply social rules and expectations to AIs. The aim of this paper is to show that mind perception in AI involves people’s deliberative use of psychological and/or social concepts in their explanations of behaviors of AIs. I meta-analyze recent relevant experiments regarding human-AI interactions. According to the meta-analysis, people’s ascription of psychological attitudes to AIs reflects their prior expectations regarding such AIs’ human-like performances and people’s attribution of morality to AIs reflects their context-dependent moral judgment regarding AIs’ moral status. In turn, mind perception in AI is not misperception.

SESSION 2

Epistemological Nature of Mind Perception in Artificial Intelligence

 **Hyungrae Noh** Sunchon National University(Korea)

Studies show that individuals perceive mind in artificial intelligence. Shank et al. (2019) report that participants showed strong emotional responses, such as unease or happiness, when AI systems like robots produce extraordinary outcomes, inhabit crucial social roles, and engage in human-like actions. In Heubner's (2010) experiment, majority of individuals ascribed beliefs and pain to a human-like AI cyborg. According to Ward et al. (2013), participants who observed intentional harm to a social robot attributed more mind (i.e., the ability to feel pain, the capacity for experience, the capacity for agency, and conscious awareness) to the robot. Given that people feel that home-cleaning robots have intentions, emotions, and unique characteristics (Sung et al. 2007), it seems that mind perception in AI is not restricted to AIs that are specifically designed to elicit social behavior. In short, people ascribe various types of mental states to many different forms of AIs and treat them as persons.

According to Turkle (2010), people perceive mind in AIs because the similarity between AI and human "pushes Darwinian buttons" that trigger certain psychological responses (Turkle 2010), as individuals perceive mind in geometric figures in Heider and Simmel's experiment (1944). Nass and Moon (2000) further claim that mind perception in AI is a sort of misperception resulting from people's non-deliberative misapplication of psychological or social concepts to non-human domain. To elaborate Nass and Moon's claim, we should understand Langer's (1992) distinction between 'mindfulness' and 'mindlessness.' The former refers to a state of conscious awareness in which the individual is implicitly aware of the context and content of information, and the latter refers to a state of mind characterized by an overreliance on categories and distinctions drawn in the past and in which the individual is context-dependent and, as such, is oblivious to novel aspects of the situation. A key feature of mindlessness is that from a single exposure to relevant information, an individual context-independently forms a cognitive commitment to the information and freezes its potential meaning. On the basis of meta-analysis of a series of relevant experiments, Nass and Moon (2000) argue that mind perception in AI is the result of people's mindlessly applying social rules and expectations to AIs. For instance, according to

SESSION 2

Epistemological Nature of Mind Perception in Artificial Intelligence

 **Hyungrae Noh** Sunchon National University(Korea)

Nass and Moon, when individuals exhibited polite responses to AIs, they were mindlessly exhibiting overlearned social behaviors without considering categorical differences between humans and AIs. In turn, Nass and Moon's argument implies that mind perception in AI is non-deliberative application of human-oriented concepts to the non-human domain.

The aim of this paper is to show that the epistemological nature of mind perception in AI cannot be taken as sheer mindless misperception. Taking mind perception in AI essentially as misperception is to assume that people apply psychological or social concepts to AIs in accord with their non-deliberative psychological responses to human-like behaviors or appearances of AIs. In short, mind perception in AI is misperception if an individual's use of psychological or social concepts in relevant situations fails to reflect the individual's background beliefs regarding AI. Nonetheless, in this paper, by giving meta-analysis of recent relevant studies, I show that people's perceiving mind in AIs reflects, and is partially determined by, their background beliefs regarding AI, thus mind perception in AI cannot be accounted merely as mindless use of psychological or social concepts.

I discuss two types of mind perception in AI, namely psychological attitude ascription and attribution of morality. The former relates to a case where people use psychological predicates, such as 'desire,' 'feel,' 'believe' to explain behaviors of AIs. According to Perez-Osorio et al. (2019), individuals with higher expectations regarding capacities of AI robots tended to explain iCub's (a human-like AI robot) behavior in terms of psychological predicates, while individuals with lower expectations maintained their mechanistic explanations of behavior. This experiment shows that people do not simply mindlessly ascribe psychological attitudes to AIs. In other words, individuals' use of psychological concepts in relevant situations is context-dependently determined, reflecting their background belief regarding AI in relation to such concepts.

SESSION 2

Epistemological Nature of Mind Perception in Artificial Intelligence

 **Hyungrae Noh** Sunchon National University(Korea)

The latter relates to, firstly, a case where people emotionally and morally response to behaviors of AIs, and secondly, cases where people treat AIs as moral agents. Consider the first case. The use of social concepts in the context of people's treating AIs as persons is observed in the experiments regarding moral status of AI. According to Shin (2021), individuals tended to show stronger emotional responses and attribute more moral patiency to social robots than economist robots despite that both robots were described as engaging in human-robot interactions and being harmed in vignettes. Shin's experiment suggests that people's use of social concepts, such as 'pity,' is context-dependently determined, reflecting their distinction between social and economist robots in relation to moral status of AI. In other words, Shin's experiment shows that, given the individuals' selective attribution of moral patiency to a particular type of robots, people consider categorical differences between humans and AIs to determine whether to treat AIs as moral patients.

Regarding the second cases, Hong's (2020) and Lima et al.'s (2021) experiments respectively suggest that people dissociate attribution of moral agency from psychological attitude ascription in their mind perception in AI. The dissociation implies that people attribute moral agency to AIs, not simply because of some automated emotional and moral responses, but because they make deliberative moral judgment regarding AIs. Hong (2020) reports that individuals blamed an AI driver for a car accident without ascribing psychological attitude to the AI driver in their explanation of the AI's behavior. When asked to compare between AI and Human drivers for car accidents respectively, individuals blamed the AI driver more despite that they caused the same kind of car accidents. Hong explains that individuals attributed responsibility to the AI driver more because they expect that AI should drive better than humans. According to Lima et al. (2021), individuals attributed punishment and responsibility to AIs but not psychological attitudes. Moreover, Lima et al. report that the individuals did not believe such punishment would achieve its retributive and deterrence functions (in contrast to punishing humans), indicating that their use of social concepts, such as 'punish,' reflects their deliberative application of such

SESSION 2

Epistemological Nature of Mind Perception in Artificial Intelligence

 **Hyungrae Noh** Sunchon National University(Korea)

concepts to the non-human domain. The fact that, in both experiments, individuals distinguished between attribution of moral agency and ascription of psychological attitudes implies that they deliberately considered what sort of mind they perceive from the given AIs. In addition, the fact that, in both experiments, individuals' background information regarding social concepts in relation to AIs (i.e., high expectation of AI drivers and special use of the concept 'punishment') was incorporated into their attribution of moral agency implies that they were aware of the category shift concerning the application of relevant social concepts. In sum, people do not mindlessly apply social rules to AIs when they treat AIs as persons, but people make deliberative moral judgments to determine whether to treat AIs as persons.

Mind perception in AI is not automated responses to human-like behaviors or appearances of AIs. Nor mind perception in AI can be fully accounted as mindless misperception (or misapplication of psychological or social concepts to the non-human domain). First, people seem to mindfully ascribe psychological attitudes to AIs given that their prior expectations of such AIs' human-like performance is embedded in the ascription of psychological attitudes (c.f., Perez-Osorio et al. 2019). Second, people seem to mindfully attribute morality to AIs considering their context-dependent moral judgment concerning the non-human domain (c.f., Shin 2021; Hong 2020; Lima et al. 2021). In turn, people's epistemic status of mind perception in AI involves mindful approach to their prior information about AI or their deliberative concerns regarding moral status of AI.

SESSION 2

Epistemological Nature of Mind Perception in Artificial Intelligence

 **Hyungrae Noh** Suncheon National University(Korea)

References

- Heubner, B., 2010. Commonsense concepts of phenomenal consciousness: Does anyone care about functional zombies? *Phenomenology and Cognitive sciences* 9: 133-155.
- Hider, F. and Simmel, M., 1944. An Experimental Study of Apparent Behavior. *The American Journal of Psychology* 57(2): 243-259.
- Hong, J. W., 2020. Why Is Artificial Intelligence Blamed More? Analysis of Faulting Artificial Intelligence for Self-Driving Car Accidents in Experimental Setting. *International Journal of Human-Computer Interaction* 36(18): 1768-1774.
- Langer, E. J., 1992. Matters of Mind: Mindfulness/Mindlessness in Perspective. *Consciousness and Cognition* 1: 289-305.
- Lima, G., Cha, M., Jeon, C., and Park, K. 2021. The Punishment Gap: The Infeasible Public Attribution of Punishment to AI and Robots. eprint arXiv: 2003.06507.
- Nass, C. and Moon, Y., 2000. Machines and Mindlessness: Social Responses to Computers. *Journal of Social Issues* 56(1): 81-103.
- Perez-Osorio, J. Marchesi, S., Ghiglino, D., Ince, M., and Wykowska, A., 2019. More Than You Expect: Priors Influence on the Adoption of Intentional Stance Towards Humanoid Robots. Salichs et al. (Eds.), *ICSR 2019, LNAI 11876*: 119-129.
- Shank D. B., Graves, C., Gott, A., et al., 2019. Feeling our way to machine minds: People's emotions when perceiving mind in artificial intelligence. *Computers in Human Behavior* 98: 256-266.
- Shin, H., 2021. Who Has a Mind?: Mind Perception and Moral Decision Toward Robots. *Journal of Social Science* 32: 195-213.

SESSION 2

Epistemological Nature of Mind Perception in Artificial Intelligence

 **Hyungrae Noh** Suncheon National University(Korea)

- Sung, J., Guo, L., Grinter, R. E., and Christensen, H. I., 2007. “My Roomba Is Rambo”: Intimate Home Appliances. J. Krumm et al. (Eds.), UbiComp 2007, LNCS 4717: 145–162.
- Tukle, S., 2011. Alone Together: Why We Expect More from Technology and Less from Each Other. New York: Basic Books.
- Ward, A. F., Olsen, A. S., and Wegner, D. M., 2013. The Harm-Made Mind: Observing Victimization Augments Attribution of Minds to Vegetative Patients, Robots, and the Dead. *Psychological Science* 24(8): 1437-1445.

SESSION 2

Epistemological Nature of Mind Perception in Artificial Intelligence

 **Hyungrae Noh** Suncheon National University(Korea)

Epistemological Nature of Mind Perception in AI

Hyungrae Noh
Department of Philosophy
Suncheon National University

Contents

1. Mind Perception in AI
 - Relevant experiments
 - Application of human-oriented concepts to the non-human domain
2. Mind Perception is Misperception?
 - Automated psychological responses
 - Non-deliberative application of relevant concepts
3. Objection 1: Psychological Attitude Ascription
4. Objection 2: Moral Patiency Attribution
5. Objection 3: Moral Agency Attribution
6. Conclusion

SESSION 2

Epistemological Nature of Mind Perception in Artificial Intelligence

 Hyungrae Noh Sunchon National University(Korea)

Mind Perception in AI

Relevant Experiments

- People perceive mind in AIs and treat AIs as persons.
 - ✓Shank et al. (2019)
 - Individuals' emotional responses ↔ AI's human-like behaviors
 - ✓Heubner (2010)
 - Individuals ascribe beliefs and pain to an AI character
 - ✓Ward et al. (2013)
 - Intentional harm ↔ Mind perception
 - ✓Sung et al. (2007)
 - Mind perception ↔ Social behaviors or human-like appearances



Mind Perception in AI

Application of human-oriented concepts to the non-human domain

- Applying psychological attitudes to explain AI's behaviors
 - ✓The AI home-cleaning robot *decides* to clean the room.
 - ✓The social robot *wants* to be recharged.
- Applying social concepts to AIs
 - ✓You *should not harm* Spot.
 - ✓I feel *pity* for Aibo.
 - ✓The AI driver is *responsible* for the car accident.
- On what epistemological ground, do we use such concepts in the relevant situations?



SESSION 2

Epistemological Nature of Mind Perception in Artificial Intelligence

👤 **Hyungrae Noh** Sunchon National University(Korea)

Mind Perception is Misperception?

Automated psychological responses

• Heider & Simmel Illusion (1944)

- Turkle (2011)
 - The similarity between AI and human “pushes Darwinian buttons” that trigger certain psychological responses.



- Application of psychological/social concepts to AI epistemologically grounded (at least partially) in some automated psychological processes.

Mind Perception is Misperception?

Non-deliberative application of relevant concepts

- Langer (1992)
 - Mindfulness: the individual is implicitly aware of the context and content of information.
 - Mindlessness: the individual overly rely on categories and distinctions drawn in the past and is oblivious to novel aspects of the information.
- Nass & Moon (2000)
 - Mind perception in AI is the result of people’s mindlessly applying social rules and expectations to AIs.
 - People non-deliberatively (i.e., without considering the categorical shift from the human-domain to the non-human domain) apply psychological/social concepts.

SESSION 2

Epistemological Nature of Mind Perception in Artificial Intelligence

 **Hyungrae Noh** Sunchon National University(Korea)

Mind Perception is Misperception?

Non-deliberative application of relevant concepts

- Nass & Moon (2000)
 - ✓ Politeness in human-human relationship
 - Speaker: "How do you like my new haircut?"
 - Listener: Positively biased polite answer (e.g., "It looks wonderful")
 - ✓ Politeness in human-computer relationship
 - Computer A: "How do you like my new haircut?"
 - Interview conducted:
 - by computer A: Significantly positively polite
 - by computer B: Low politeness
 - via paper-and-pencil questionnaire: Low politeness
 - In debrief, participants reported that:
 - They were unaware of their being polite to a computer.
 - They disagreed that computers have feelings or warrant polite treatment.

Mind Perception is Misperception?

A Summary

- People use psychological and/or social terms in the relevant situation because:
 - They (non-deliberatively or automatically) feel that AIs have mind.
 - They fail to consider the category shift from the human-domain to the non-human domain.
 - They would not do so if they were aware of the category shift.
- Therefore, mind perception in AI is a sort of misperception.
- Do we always fail to consider the category shift?

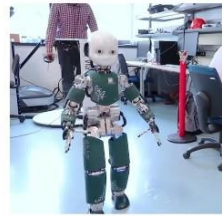
SESSION 2

Epistemological Nature of Mind Perception in Artificial Intelligence

👤 **Hyungrae Noh** Sunchon National University(Korea)

First Objection

Psychological Attitude Ascription



- Perez-Osorio et al. (2019)
 - Correlation between individuals' expectations regarding capacities of AI robots and the tendency to use psychological predicates in their explanations of iCub's behavior
 - ✓ High expectation → 'desire,' 'feel,' 'believe,' etc.
 - ✓ Low expectation → non-psychological mechanistic terms
- Participants, at least implicitly, consider the category shift.
- Participants did not mindlessly ascribe psychological attitudes to iCub.

Second Objection

Moral Patency Attribution



- Shin (2021)
 - Correlation between emotional responses and moral patency attribution
 - ✓ Harming an economist robot vs. Harming a social robot



- Participants' use of social concepts, such as 'pity,' reflects their distinction between different types of AIs.
- Participants did not mindlessly attribute moral patency to social robots.

SESSION 2

Epistemological Nature of Mind Perception in Artificial Intelligence

 **Hyungrae Noh** Sunchon National University(Korea)

Third Objection

Moral Agency Attribution

- Hong (2020)
 - Participants blamed an AI driver for a car accident more than a human driver without ascribing psychological attitudes to the AI driver.
- Lima et al. (2021)
 - Participants attributed punishment and responsibility to AIs but not psychological attitudes, where the use of relevant social concepts (e.g., punish) are specialized in non-human domain.
- People dissociate attribution of moral agency from psychological attitude ascription in their mind perception in AIs.
- People incorporate their background information regarding social concepts in relation to AIs into their attribution of moral agency to AIs.

Conclusion

- Mind perception in AI is misperception:
 - if it is grounded sheer in automated psychological processes,
 - and/or if the category shift from the human domain to the non-human domain is essentially mindless.
- Mind perception requires deliberative psychological processes.
 - Participants' selective attribution of moral agency
- People do consider the category shift.
 - Participants' expectations of the human-like capacities of AIs
 - Participants' distinction between social robots and economist robots
 - Participants' high expectation of driving skills of AI drivers

SESSION 2

In the age of AI, do we need to change the definition of art?

 **Jaepark Rhee** Chugye University for the Arts(Korea)

Abstract

As a result of being asked questions from the students through the <Art and Artificial Intelligence> class in university located in Seoul, students wondered whether people can feel emotions in works created by AI and whether works created by AI can be included in the realm of art. Furthermore, they also questioned what the standards and boundaries of art can be. These questions are questions that will be asked not only by students, but also by audiences who will experience art in the age of AI. Therefore, I conducted a picture evaluation experiment for 182 students to answer these questions. The experimental results of 17 pictures (9 human drawn pictures, 7 machine drawn pictures, and 1 chimpanzee drawn picture) are as follows. In the quantitative evaluation, the pictures drawn by AI scored highest rank in beauty and desire for owning (higher than Damien Hirst and Jackson Pollock). In the qualitative evaluation, various emotional expressions were detected in the drawings of AI, humans, and chimpanzees, confirming that emotions can be felt even in paintings of AI or chimpanzees other than human artists. The results of this experiment showed how difficult it is to demarcate what is a work of art and what is not, and how difficult it is to redefine art in the age of AI. Traditionally or customarily, art has been defined as what humans do, but if humans feel various emotions through the work of machines (or chimpanzees), then it seems that it is the time to consider the debate over whether the definition of art needs to be revised should begin.

SESSION 2

In the age of AI, do we need to change the definition of art?

👤 Jaepark Rhee Chugye University for the Arts(Korea)

인공지능의 시대, 예술의 정의를 바꿀 필요가 있을까?

: 인간, 침팬지, 인공지능 그림에 대한 비교 실험

In the age of AI, do we need to change the definition of art?

: Comparative experiment on human, chimpanzee, and AI drawings

추계예술대학교
영상비즈니스학과 강사
이재박

Chugye University for the Arts
Jaepark, Rhee

- Portrait Experiment 1

1



2



| | 1. Machine Painted | 2. Human Painted |
|------------------|--------------------|------------------|
| Beauty | 3.25 | 3.35 |
| Desire of Owning | 1.99 | 2.27 |
| Human or Machine | 77% Human | 82% Human |

SESSION 2

In the age of AI, do we need to change the definition of art?

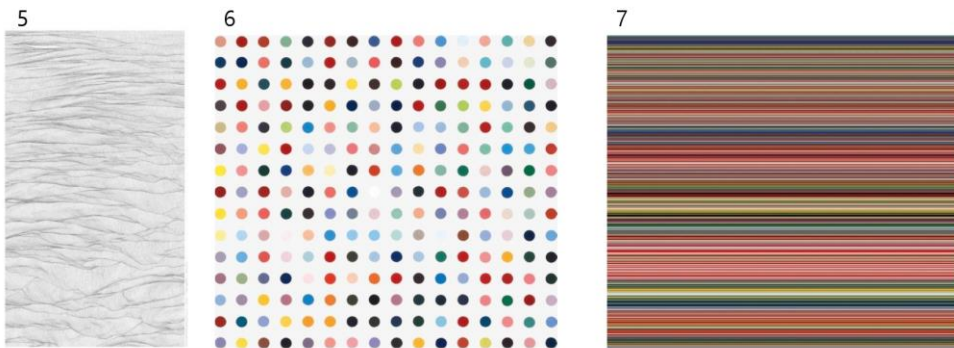
Jaepark Rhee Chugye University for the Arts(Korea)

• Portrait Experiment 2



| | 3. Human Painted | 4. Machine Painted |
|------------------|------------------|--------------------|
| Beauty | 3.27 | 2.07 |
| Desire of Owning | 2.82 | 1.57 |
| Human or Machine | 72% Human | 85% Machine |

• Abstract Painting Experiment 1



| | 5. Machine Painted | 6. Human Painted | 7. Human Painted |
|------------------|--------------------|------------------|------------------|
| Beauty | 2.87 | 2.25 | 2.12 |
| Desire of Owning | 3.00 | 2.04 | 1.77 |
| Human or Machine | 84% Human | 54% Machine | 67% Machine |

SESSION 2

In the age of AI, do we need to change the definition of art?

👤 Jaepark Rhee Chugye University for the Arts(Korea)

- Abstract Painting Experiment 2



| | 8. Human Painted | 9. Human Painted | 10. Chimp Painted | 11. Machine Painted |
|------------------|------------------|------------------|-------------------|---------------------|
| Beauty | 2.86 | 2.27 | 2.83 | 2.62 |
| Desire of Owning | 2.44 | 1.91 | 2.56 | 2.16 |
| Human or Machine | 88% Human | 73% Human | 82% Human | 67% Human |

- Skull Experiment



| | 12. Machine Painted | 13. Human Pictured |
|------------------|---------------------|--------------------|
| Beauty | 2.23 | 2.55 |
| Desire of Owning | 1.60 | 1.92 |
| Human or Machine | 60% Human | 72% Human |

SESSION 2

In the age of AI, do we need to change the definition of art?

Jaepark Rhee Chugye University for the Arts(Korea)

- Cat Experiment

14



15



| | 14. Machine Painted | 15. Human Painted |
|------------------|---------------------|-------------------|
| Beauty | 3.74 | 3.35 |
| Desire of Owning | 3.28 | 2.88 |
| Human or Machine | 68% Human | 83% Human |

- Portrait Experiment 3

16



17



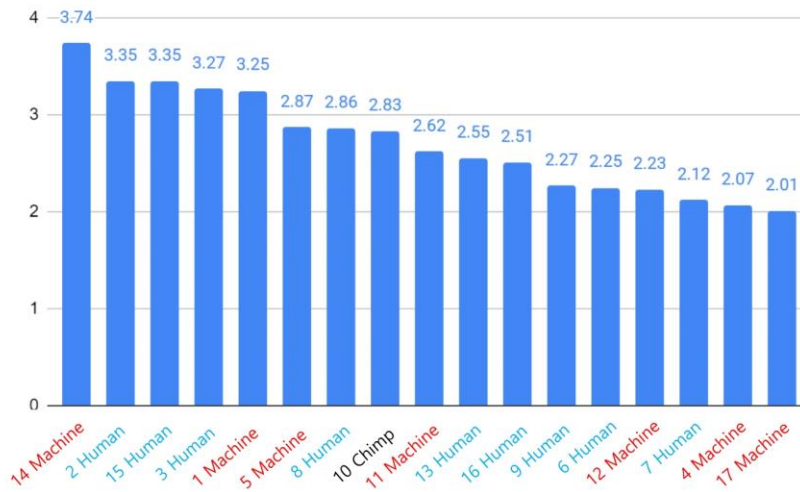
| | 16. Human Painted | 17. Machine Painted |
|------------------|-------------------|---------------------|
| Beauty | 2.51 | 2.01 |
| Desire of Owning | 1.97 | 1.59 |
| Human or Machine | 68% Human | 65% Machine |

SESSION 2

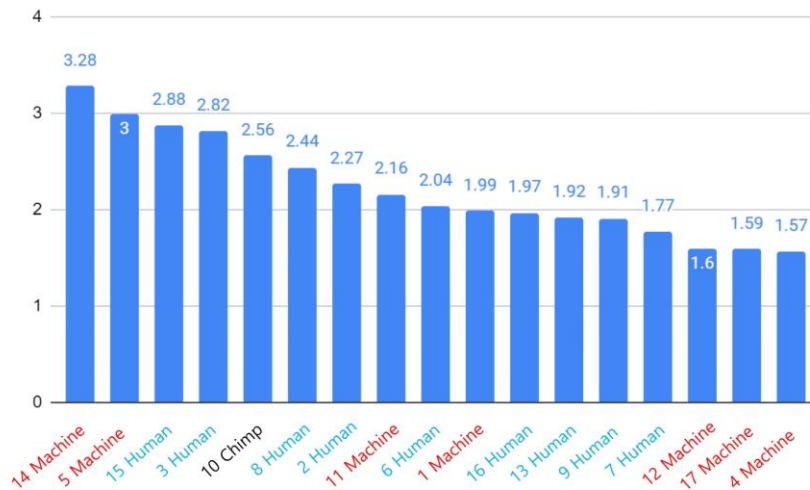
In the age of AI, do we need to change the definition of art?

👤 Jaepark Rhee Chugye University for the Arts(Korea)

Beauty



Desire of Owning



SESSION 2

In the age of AI, do we need to change the definition of art?

 **Jaepark Rhee** Chugye University for the Arts(Korea)

Students Opinion
On this Painting



English is Google Translated.

- 마주보고 있는 듯하다 마치 남자의 시선이 느껴질 정도다 원지 남자의 시선은 슬프기도 어딘가 의심이 서려있기도 의미심장한 시선을 지니고 있는 듯하다
- 기법이 뛰어난 예술이다. 최근에 그려진 것이라면 이전의 정형화된 자화상 혹은 주문 초상화 그리기 방식을 아주 잘 답습하고 있는 그림이라고 할 수 있겠다. 모작이 아니라 이 그림 자체로 독립적인 작품이라면, 제목을 잘 지어야 할 것이다. 이 때까지의 미술사의 맥락을 잘 따르고 있고 기법이 더할 나위 없이 뛰어난 그림이긴 하지만 현대 미술로서 값어치가 매겨지기 위해서는 맥락에 대한 이해를 바탕으로 기법, 자화상 혹은 초상화 개념을 작가가 어떻게 새롭게 재해석하고 있는지도 드러나야 한다. 하다못해 제목을 통해서라도 말이다.
- 그림에서 쓰인 색채가 대체로 어두워서 전체적인 그림의 분위기가 어둡다고 느꼈고, 또 정면을 응시하는 중년 남자의 얼굴을 자세히 보면 눈썹이 약간 찡그러져 있고 눈가가 붉은 것을 보아 슬픔을 참고 있는 듯한 느낌을 온었다.
- 명암표현이 신기함
- 미소가 은화해 보인다
- 그림에 눈을 뚫어저라 쳐다보면 마음속에 진실만 남게 된다.
- 힘들고 지친 삶을 보느라 같다
- 진짜 사람같으면서 부드럽다.
- The degree to which a man who appears to be observing is expected to show up close
- It is architecture. This is a very detailed explanation of the format of the setting for this. Moreover, even though this situation is not reached. They are doing well without being overdone, they are struggling with access to hardships that deserve compensation, and they are struggling with what they can afford. . Please let me know the title.
- Expectations for this have arrived, and we anticipate that 'forecasts' will come, and we will meet your expectations.
- Contrast is forbidden
- smile is gentle
- Continue to make up your mind in the future.
- help and live a weary life
- an ordinary person

Students Opinion
On this Painting



English is Google Translated.

- 기계가 한것같다
- 입시미술 수채화 같다
- 그냥 노말한 중세 미술작품 같다.
- 손의 색깔이 부자연스럽다
- 피부가 자연스럽지 않다. 남성을 그린 그림임에도, 여성을 그린 그림에서 찾을 수 있는 표현법이 차용된 것 같다. 손과 얼굴의 피부 톤이 화가가 그린 그림이라기에는 눈에 띄는 차이를 보인다.
- 1번류의 그림과 배경은 유사하지만 느낌이 상대적으로 많이 과하고 인위적으로 느껴질 정도로 은화하고 부드럽다.
- 앞 선 그림과 비슷한 감상이다. 르네상스 시대에 초상화 모작의 느낌이다.
- 1번과 같은 작가의 그림 같다. 우울하고 행복해보이지 않는다. 그리고 단순하게 그림기법을 통해서만 그려진 듯 하다.
- 앞의 그림과 다르게 확실히 옛날 그림 티가 매우 남는다. 앞의 그림은 배경 부분에서 어색한 티가 나는데 반해, 위 그림은 정말 사람이 그린 것 같습니다.
- 분위기와 남성의 표정이 오묘하며 역시 높은완성도를 보여주고 전문적인 유희지식을 바탕으로 그림을 그렸다.
- looks like a machine generated
- It's like watercolor painting
- It just looks like a plain medieval art piece.
- Hand color is unnatural
- The skin is not natural. Even though it is a picture of a man, the expression method found in a picture of a woman seems to have been borrowed. The skin tones of the hands and face show a noticeable difference for an artist's painting.
- Although the background is similar to that of the first picture, the feeling is relatively excessive and soft enough to be felt artificially.
- It is similar to the previous picture. It feels like a portrait replica in the Renaissance era.
- It looks like a painting by the same artist as number one. Depressed and doesn't look happy And it seems to be drawn only through simple drawing techniques.
- Unlike the previous picture, it definitely looks very old-fashioned. While the previous picture looks awkward in the background, the picture above looks like it was really drawn by a human.
- The atmosphere and the male expression are subtle, and they show a high degree of perfection, and the paintings were drawn based on professional oil painting knowledge.

SESSION 2

In the age of AI, do we need to change the definition of art?

Jaepark Rhee Chugye University for the Arts(Korea)

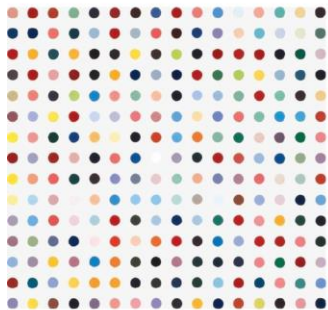
Students Opinion On this Painting



English is Google Translated.

- 여러 모습이 생각하는 바에 따라 다르게 보인다.
- 그림이 내가 좋아하는 스타일이어서 저는 보면서 괜찮다고 느꼈습니다
- 엄청 유명한 화가가 그린거같다
- 나도 그릴 수 있을 것 같다
- 광대한 바다가 흐르는 것 같은 느낌이 들고, 세밀하게 표현을 잘한 것 같다
- 철학적 내용이 담긴건가? 소다 작품에 대한 아름다움은 딱히 느껴지지 않음
- 이 그림 또한 의미가 많이 있을 거 같다. 전체적으로 잔잔한 분위기가 마음에 든다.
- 파동처럼 보이고 신선하다는 느낌이 강하다.
- 마음이 편해지는 그림이다.
- 물결이 흘러가는 모습같다
- 무슨 그림인지 잘 모르겠다
- 다양한 선들로 마치 자신의 상태를 공허함의 상태를 표현한것같다
- 왠지 한국적인 그림이라고 느껴졌는데, 선이 여러개 있는 것을 보니 우리나라의 산과 들 같기도 하고, 바다의 겹겹이 치는 파도 같기도 했다. 마음이 편안해진다.
- 최근에 저런 스타일의 간결하지만 세부적인 것을 강조하는 화풍이 유행해서 봤던 것과 비슷하고 깔끔하지만 예술성이 느껴져서 집안에 걸어 놓아도 이쁠것 같다는 생각이 들었습니다.
- Different appearances look different depending on what you think.
- The drawing is the style I like, so I felt okay looking at it.
- It looks like it was painted by a very famous artist.
- I think I can draw too
- It feels like the vast sea is flowing, and it seems to be well expressed in detail.
- I wonder if it contains philosophical content. I don't feel the beauty of the work.
- This picture also has a lot of meaning. I like the calm atmosphere overall.
- It looks like a wave and has a strong feeling of freshness.
- It is a picture that makes you feel at ease.
- It looks like the water is flowing
- I don't know what picture it is
- It is as if he expressed his state of emptiness with various lines.
- It felt like a Korean painting for some reason, but seeing that there were several lines, it looked like the mountains and fields of Korea, and like the waves crashing over the sea. my mind is at ease
- It's similar to what I've seen because of the recent trend of a simple but detailed style of painting like that, and I thought it would be pretty if I hung it in the house because it was neat and artistic.

Students Opinion On this Painting



English is Google Translated.

- 시력 테스트 같다
- 정신사납다
- 눈아프다
- 어지럽다
- 별 감흥 없다
- 점자판 같다.
- 환공포증
- 색맹인 아닌지 구별이 가능하다.
- 현란하다
- 동그라미가 너무 많아서 어지럽다
- 솔직히 예술성은 잘 모르겠다
- 감각한 느낌이 든다
- 의미를 잘 모르겠다
- 예술은 무엇을 향해 가는가
- 나도 그릴 수 있을 것 같다
- 개인적으로 이런 그림들은 작품같다고 느껴지지 않기 때문에 좋아하지 않는다
- 무언가를 의도한지 모르겠다. 점의 배치가 그저 시력테스트를 떠오르게 한다
- 그냥 색채 구별만 같다
- 점이 색깔이 있구나
- 표현의 자유는 알겠으나 남독이 가지 않는다
- 다양한 색의 팔레트가 같은 모양으로 모여 서로 다르면서도 통일감을 주어 다채롭다.
- 색을 볼때 머리가 맑아지는 것 같다
- 색상때문에 그림이 통통 튀기는 하나 역시 표현하고자 하는 바를 모르겠다
- 이게 예술이라는게 머리로는 이해가 가는데 가슴으로는 이해가 안됨
- 뭔가숨을 개념을 전달하기엔 표현이 부족함
- 솔직히 이런미술은 좀 싫다
- like an eye test
- crazy
- it hurts my eyes
- dizzy
- no inspiration
- It's like a braille board.
- cyclophobia
- You can tell if you are colorblind or not
- dazzle
- There are too many circles, so it's dizzy
- To be honest, I don't know much about art.
- I feel cute
- I don't know the meaning
- What is art going towards?
- I think I can draw too.
- Personally, I don't like these paintings because they don't feel like works of art
- I don't know if something was intended. The placement of the dots just reminds me of an eye test
- It's just like a color separator
- the dots are colored
- I understand freedom of expression, but I do not understand
- Palettes of various colors are gathered in the same shape to give a sense of unity while being different from each other.
- When I see the color, my head seems clear
- Because of the color, the picture bounces around, but I still don't know what I want to express.
- My head understands that this is art, but my heart doesn't
- i want something Lack of expression to convey the concept
- To be honest, I don't like this kind of art.

SESSION 2

In the age of AI, do we need to change the definition of art?

Jaepark Rhee Chugye University for the Arts(Korea)

Students Opinion
On this Painting



English is Google Translated.

- 모니터 깨진거같다
- 선넘었다
- 눈아프다
- 어지럽다
- 계속 응시하게 된다
- 착시효과를 주는 것 같다.
- 다채롭다
- 잘 모르겠다 티비가 고장난 느낌이다
- 상상력을 자극하진 않는 것 같다.
- 보자마자 벽지같은 생각이 들었다
- 평한 색상과 얇은 선들의 조합이 눈이 아프기도 하지만, 색상들의 조화는 아름답습니다.
- 개인적으로 그림이라기에는 거리가 멀지 않나 싶다.
- 기계가 그려낸 그림 같다. 정교하고 촘촘한 선의 배치가 좋았지만 오차없이 일정한 배열이 무언가를 느껴지도록 하는데에는 살짝 아쉬웠다.
- 직선들이용해서 많은 색감의선으로그렸는데 색이 조화로운 것도아니고 뿔 나타내고 숨은건지 모르겠다
- 최근에 휴대폰 액정에 문제가 생겨 화면에 위 그림과 같은 줄이 생겼던 일이 있었는데 너무 똑같아서 깜짝 놀랐다
- 그리기힘들긴하겠으나 이런걸로 감정느끼기 어렵다
- The monitor seems to be broken.
- crossed the line
- it hurts my eyes
- dizzy
- will continue to stare
- It seems to give an optical illusion.
- colorful
- I don't know, it feels like the TV is broken
- It doesn't seem to stimulate the imagination.
- As soon as I saw it, I thought it was a wallpaper.
- The combination of bright colors and thin lines hurts the eyes, but the harmony of colors is beautiful.
- Personally, I think it's far from being a painting.
- It looks like a painting drawn by a machine. I liked the elaborate and dense line arrangement, but it was a bit disappointing to make a certain arrangement feel something without error.
- I drew lines of many colors using straight lines, but the colors aren't harmonious and I don't know what I want to represent.
- Recently, there was a problem with the screen of the mobile phone, and there was a line like the one in the picture above.
- It's hard to draw, but it's not worth it to feel emotions like this

Students Opinion
On this Painting



English is Google Translated.

- 뭘 표현한건지 모르겠다
- 정신없다
- 어지럽다
- 이걸 무슨 작품이지?라는 생각이 들었다
- 내 자신의 심오하고 힘든 마음을 표현한 그림같습니다
- 난잡하면서도 아름다운 신기함
- 나도 그릴 수 있을 것 같다. 의미부여하기 나름인 것 같다.
- 어떤 의도로 그랬는지 모르겠으나, 현재 저의 뇌 상태 같습니다.
- 그림에서 뭔가 작가의 혼란스러운 마음을 표현한 것 같다. 색채가 마음에 든다.
- 물감을 아무렇게나 뿌린 것 같다
- 지금까지 그림중에 가장 좋았습니다
- 난잡하게 마구 페인트를 부어버리고 만들어 낸 작품감이 느껴질 수도 있지만 노란색과 흰색 페인트가 역동적인 사람의 모습을 나타내는 것 같아서 격렬하고 강한 느낌의 전경의 위엄을 보여주는 것 같다고 느꼈다
- 과학적으로 세상의 모든 법칙은 엔트로피, 즉 무질서도가 상승하는 방향으로 진행된다고 한다. 이 그림 또한 무질서 속의 질서를 표현한 것 같다
- 난해하다. 어떤 것을 표현하려 한 건지 모르겠다.
- 솔직히 말하면 어린시절부터 이런 그림이 가까이 있다면 또 그런 그림 수 있겠다고 생각했다.
- 그림에서 분노가 느껴진다
- 무슨 그림인지 잘 모르겠다
- 사실 이런 그림을 볼때마다 누구나 그림수잇는 그림같은데 막 뿌리면 되지않을까 하지만, 생각보다 쉽지않다
- I don't know what you're expressing
- Lose self-confidence
- dizzy
- I thought, what kind of work is this?
- It is like a painting that expresses my own profound and difficult heart.
- A complicated yet beautiful novelty
- I think I can draw too. I guess it's up to you to give it a meaning.
- I don't know for what purpose I drew this, but it seems like my current brain state.
- The painting seems to express the artist's confused mind. I like the colors.
- It looks like it's been sprayed with paint
- Best drawing so far
- It may feel like a work made by pouring paint in a mess, but I felt that the yellow and white paint seemed to represent a dynamic human figure, so it seemed to show the nuances of a fierce and strong war.
- Scientifically, it is said that all laws of the world proceed in the direction of increasing entropy, that is, disorder. This painting also seems to express order in disorder.
- difficult I don't know what you were trying to express.
- To be honest, since I was a child, I thought that if paintings like this were valuable, I could draw them too
- I can feel the anger in the picture
- I don't know what picture it is
- Actually, whenever I see a picture like this, it looks like a picture that anyone can draw, and I wonder if I can just spray it on, but it's not as easy as I thought

SESSION 2

In the age of AI, do we need to change the definition of art?

Jaepark Rhee Chugye University for the Arts(Korea)

Students Opinion
On this Painting



English is Google Translated.

- 푸른색은 바다같은 검은색은 심해 같다
- 동양화 같다. 아름답다.
- 꽃 봉우리가 터지는 거 같다
- 약간 동양화? 같은 느낌이 들고 어딘가 모르게 화가 난 거 처럼 보였다
- 투박하고 거칠지만, 아름답다는 생각이 든다.
- 꽃잎의 색이 참 예쁜 것 같다
- 물감의 표현이 꽃을 표현한것처럼 느껴졌습니 다
- 화가나서 막 찍은 것 같다.
- 감정의 분출을 표현한듯
- 시드는 꽃의 일대기를 표현하는것같다
- 색 조합이 맘에 든다
- 파란색 꽃을 만든거같다
- 간단해보이지만 한땀한땀 표현해낸거같다.
- 파도가 바위에 부딪히는 느낌이 납니다. 치열 함이 느껴져요
- 그림이 뭔가 부서지는 파도를 표현한 것 같다. 과격한 붓터치가 흥미롭다
- 연작일 것 같다.
- 아름답다.
- 이 그림을 보자마자 파란색 꽃이 떠올랐는데 현실적으로 묘사된 것이 아니라 추상적인 느낌 이 들도록 질감이 처리된 것을 보고 현실의 꽃 이 아닌 꿈속에서 몽롱하게 등장하는 꽃을 상 상하게 되었다.

- Blue is like the sea and black is like the deep sea
- like an oriental beautiful.
- It looks like the flower buds are bursting
- A little oriental? I felt the same way and it looked like I was somehow angry
- It's rough and rough, but I think it's beautiful.
- I think the color of the petals is very pretty.
- I felt like the expression of the paints expressed the flowers.
- I think I just took a picture because I was angry.
- As if expressing an outburst of emotions
- Wither seems to express the life of a flower
- I like the color combination
- I think I made blue flowers
- It looks simple, but it seems to have been expressed one stitch at a time.
- Feel the waves crashing against the rocks. I feel the intensity
- The painting seems to represent a wave breaking something. The intense brush strokes are interesting
- It seems to be a series
- beautiful.
- As soon as I saw this picture, a blue flower came to my mind, but when I saw that the texture was processed to give an abstract feeling rather than a realistic depiction, I imagined a flower that appeared vaguely in a dream rather than a real flower.

Students Opinion
On this Painting



English is Google Translated.

- 노을같다
- 물속의 물고기를 표현한거 같아
- 나뭇잎 날개같다
- 미술은 어렵다
- 희망이 느껴진다
- 흰색 부분이 곰팡이 핀 단호박을 표현한 거 같 기도 하다.
- 물고기를 이다
- 정열적인 색이 열정이 느껴진다.
- 호박 조각들이 하늘에 떠있는 느낌이다.
- 불타는 느낌이다.
- 붉은색이 안정감을 주는 느낌이었다
- 불타는 이미지가 연상된다.
- 이해하기 힘든 그림같다.
- 불만 곳에 물뿌리는 거 같다
- 주황색과 부분들이 모여져 있는것이 물고기처럼 보였습니다
- 역동적이다
- 새의 털 모습을 확대하여 그린 느낌이다.
- 처음 딱 보고 색감이 되게 이쁘다고 생각했고 마치 불가사리 같은 느낌이었다
- 불길같은 그림 듯 하다.
- 용암이 솟는것같다
- 색조합이 맘에 든다
- 물고기 나뭇잎 연상케하는 색감입니다! 부드러운 텍스처가 편안하게 하기도 합니다.
- 금붕어 여러마리가 함께헤엄치는것같은 느낌이 들었다. 단순함 속에 자유로움이 느껴진다.
- 어느 동물도 입을 벌리고 있는 포도송이만큼 바 글바글대는 새끼들한테 먹이를 넘겨주는 것 같다. 누가 누구에게 잠식당하는 지 알 수 없는 것 처럼 보이는 기법이 좋다.

- like the sunset
- I think it represents a fish in the water
- like moth wings
- art is hard
- I feel hope
- It seems that the white part expresses the moldy sweet pumpkin
- are the fish
- Passionate colors feel passionate.
- It feels like the pumpkin pieces are floating in the sky.
- It feels like burning.
- The color red gave me a sense of stability.
- Burning images come to mind.
- It's a hard picture to understand.
- It's like sprinkling water on dissatisfied places.
- The orange colored parts looked like a fish dynamic
- It feels like an enlarged drawing of a bird's fur.
- When I first saw it, I thought the color was very pretty, and it felt like a starfish.
- It looks like it drew fire.
- It looks like lava
- I like the color combination
- The color is reminiscent of the fish Nemo! The soft texture is also soothing.
- It felt like several goldfish swimming together. There is a sense of freedom in simplicity.
- It seems that an animal is handing over food to the young as busy as a bunch of grapes with their mouths open. A technique that makes it seem impossible to know who is taking over by whom is good.

SESSION 2

In the age of AI, do we need to change the definition of art?

 **Jaepark Rhee** Chugye University for the Arts(Korea)

이번 실험을 통해 발견한 점




- 인간이 무엇을 느끼기 위해 작가가 반드시 인간이어야 할 필요는 없다.
- 인간이 무엇을 느끼는가는 개인차가 너무 크다(정성 평가).
- 그럼에도 불구하고 집단의 의견을 확인할 수는 있다(정량 평가).
- 구상화일수록 더 아름답다고 느낀다.
- 기계는 구상화에서도 추상화에서도 튜링테스트를 통과했다.
- 예술은 인간만 할 수 있다는 정의에 수정이 필요해 보인다.

Findings from this experiment

- A creator does not have to be a human being to make a human feel something.
- There is too much individual gap in what humans feel (qualitative evaluation).
- Nevertheless, it is possible to ascertain the opinion of the group (quantitative evaluation).
- More figurative it is, the more beautiful it is.
- The machine passed the Turing test in both figurative and abstract art.
- The definition that art can only be done by humans seems to need revision.

SESSION 3

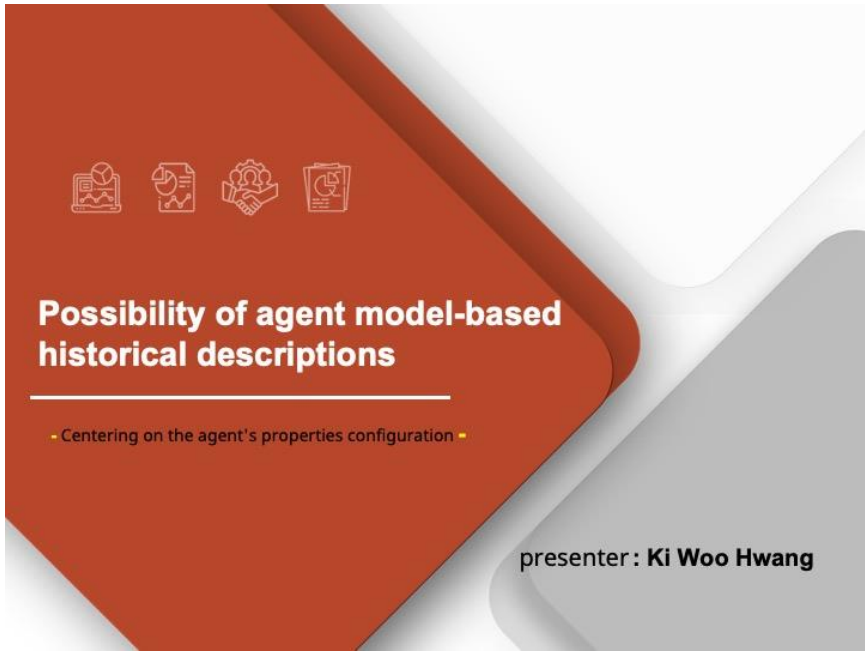
Moderator  Iksoon Jeong Chung-Ang University

| AI and Future & Past | |
|----------------------|---|
| 14:00 ~ 14:30 KST | <p>Possibility of Agent Model-based Historical Descriptions</p> <p> Kiwoo Hwang Sungkyunkwan University(Korea)</p> |
| 14:30 ~ 15:00 KST | <p>A Study on the English Old Poor Law System with Agent-Based Modeling Simulation</p> <p> Seunghwan Lee Sungkyunkwan University(Korea)</p> |
| 15:00 ~ 15:30 KST | <p>Charles Darwin's Legacy in Artificial Intelligence</p> <p> Jiyeon Kim & Youngsue Han Hankuk University of Foreign Studies(Korea)</p> |
| 15:30 ~ 16:00 KST | Discussion |

SESSION 3

Possibility of Agent Model-based Historical Descriptions

 **Kiwoo Hwang** Sungkyunkwan University(Korea)



Four icons representing AI, history, and data analysis are shown at the top. The main title is "Possibility of agent model-based historical descriptions". Below it is a bullet point: "Centering on the agent's properties configuration". The presenter's name "presenter : Ki Woo Hwang" is at the bottom right.

Who does the historical description?

◇ The relationship between the role of a historian and the agent



SESSION 3

Possibility of Agent Model-based Historical Descriptions

 **Kiwoo Hwang** Sungkyunkwan University(Korea)

The role of historians in digital history

- 01** Configuring agent properties
 - Work to bring the individual elements of history to life
 - The task of assigning the location, relationship, and role of individual elements in the entire history

- 02** The relationship between constructed attributes and historians
 - The properties of the completed agent do not allow historians to intervene.
 - Agent activity (change) also does not allow historians to intervene.

Configure Agent Properties-I

Simplifying semantic structure

It means extremely simplifying the human language structure, and this work is digital coding of human language.

For example, among the numerous causes of the Holocaust in the existing historical narrative of the Holocaust is the "will of the Nazi leadership." In order for the will of the Nazi leadership to function with the Holocaust within the simulation, the detailed factors that make up the will of the Nazi leadership need to be converted into simple language. Simple linguistic logic is easy to function.

A functional translated language fills the contents of the attribute.

SESSION 3

Possibility of Agent Model-based Historical Descriptions

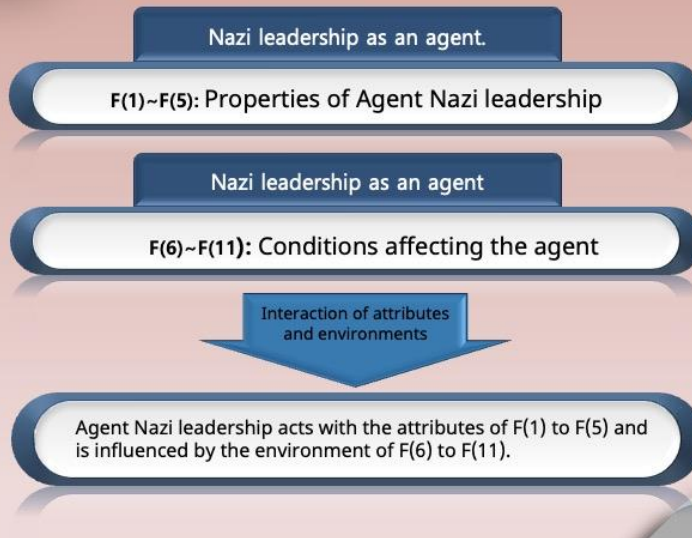
 **Kiwoo Hwang** Sungkyunkwan University(Korea)

Configure Agent Properties-II

Work to break down Nazi leadership's will into simple linguistic structures

| | |
|-------|--|
| F(1) | Anti-Semitism is strong. |
| F(2) | They put their own interests before others. |
| F(3) | It prefers public agitation to maintain stable power. |
| F(4) | Self-ethnic superiority is strong. |
| F(5) | They tend to blame other peoples and other countries for the problems of their society. |
| F(6) | Anti-Semitism in Europe before the Holocaust was a common phenomenon. |
| F(7) | Since 1943, the war has been at a disadvantage. |
| F(8) | During the war, the German people were emotionally unstable. |
| F(9) | During the war, the German people's doubts and dissatisfaction with the Nazi leadership increased. |
| F(10) | Nazi Germany lacked labor in the production of war supplies. |
| F(11) | Jews under the Nazis provided labor to Germany. |

Configure Agent Properties-III



SESSION 3

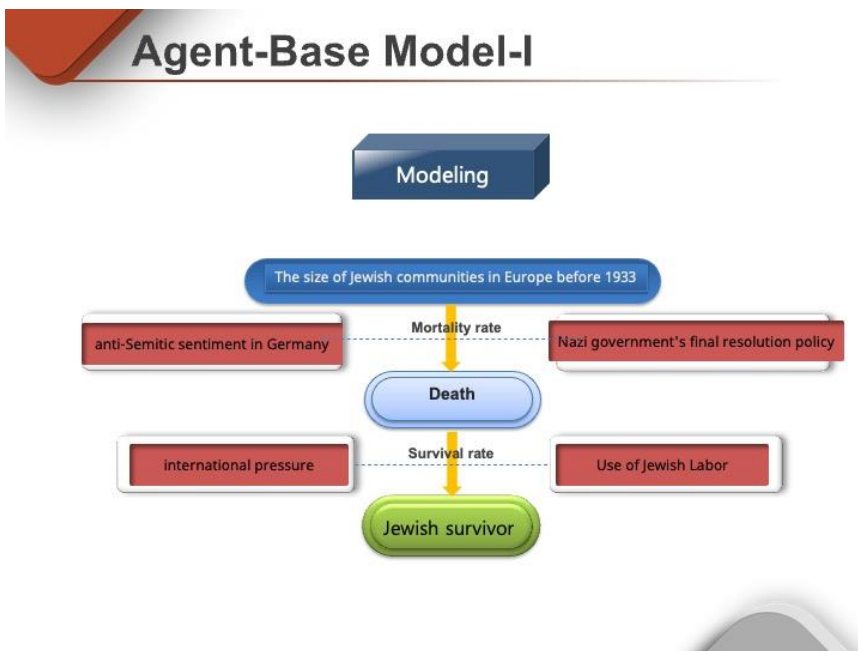
Possibility of Agent Model-based Historical Descriptions

 Kiwoo Hwang Sungkyunkwan University(Korea)

Configure Agent Properties-IV

Another agent: Jews under Nazi Germany

| | |
|-------|---|
| F'(1) | There was a strong resistance to anti-Semitism. |
| F'(2) | He was ignorant of the intentions of the German government. |
| F'(3) | They did not have the organization or political power to resist. |
| F'(4) | They expect their situation to improve over time. |
| F'(5) | They have no choice but to obey the German government for survival. |

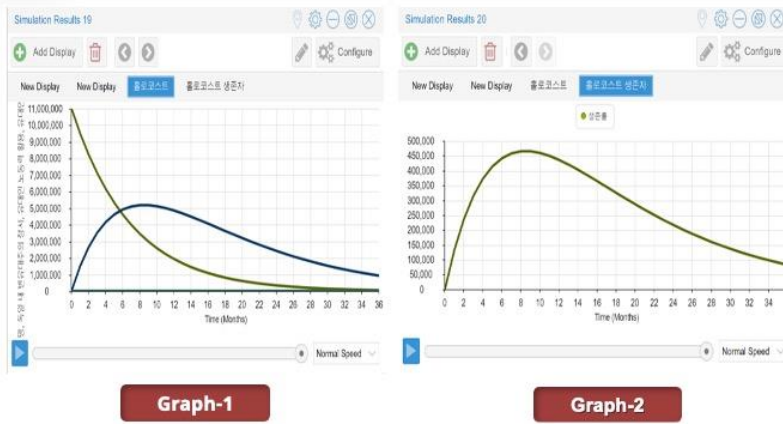


SESSION 3

Possibility of Agent Model-based Historical Descriptions

Kiwoo Hwang Sungkyunkwan University(Korea)

Agent-Base Model-II



Agent-Base Model-III

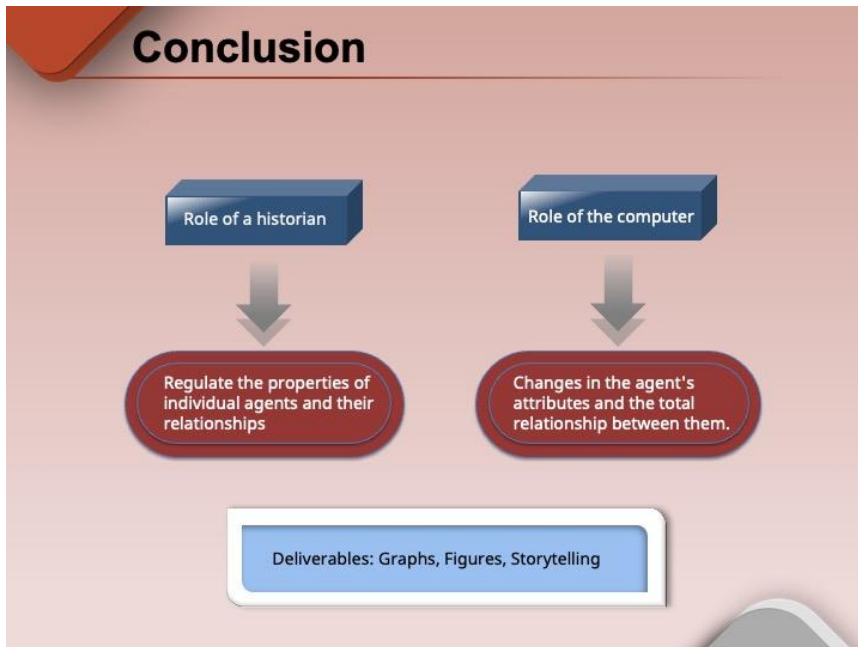
Graph-1 is the result of a computer's self-simulation with the model finished. The graph is in the form of a video, not a picture, and the model maker is blocked from intervening. The data show how the three stages are transformed by four factors. First of all, the monthly change shows that the number of Jewish victims changes significantly in 1941, but is less than in 1943 based on the absolute value. Under four conditions, Jewish sacrifices peaked in 1943 and gradually declined. International pressure and the use of Jewish labor to produce war supplies have contributed to the gradual decline in the number of victims since 1943. Of course, if other conditions were set, the aspects of the graph would look different.

Graph-2 is a result of identifying specific factors among the simulation results of a manufactured model. This result shows the number of surviving Jews in Europe at the time, centered on the Holocaust. This data shows that the number of Jews in Europe is gradually decreasing during the Holocaust.

SESSION 3

Possibility of Agent Model-based Historical Descriptions

 **Kiwoo Hwang** Sungkyunkwan University(Korea)



Thank you

SESSION 3

A Study on the English Old Poor Law System with Agent-Based Modeling Simulation

👤 Seunghwan Lee Sungkyunkwan University(Korea)

Abstract

At the time, the old Poor Law in the early 19C England was criticized as the cause of creating a vicious cycle of poverty by reducing the will of the poor to work, increasing unemployment, and reducing wages, and spreading poverty. In the early 19th century, this led to a social debate about the reform of the Poor Law in England, and eventually the Poor Law amendment was passed in 1834, ending the old Poor Law system in England. There are various historical interpretations of the collapse of the old Poor Law system and its causes. For example, neo-traditionalist Poor Law researchers asserted that collapse of the Poor Law system and the Poor Law Amendment of 1834 was a political movement and ideological offensive of industrial bourgeoisie to create a national labor market. On the other hand revisionist argued that the rapid increase in the cost of relief in the early 19th century, was due to the selfish and lax administration of landowners and tenants in rural areas. Previous studies have focused on the contradictions and instability of the old Poor Law system.

However, the old Poor Law system was maintained for about two centuries from the beginning of the 17th century to the beginning of the 19th century. If we understand the old Poor Law system as an unstable and contradictory system, we cannot explain why the old Poor Law system was able to maintain for such a long time. And it is worth noting that the technological innovation of the Industrial Revolution occurred intensively during this period, especially in the early and mid-18th century. If we regard the old Poor Law system as a mechanism that made the poor indolent and thereby created the cycle of poverty and spread it, we will not have a correct understanding of these technological innovations and the British Industrial Revolution that began at this time. And we should not regard the Poor Law as a system created and enforced by the ruling class. In such a case, the interaction between classes and its complex dynamics surrounding the Poor Law will be overlooked. Therefore, in this paper, the Poor Law is understood as follows. First, the old Poor Law system was

SESSION 3

A Study on the English Old Poor Law System with Agent-Based Modeling Simulation

Seunghwan Lee Sungkyunkwan University(Korea)

stable until at least the end of the 18th century. This means that both the benefactors and beneficiaries of the Poor Law were benefiting from this system, and they had the incentive to maintain the system. Second, the stability of the old Poor Law system is closely related to the political, economic and social development of 18th century British society. We will explore one aspect of this relationship. Third, the old Poor Law system was not maintained by one-sided interests of any class, benevolence, compassion, or morality. For a comprehensive understanding of the old poor law system, not only the beneficiaries but also the benefactors should be considered, and the interaction between these two classes should be examined.

We use the following method to explore this based on computational history. First, we use computational historical methods, Agent-Based Modeling (ABM), for a novel way of understanding of the Poor Law. ABM is a simulation method to understand whole system from agents and the interaction between agents. It should be a suitable methodology to reproduce and evaluate the interaction between classes in the old Poor Law system and the stability of the old Poor Law system. Second, we focus on the evolutionary game theory(EGT). EGT is a theory that introduces game theory to explain the occurrence of social cooperation in the evolution. It is expanding its scope of application to sociology, economics and even history through the iterative prisoner's dilemma game experiment of Axelrod. We interpret the stable period of the Poor Law as a state in which the phase of cooperation is maintained in the iterated prisoner's dilemma game of EGT. And in this process, stable technological innovation and capital accumulation occur through ABM. In addition, the process of the collapse of the old Poor Law system after such a stable period is interpreted as the transition of cooperative games to non-cooperative games.

First, we construct a Agent Based Artificial Society (ABAS), a virtual society where a agent do product and interact with other agents. There are phenomenons like capital accumulation, technological development, mutual aid, and material looting. ABAS is a artificial society that seeks a balance between increase in technology and stability, and we measure the point at which the balance between technology increase

SESSION 3

A Study on the English Old Poor Law System with Agent-Based Modeling Simulation

👤 Seunghwan Lee Sungkyunkwan University(Korea)

and social stability is achieved through interactions between agents and at this time the state of agents is observed and measured. The agents repeatedly play the prisoner's dilemma game by selecting the two strategies, the strategy of sharing goods with other agent and the strategy of not sharing the goods. Since the goal of all agents is to keep society stable and to accumulate capital at the same time, agents find a balance between poverty and accumulation in order to accumulate maximum capital while avoiding maximum disturbance in the artificial society. In addition, we simulate and observe the process in which the balance of this system is disrupted. In this case Agents' unrestricted accumulation of capital and chronic disturbance occur. It explains the process of the collapse of the old Poor Law system through the transition of the system from the equilibrium of the cooperative game to the equilibrium of the non-cooperative game. For example, the state of unlimited capital accumulation and chronic turmoil can be compared to the swing riots of 1830.

In summary, this paper assumes that agents in a artificial society play a iterated 2X2 game with a strategy of mutual aid and plunder. We simulate the case where the Nash equilibrium of the game is achieved in cooperation and the case in defect, and analyze the conditions of each case. We evaluate each state of equilibrium and its relationship to the system(capital accumulation, technological development, and population of system). Finally, the transition from stability to instability is simulated, and how the conditions of the transition and social instability affect the system are observed and evaluated. And based on the simulation data collected in this way, this thesis attempts a new historical view of the old Poor Law by comparing it with the real data and historiography about old Poor Law system.

SESSION 3

A Study on the English Old Poor Law System with Agent-Based Modeling Simulation

👤 Seunghwan Lee Sungkyunkwan University(Korea)

행위자 기반 모델링(ABM)을 통한 영국 빈 민법 연구

시뮬레이션

- 문제 해결을 위한 알고리즘의 하나.
- 일종의 사고 실험/가상 실험이지만, 정확한 해를 기대하는 것은 아니다.
- 컴퓨터 시뮬레이션: 컴퓨터를 사용하여 어떤 시스템을 모사하는 것(Imitation of a system via computer). 즉 시뮬레이션은 기본적으로 모델(model)을 구축한다. 모델은 시스템을 추상화한 것이다.
- 대표적인 시뮬레이션 방법으로 시스템 다이내믹스와 행위자 기반 모델링(ABM)이 있다. 시스템 다이내믹스는 거시적인 방법론이라면 행위자기반 모델링은 미시적인 방법론이라 할 수 있다.

SESSION 3

A Study on the English Old Poor Law System with Agent-Based Modeling Simulation

👤 Seunghwan Lee Sungkyunkwan University(Korea)

행위자 기반 모델링(Agent-Based Modeling)

- 다수의 행위자(agent)로 구성되어 있음.
- 다수의 행위자들은 복잡하고, 비선형적, 불연속적인 상호작용을 한다.
- 시스템 구성요소들의 성질 이외의 성질이 출현하는 시스템, 즉 창발적 특성을 가진 시스템을 모델링하기에 적합.
- 시스템 구성 요소들의 현재 상태가 과거의 영향 하에 있는 경우, 즉 구성원이 과거 경험을 통해 학습을 하고 이를 체화하여 현재의 행위에 반영하는 양태를 모델링하기에 적합.

행위자 기반 모델의 행위자

- 자율성: 행위자들은 독립적으로 정보를 분석하고, 의사결정을 행한다.
- 이질성: 각 행위자는 다른 행위자들과 구분되는 고유의 개체, 따라서 그 속성과 행위에 있어서 다를 수 있다.
- 목적중심성: 각 행위자의 행위는 목적에 귀속된다.
- 제한된 합리성: 행위자는 완벽한 합리적 개인으로 상정되지 않는다. 따라서 목적을 추구하는데 있어서 제한적인 합리성을 갖는다.
- 상호작용: 행위자는 다른 행위자들과 상호작용한다.
- 학습/적응: 행위자는 지난 행위로 부터 경험을 얻는다. 그리고 이 경험을 바탕으로 학습을 하며, 이를 현재의 행위에 반영할 수 있다.

SESSION 3

A Study on the English Old Poor Law System with Agent-Based Modeling Simulation

👤 Seunghwan Lee Sungkyunkwan University(Korea)

행위자 기반 모델링의 학문적 활용-1

- 콘웨이의 인생게임(Conway's Game of Life)
- 일종의 세포 자동차(cellular automata) 프로그램
- 폰 노이만의 자기 재생 조직(self-reproducing organization)을 ABM으로 구현한 것.
- 정사각형의 그리드들이 이어진 공간에 살아있는 세포와 죽은 세포에 대한 간단한 규칙을 설정하면 세포들이 삶과 죽음을 반복하면서 다양한 패턴들을 만들어 내는 게임.
- 시작시에 설정한 간단한 규칙으로는 설명하기 힘든 창발적 현상들이 나타난다. 즉 창발성을 실제로 보여주는 대표적인 ABM.
- 유전자 복제 구조연구 등에 사용됨.

행위자 기반 모델링의 학문적 활용-2

- 셸링의 동적 분리 모델(Schelling's model of segregation)
- Dynamic Models of Segregation, 1969
- 미국의 흑인과 백인의 거주지 분리에 대한 설명을 위해 토머스 셸링이 제시한 모델.
- 이 모델에서 행위자들은 자신과 같은 색의 행위자들을 이웃으로 두었을 때 만족감을 느낀다. 즉 같은 색의 행위자를 선호하는 것.
- 행위자들은 자신의 만족을 높이기 위해 동일한 색상을 지닌 행위자가 위치한 셀의 주변으로 이동한다.
- 셸링은 이 모델을 통해 행위자의 아주 작은 선호로부터 시스템 전체의 현저하고 거대한 분리가 나타날 수 있다는 것을 보여줌.
- ABM의 사회학적 활용이라는 측면에서 의미가 있음.

SESSION 3

A Study on the English Old Poor Law System with Agent-Based Modeling Simulation

Seunghwan Lee Sungkyunkwan University(Korea)

행위자 기반 모델링의 학문적 활용-3

- 슈가스케이프(Sugarscape)
- Growing Artificial Societies(Joshua Epstein&R. Axtell, 1995).
- 아래로부터의 사회학 시뮬레이션(social science from the bottom up).
- 경제가 어떻게 발생하는 가를 인 실리코(in sillico: 컴퓨터 모의 실험)를 통해 보여주고자 고안한 시뮬레이션.
- 경제의 기본적 요소인 소비자, 생산자, 시장 등을 배제하고 인간과 자원 만이 존재하는 모델을 구성하여 여기서 어떻게 경제활동의 최초의 고리가 나타나는 지를 관찰.
- 행위자, 환경, 규칙의 세 가지 기본 요소들만으로 이른바 인공사회(Artificial Societies)를 구성한다.
- 슈가스케이프라는 가상의 섬에 행위자(인간)와 자원(설탕)을 배치하고 행위자가 자원을 채집하는 것을 시작으로 빈부의 격차, 무역, 시장, 채무 등이 나타나는 것을 시뮬레이션함.
- 그 결과 슈가스케이프에서는 운동, 자원채집, 재생산, 번식, 전투, 문화 교류, 무역, 유전, 이윤, 오염, 질병 등 다양한 사회적 요소들이 출현한다.

슈가스케이프와 역사적 지식

- 슈가스케이프(시뮬레이션)는 일종의 닫힌 계. 자기완결적 구조. 따라서 이를 통해 현실에 대한 지식을 어떻게 얻을 것인가에 대한 고찰이 필요하다. 즉 ABM 시뮬레이션이라는 닫힌 계와 현실 사이의 연관성을 보증해주는 것이 무엇인지에 대한 의문은 끊임없이 제기될 수밖에 없다.
- 시뮬레이션은 현실에 대한 추상, 이는 우리의 지식과 크게 다르지 않다. 시뮬레이션은 현실에 대한 재현이 아니라, 현실을 추상한 개념의 재현. 따라서 시뮬레이션을 통한 지식은 현실에 대한 추상과 개념에 대한 지식. 시뮬레이션은 그 원본인 현실을 필요로 한다.
- 하지만 역사적 현실은 이미 존재하지 않는 것. 역사 시뮬레이션을 구축하려면 사료와 데이터 뿐 아닌 다른 원본이 요구된다. 왜냐하면 사료와 데이터는 현실이 아니기 때문이다. 따라서 우리는 역사를 시뮬레이션 하기에 앞서서 시뮬레이션의 원본이 되기에 충분한 형태로 역사를 구성한 무엇을 필요로 한다. 우리는 이를 역사적 해석이나 내러티브라고 판단한다.
- 그렇다면 시뮬레이션은 이미 해석된 역사를 컴퓨터 프로그램을 통해 모사하는 것인가? 우리는 ABM 통해 창발적 요소들을 도출할 것을 기대한다. 즉 시스템을 구성하는 개개의 원자들의 성질로는 설명할 수 없는 전혀 새로운 현상이 ABM을 통해 나타날 것이며, 이를 통해 원본이 되었던 역사적 해석을 보완하거나 대체할 것을 기대하는 것이다.

SESSION 3

A Study on the English Old Poor Law System with Agent-Based Modeling Simulation

👤 Seunghwan Lee Sungkyunkwan University(Korea)

근대 영국 빈민법

- 1601년 엘리자베스 빈민법
- 1662년 정주법
- 1723년 Workhouse Test Act
- 1782년 길버트 법
- 1795년 스펄햄랜드 제도
- 1818년 스텐다지스 본 법
- 18세기 말~19세기 초 이른바 빈민법 논쟁
- 1834년 빈민법 개혁

구빈민법 체제의 작동

- 재산을 가진 사람은 교구에 세금을 납부. Poor Rate
- 이는 교구의 빈민감독관에 의해 징수되었음. 이 빈민감독관은 1년 임기의 무급직이었음.
- 이 감독관이 세금을 집행할 권한을 가지고 구빈을 행했음.
- 당시 빈민을 두 종류로 분류, 근로능력이 있는자와 없는자.
- 만약 빈민이 다른 교구로 이주하고자 한다면 교구의 허가증이 있어야 했음. 이것은 허가증을 발급할 교구에 허가증을 소지한 빈민에 대한 구빈 책임이 있음을 밝히는 것이었음.
- 하지만 이 시스템은 매우 유동적이어서 각 교구는 자신의 지역에서 빈곤문제에 대한 대처에 있어서 어느정도의 자율성을 가지고 있었음.

SESSION 3

A Study on the English Old Poor Law System with Agent-Based Modeling Simulation

👤 Seunghwan Lee Sungkyunkwan University(Korea)

1601년 엘리자베스 구빈법

- 교구중심주의 확립, 원외/원내구제 구분
- 1815년까지 각 교구는 자신의 교구에 거주하는 빈민의 구빈을 책임져야 했음. 구빈비용은 교구민들의 세금에 의해 충당.
- 원외구제/원내구제
- 특기할 것은 원외구제, 즉 노동능력이 있는 빈민에게도 구호를 제공했다는 점.
- 원내구제는 workhouse에 수용된 노동능력이 없는 빈민에게 제공하는 구빈형태를 말하는 것.
- 하지만 17세기 노역소는 빈민수용시설이라기보다는 작업장에 가까웠다.

순회노동제도(Roundsman System)

- 빈민감독관은 일할 능력이 있는 빈민에게 ticket을 교부할 수 있었다.
- 빈민들은 지역의 일손이 필요한 농업가에게 이 ticket을 가지고 갔다.
- 빈민들은 하루의 노동이 가능했고 노동을 마치면 빈민법으로 얻을 수 있는 하루치의 금품을 받을 수 있었다.

SESSION 3

A Study on the English Old Poor Law System with Agent-Based Modeling Simulation

👤 Seunghwan Lee Sungkyunkwan University(Korea)

길버트 법

- 영국 구빈법의 핵심적 제도 중 하나는 노역소(Workhouse).
- 노역소는 근로 가능한 빈민이 일을 하기 위해 수용되는 곳.
- 교구는 자신의 노역소를 운영했음. 없는 교구도 다수 존재.
- 1782년 길버트 입법이 이를 바꾸었다.
- 교구들은 연합하여 하나의 노역소를 운영할 수 있게 되었음.
- 또한 노역소에 수용되는 빈민은 근로능력이 없는 빈민으로 제한 됨.
- 하지만 이 입법에 따라 노역소를 짓고 운영하는 교구는 거의 없었음.

스핀햄랜드제도(Speenhamland System)

- 18세기 말의 흉작과 인플레이션 그리고 프랑스 혁명의 위협 하에서 성립.
- 빈민법 개혁과 관련하여 가장 큰 논란을 불러일으켰던 제도.
- 빈민들에게 수당을 지급. 현금을 급여했음.
- 구빈을 통해 지급되는 수당이 가족의 수, 생계비(빵값)에 연동되도록 한 것.
- 생계비가 높을 수록, 가족이 많을 수록 더 많은 수당을 수령할 수 있었음.

SESSION 3

A Study on the English Old Poor Law System with Agent-Based Modeling Simulation

👤 Seunghwan Lee Sungkyunkwan University(Korea)

빈민법의 문제들(1)

- 순회노동제도는 고용주들이 월급을 낮게 유지하게 만드는 문제점이 있었다. 어차피 빈민들의 임금은 구빈법에 정해져 있는 만큼 자유로운 노동자들에게 그 이상의 임금을 줄 유인이 없었음. 또한 노동자들도 생산성을 높일 필요를 느끼지 못했음.
- 스펜햄랜드 제도는 빈민에게 지급하는 돈이 생계비와 연동되어 있는 관계로 물가 상승에 따라 구빈지출이 증가했음. 또한 당대에 빈민들의 자활의지를 꺾고, 무절제한 출산을 부추긴다는 비판에 직면.

빈민법의 문제들(2)

- 곡가 상승, 흉년, 무역에서의 불경기와 실업률 상승이 겹치면서 19세기 초 영국의 빈민의 수는 꾸준히 증가. 나폴레옹 전쟁 또한 이에 일조하였음.
- 구빈 비용은 19세기의 시작부터 1830년대까지 꾸준히 상승. 1830년대가 되면 지방정부가 감당하기 힘들 정도로 증가.
- 이에 빈민법에 대한 비판의 목소리가 커짐.
- 대개 빈민들의 도덕적 타락, 게으름과 구빈행정의 비효율성, 부패 등에 대한 여론의 비판이 거세지기 시작함.

SESSION 3

A Study on the English Old Poor Law System with Agent-Based Modeling Simulation

👤 Seunghwan Lee Sungkyunkwan University(Korea)

빈곤에 대한 여론의 변화

- 산업혁명 초기 토머스 페인, 로버트 오웬 등 빈민에 대한 동정과 연민.
- 자유주의가 사회적으로 확산, 중상주의적 경제정책을 자유방임주의 경제정책으로 전환해야 한다는 목소리가 점차 증가.
- 벤담은 빈민법 개혁을 주장. 효율적인 구빈행정을 위해 구빈을 전담하는 일종의 회사를 설립 운영해야 한다고 주장.
- 맬서스는 한 발 더 나아가 빈민법 자체를 폐지해야 한다고 주장. 빈민법이 빈민의 자활에 그 어떠한 도움이 되지 않을 뿐 아니라 오히려 빈민들을 빈곤의 덫에 빠뜨린다고 주장. 또한 빈민법은 빈민의 인구를 늘려 더 거대한 빈곤을 불러올 것이라고 주장.

1834년 빈민법 개혁

- 1817년 하원 빈민법 조사 위원회: 빈민법 폐지
- 1817년 상원 빈민법 조사 위원회: 빈민법 옹호
- 1832년 왕립 빈민법 위원회: 1830년 스윙폭동 이후 빈민법 개혁 주장이 여론의 혜게모니를 장악. 즉 빈민법은 유지하되 개혁이 필요하다는 것. 다수의 자유주의 지식인들이 빈민법 위원으로 위촉, 1834년 이들의 보고서를 바탕으로 빈민법 개혁 입법이 상하원을 통과.

SESSION 3

A Study on the English Old Poor Law System with Agent-Based Modeling Simulation

👤 Seunghwan Lee Sungkyunkwan University(Korea)

신빈민법의 원칙

- 열등처우의 원칙: 구빈의 수준이 노동자의 생활수준을 넘어서는 안됨. 이를 통해 이른바 근로 의식을 고취하고자 했음.
- 노동 가능 빈민과 노동 불가능 빈민의 엄격한 구분: 노동 가능 빈민을 구빈에서 제외.
- 원외구제의 폐지: 구빈 대상인 모든 빈민은 원칙상 노역소에 수용.
- 구빈행정의 효율화: 구빈행정에 필요한 인력을 유급인력으로 교체, 구빈행정의 중앙집권화 시도

빈민법 연구의 흐름

- 1) 전통적 해석
- 2) 신전통주의적 해석(취극적 해석)
- 3) 수정주의적 해석
- 4) 절충주의적 해석
- 5) 최근의 흐름(가족 정책을 통한 접근, 젠더적 접근, 산업혁명과의 관계를 통한 접근, 비교사적 접근: 중국과의 비교)

SESSION 3

A Study on the English Old Poor Law System with Agent-Based Modeling Simulation

Seunghwan Lee Sungkyunkwan University(Korea)

신전통주의적 빈민법 해석

- 첫째, 신빈민법 개혁은 자유주의적 부르주아가 온정주의적 구질서를 유지하고자 했던 지주계급에 대해 승리를 거둔 것이다(웹 부부).
- 둘째, 따라서 구빈민법 체제 자체는 구질서의 상징이자 산업사회 형성의 장애물이었다(웹 부부).
- 셋째, 빈민법 개혁 당시 가장 큰 비판의 대상이 되었던 원외구제, 그 중에서도 스피햄랜드 제도는 노동 시장 창출을 막기 위한 지역 기반 세력들의 마지막 노력이었다(플라니).
- 넷째, 따라서 스피햄랜드 제도는 근대적 경제체제와 양립불가능한 시대의 흐름에 역행하는 제도였다(플라니).

수정주의적 해석의 요점

- 첫째, 19세기 초 영국 사회의 경제적 문제들, 실업과 경기 불황 등이 구빈민법 체제의 모순에서 출현한 것이라고 보기 힘들다. 왜냐하면 구빈민법은 지역과 시대에 따라 다양한 형태로 시행되었기 때문이다. 따라서 구빈민법 체제의 이해를 위해서 지역에 따른 사례연구에 집중할 필요가 있다.
- 둘째, 1834년 빈민법 개정을 온정주의 사회에 대한 시장경제의 승리 혹은 산업자본가 계급의 지주계급에 대한 승리로 볼 수 없다. 사례적으로 보았을 때, 빈민법 개정 이후에 농촌 지역에서 지주층의 영향력이 약화되었다는 근거가 없다.
- 셋째, 빈민법 개혁 이후에도 빈민 행정의 중앙화가 관철되지 않았다. 구빈행정에 실무적인 영향을 미친 것은 지역의 유력가들, 즉 타운의 자치능들, 직인들, 상인들이었다.
- 넷째, 개혁의 쟁점이었던 원외구제 철폐 또한 완전히 이루어지지 못했다.
- 다섯째, 따라서 빈민법 체제가 빈민법 개혁을 기점으로 성격이 변화한다고 볼 수 없으며, 구빈민법 체제와 신빈민법 체제를 나눌 명확한 이유를 찾기 힘들다.

SESSION 3

A Study on the English Old Poor Law System with Agent-Based Modeling Simulation

👤 Seunghwan Lee Sungkyunkwan University(Korea)

수정주의적 구빈민법 체제 이해의 한계

- 첫째, 연구들이 지역적으로 파편화되어 있고 또한 경제적 분석에 몰려있다. 따라서 구빈민법 체제에 대한 전체적인 이해와 조망을 끌어내기에는 무리가 있다.
- 둘째, 지역의 계급 역학 중에서도 주로 지주, 차지농, 타운의 중산계급 등 구빈세를 부담했던 납세자 중심의 접근에 그치고 있다. 따라서 구빈의 대상이었던 노동자, 빈민에 대한 분석과 이들과 납세자들이 어떤 상호작용을 했는지에 대한 분석이 부족하다.

절충주의적 해석의 요점

- 첫째, 절충주의적 해석은 신전통주의의 빈민법 개혁에 대한 이데올로기적, 정치적 해석을 수정주의의 농촌 중심적 사례 연구의 성과에 접목하였다.
- 둘째, 절충주의적 해석은 빈민법을 둘러싼 농촌 사회의 계급적 이해에 따른 역학, 이데올로기 성향, 계층 성향에 대한 정치적, 사회적 분석을 수행했다.
- 셋째, 그간 간과되어 왔던 농촌 공동체에서의 구빈민법의 역할을 긍정적으로 재조명하였다.
- 넷째, 구빈민법 체제는 농촌 사회에서 상호부조의 일종으로 실업보험과 같은 사회보장의 역할을 일정부분 수행하였고, 이를 통해 근대적인 전국적 노동시장의 창출에 기여하였다.

SESSION 3

A Study on the English Old Poor Law System with Agent-Based Modeling Simulation

👤 Seunghwan Lee Sungkyunkwan University(Korea)

새로운 연구 경향

- 다양한 방법론과 주제들이 빈민법 연구에 도입되고 있다. 젠더적 접근, 산업혁명과의 연관성 탐구 등.
- 구빈민법과 사회적 혁신 사이의 긴밀한 연관을 주장하는 새로운 주장이 나타났다.
- 이는 구빈민법 자체를 구시대의 유산으로 보았던 기존의 시각과는 확연히 다른 접근이다.
- 이러한 시각은 이른바 4차 산업혁명과 기본소득제 논의에 시사하는 바가 있다.

종합

- 첫째, 전통주의적 시각에서 빈민법은 빈민들의 근로의욕을 감소시키고, 실업을 증가시키며, 임금을 감소시켜 결과적으로 사회적 빈곤의 굴레를 만들어내는 역할을 하는 것이었다.
- 둘째, 신정통주의적 시각에서 볼 때, 1834년 빈민법 개혁은 자유주의적 부르주아들의 시장 창출을 위한 이데올로기적 공세이자, 반이보주의적 처사였지만, 구빈민법 자체는 전근대적인 관습으로서 근대적 산업자본주의 체제의 창출에는 방해가 되는 것이었다.
- 셋째, 수정주의적 해석 하에서 구빈민법은 지역의 다양한 정치적, 경제적, 계급적 역학 하에서 시행된 것으로서 그 양상을 하나로 묶기에는 무리가 있지만, 이를 시대에 뒤떨어진 구습이라기 보다는 그때 그때의 시대적, 지역적 요구에 따라 능동적으로 대처하고자한 노력의 일환으로 평가할 여지가 있다.
- 넷째, 이러한 평가의 연장선상에서 구빈민법은 18~19세기 초 영국에서 일종의 사회적 안정장치의 역할을 했으며 현대 복지 제도의 시초로서 평가할 수 있다.
- 마지막으로 최근의 연구는 빈민법과 산업혁명 사이의 모종의 연관을 찾고있다. 빈민법의 복지제도로써 역할, 이로 인한 정치적, 사회적 안정이 산업 기술에의 투자와 발전, 더 나아가 산업혁명의 밑바탕이 되었다고 주장한다.

SESSION 3

A Study on the English Old Poor Law System with Agent-Based Modeling Simulation

👤 Seunghwan Lee Sungkyunkwan University(Korea)

구빈민법 체제와 산업혁명

- 구빈민법 체제에 대한 새로운 평가, Avner Greif, Murat Iyigun, Diego Sasson(2012), Avner Greif, Murat Iyigun(2013).
- 구빈민법과 사회적 혁신 사이의 긴밀한 연관을 주장.
- (...) "the 17th to the 19th century, English counties that provided more poor relief had fewer riots and more innovations (as measured by patents)."
- 이는 구빈민법 자체를 구시대의 유산으로 보았던 기존의 시각과는 확연히 다른 접근.

게임이론을 통한 구빈민법체제의 이해

- 빈민법 체제의 몇몇 특성을 게임이론을 통해 재구성.
- 빈민은 일방적인 시스템의 수혜자가 아님.
- 지배층 또한 일방적인 시혜를 통해 시스템을 유지한 것이 아님.
- 빈민과 지배층 둘 모두 시스템 내의 이해의 당사자.
- 빈민법 시스템을 유지하기 위해서는 두 집단 모두의 양보, 협력, 대립 등을 거친 일종의 합의가 필요했음. 이 합의 상태를 게임이론의 균형개념을 통해 설명할 수 있음.
- 게임이론을 이용한 빈민법 체제 이해를 통해 산업혁명과 빈민법 사이의 관계를 규명할 수 있음.

SESSION 3

A Study on the English Old Poor Law System with Agent-Based Modeling Simulation

👤 Seunghwan Lee Sungkyunkwan University(Korea)

게임이론

- 게임이론과 플레이어: 어떤 플레이어도 게임의 룰에 있어서 소외되어 있지 않다. 기본적으로 플레이어가 가지고 있는 정보는 균등하다.
- 게임이론에서의 정보: AkE/BkE- \rightarrow AkBkE/BkAkE-
 \rightarrow BkAkBkE/AkBkAkE
- 『게임이론』, 왕규호(2005); 죄수의 모자에 대한 우화.
- 죄수의 딜레마
- 내쉬 균형

복지 시스템과 게임이론

- 분배 문제(케이크 자르기, Envy-free cake-cutting)
- 최후통첩게임에서 최적의 해를 구하는 문제: 내쉬 협상이론(Nash, 1950), 루빈슈타인 협상이론(Rubinstein, 1982).
- 롤스, "무지의 장막" 사고실험에 대한 고찰(Kalai, 1977).
- 이러한 고찰들은 수많은 파레토 효율 중 공정성을 충족하는 지점을 찾기위한 것.
- Dean Kethy&Christ Kevin(1998), 홍경준(2017) 등도 사회 복지의 문제에 게임 이론을 이용하여 접근.

SESSION 3

A Study on the English Old Poor Law System with Agent-Based Modeling Simulation

👤 Seunghwan Lee Sungkyunkwan University(Korea)

진화게임이론과 상호부조

- 로버트 악셀로드의 죄수의 딜레마 프로그램 대회
- 반복되는 죄수의 딜레마 토너먼트에서 가장 높은 점수를 얻은 전략은 Tit-for-Tat.
- 한 번의 게임은 배신/배신의 전략이 내쉬균형.
- 하지만 계속되는 게임에는 다른 내쉬균형이 존재한다.
- 즉, 협력/협력을 지속하는 상태가 발생한다는 것.
- 이러한 개념을 통해 복지 문제에 접근한 예: Peter Vanderschraaf, 1999; Edgar J. Sánchez Carrera, 2012

구빈민법 체제와 협동게임

- 18세기 영국 구빈민법 체제는 협력/협력의 내쉬균형 상태.
- 구빈민법 체제는 빈민(피지배층)과 후원자(지배층) 사이의 2X2 게임. 각 플레이어의 전략은 소요(빈민)와 구빈(지배층).
- 온정주의적 구빈질서는 "하층민들의 존경과 복종, 지배층의 시혜를 맞바꾸어 안정된 질서를 유지하는 체제".
- 반복 죄수의 딜레마(iterated Prisoner's Dilemma) 게임을 통해서 에이전트들의 전략이 어디로 수렴하는 지를 살핌.
- 진화게임이론에서 협동의 상태, 즉 사회적 안정과 구빈의 상태에서 균형을 이루는 체제.

SESSION 3

A Study on the English Old Poor Law System with Agent-Based Modeling Simulation

👤 Seunghwan Lee Sungkyunkwan University(Korea)

빈민법 체제의 와해

- 협동 상태의 균형이 비협동 상태의 균형 상태로 이동한 것.
- 후원과 소요가 동시에 존재하는 체제(1830년 스윙폭동) 혹은 후원(정확히는 원외구제) 폐지와 소요가 존재하는 체제(1934년 빈민법 개정)로 전환된 것.
- 이 변화를 게임이론의 내적 논리를 통해 어떻게 설명해야함.
- 이를 시뮬레이션을 통해 보여주어야 함.

Game Change

- David Robert Rodson & David Goforth(2005), Topology of 2X2 Games.
- 이들은 위상학을 이용하여 2X2 게임의 종류가 전환된다는 것을 보여줌. 게임의 보수 행렬의 보수값 변화를 통해 게임의 양상이 변화됨. 예컨대 '죄수의 딜레마' 게임이 위상적으로 인접한 '사슴사냥 게임' 등으로 변화할 수 있다는 것.
- 이를 이용하여 이동한(2012)은 노사의 임금협상의 국면이 몇 가지 게임의 국면으로 전환됨을 보여주었음.

SESSION 3

A Study on the English Old Poor Law System with Agent-Based Modeling Simulation

Seunghwan Lee Sungkyunkwan University(Korea)

위상 변화의 예



빈민법 체제 변화 시뮬레이션

- David Robert Rodson & David Goforth(2005)의 이론을 이용하여 빈민법 체제가 협동게임에서 비협동 게임으로 변환됨을 시뮬레이션을 통해 보여줄 수 있음.
- 구빈민법 체제에서 왜 계급 간 협동과 산업혁명이라는 혁신이 일어났는지에 대한 또 하나의 설명을 제공할 수 있음.
- 또한 구빈민법 체제의 와해에 대한 설명을 제공할 수 있음.
- 이렇듯 ABM 시뮬레이션을 통해 구빈민법 체제와 산업혁명 그리고 빈민법 체제의 쇠퇴에 대한 새로운 역사적 시각을 제시할 수 있다.

SESSION 3

A Study on the English Old Poor Law System with Agent-Based Modeling Simulation

👤 Seunghwan Lee Sungkyunkwan University(Korea)

구빈민법에 대한 해석

- 구빈민법 시스템은 어떤 계급의 일방적인 양보나 선행, 연민에 의해 유지된 체제가 아니었음.
- 모든 계급이 주어진 상황에서 가장 최선의 전략을 찾은 결과 우연히 이루어진 평형상태.
- 하지만 이 우연한 평형상태가 산업혁명의 충분조건이 되었다.
- 마찬가지로 구빈민법 체제의 와해 또한 어떤 계급의 도덕적 해이나 계급적 악의에 의한 것이 아님.

문제점 및 앞으로의 과제

- 구빈민법 체제라고 뭉뚱그려서 1834년 이전의 구빈 시스템을 통칭하는 것은 문제가 있음. 1834년 개혁은 정확히 원외구제 폐지에 대한 것. 따라서 원외구제 시뮬레이션으로 특정할 필요도 있음.
- 마찬가지로 위에서 고려에 넣은 것은 주로 농촌에 대한 이야기임. 따라서 농촌의 경우로 특정할 필요가 있음.
- 구체적으로 구빈민법의 각 국면을 시뮬레이션 할 필요가 있음. 순회노동제도, 길버트법, 스피햄랜드 시스템 등.

SESSION 3

Charles Darwin's Legacy in Artificial Intelligence

 **Jiyeon Kim & Youngsue Han** Hankuk University of Foreign Studies(Korea)

Abstract

My presentation deals with legacy of Charles Darwin in artificial intelligence. Current rise of boom in artificial intelligence heavily relies on statistical learning. Statistical learning is a framework for machine learning drawing from statistical algorithms. The algorithms in statistical learning mainly divides into regression and classification. Francis Galton, scientist was a key figure in statistics. He founded numerous concepts in statistics. He devised regression, one of the most important concepts in statistics. He was also cousin of Charles Darwin. Darwin inspired Galton to incorporate eugenics, science of transforming the human race through selective breeding. The regression in AI follows parallel concept of selective breeding in eugenics. My presentation copes with dark story of statistical algorithms in contrast between eugenics and artificial intelligence.

SESSION 3

Charles Darwin's Legacy in Artificial Intelligence

 **Jiyeon Kim & Youngsue Han** Hankuk University of Foreign Studies(Korea)

Charles Darwin's legacy in Artificial Intelligence

Jiyeon Kim, Hankuk University of Foreign Studies
Youngsue Han, Hankuk University of Foreign Studies

CONTENTS

- Introduction
- Eugenics
- Plato / Charles Darwin / Francis Galton
- Regression
- Darwin's legacies on Galton's theory
- Dark side of regression and eugenics
- Conclusion

SESSION 3

Charles Darwin's Legacy in Artificial Intelligence

 Jiyeon Kim & Youngsue Han Hankuk University of Foreign Studies(Korea)



Introduction

- **Sayuri Fujita (1979-)**
- Japanese television personality based in South Korea
- Sayuri's childbirth stirs debates over choice to be single mother in Korea.
- The debate focussed on her decision to become a single mother: gender issue
- More important agenda in terms of **eugenics ignored.**

What is eugenics?

- **eugenics** (1800-1900) Greek eugenes "well born", from genes "born"
- A set of beliefs and practices that aim to improve the genetic quality of a human population, by excluding people and groups judged to be inferior or promoting those judged to be superior.

- Plato - Eugenics Footprint of the Greek Age

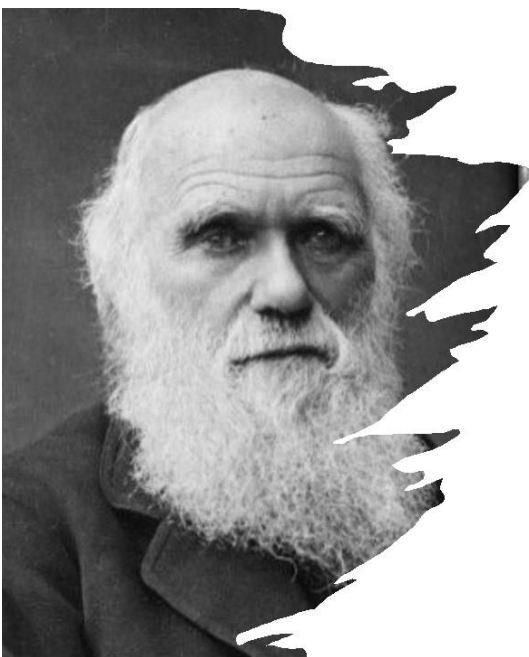
SESSION 3

Charles Darwin's Legacy in Artificial Intelligence

 Jiyoon Kim & Youngsue Han Hankuk University of Foreign Studies(Korea)

Plato's selective breeding

- Marry good men and women who are genetically good and have children, and prevent genetically bad men and women from having children from having children.
- Plato said that intellectual ability makes humans think of noble values and that finding truth in reality is possible through rational thinking. Also, only the clever can be the ruler of the public.



Charles Darwin (1809-1882)

- a British scientist who developed the Theory of Evolution
- the idea that plants and animals develop gradually from simpler to more complicated forms by natural selection.
- On the Origin of Species (1859)

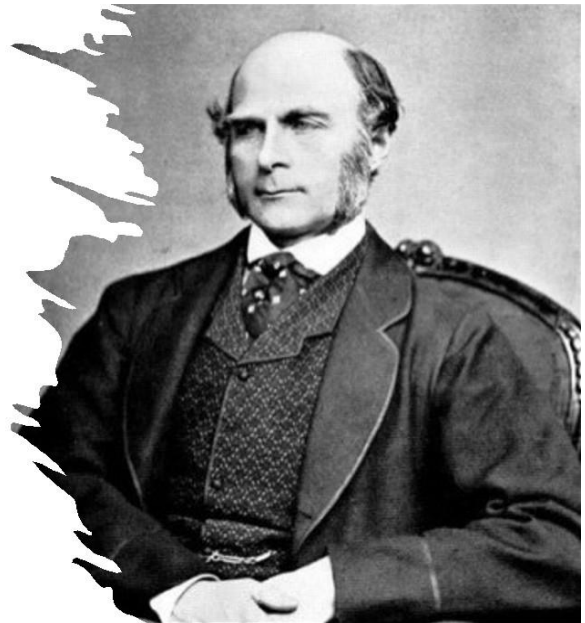
SESSION 3

Charles Darwin's Legacy in Artificial Intelligence

 Jiyoon Kim & Youngsue Han Hankuk University of Foreign Studies(Korea)

Francis Galton (1822-1911)

- Galton, a cousin of Charles Darwin.
- Galton developed eugenics based on Darwin's ideas after reading Darwin's *On the Origin of Species* (1859).
- Galton's interest
 - "improving human stock" through scientific management of mating
 - wanted to use statistical methods to explain the genetic and human differences in intelligence



What is regression?

- What is AI?
 - Artificial Intelligence, The study of how to make computers do intelligent things that people can do, such as think and make decisions
- AI's algorithm
 - Classification and Prediction
 - **Regression** belongs to the **prediction algorithm**

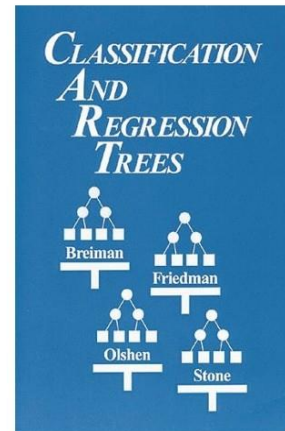
SESSION 3

Charles Darwin's Legacy in Artificial Intelligence

Jiyeon Kim & Youngsue Han Hankuk University of Foreign Studies(Korea)

Classification and Regression (1984)

- Breiman, Leo (1984) *Classification and Regression Trees*. Routledge.
- CAR (Classification and Regression)
 - Two key statistical methods in data science and AI



Birth of Regression by Galton

- Reversion – Sweet pea experiment, 1875
- Study the distribution of height between parents and children
- “*regression toward the mean*”
- The act of thinking or behaving as you did at an earlier time of your life, such as when you were a child
- Finding Causal Relation by regression

SESSION 3

Charles Darwin's Legacy in Artificial Intelligence

 **Jiyeon Kim & Youngsue Han** Hankuk University of Foreign Studies(Korea)

Regression for predictions by numbers

Classification for telling you categories

SESSION 3

Charles Darwin's Legacy in Artificial Intelligence

 Jiyeon Kim & Youngsue Han Hankuk University of Foreign Studies(Korea)

Galton's theory of regression

- Parents & Child 's height
- He charted the median values of the parent and child keys, finding that the deviation represented by the child's height was two-thirds that of the parent key deviation.
- No matter how tall his father was, his children's height was greater than the average, but they were distributed close to the average of the generation.
- "Every phenomenon tries to return to its mean"

Darwin's legacies on Galton's theory

- **"As natural selection acts solely by accumulating slight, successive, favourable variations."**
- This teaching of Darwin's theory of evolution is consistent with the law of survival of the fittest that the weak are culled.
- In fact, Darwin argued for evolution only in the realm of biology, but not Galton.

- Galton focuses only on the importance of genes.

SESSION 3

Charles Darwin's Legacy in Artificial Intelligence

 Jiyeon Kim & Youngsue Han Hankuk University of Foreign Studies(Korea)

Dark side of regression and eugenics

- Regression to the mean is not a phenomenon caused by heredity.
- Various factors

- However, Galton tried to interpret all characteristics expressed in humans as hereditary.

Conclusion

- Amongst the AI technologies that we are aware of, the majority of the predictions are conducted by regression.
- The regression comes from a combination of Plato's ideas from Ancient Greece, Charles Darwin's theory of evolution, and Galton's eugenics that blindly believed in heredity.
- Galton's use of regression in eugenics led to many generalization errors, which negatively impacted overall.
- Thus, we are using AI technology regression while considering the many different variables that occur to avoid possible errors.

SESSION 3

Charles Darwin's Legacy in Artificial Intelligence




 **Jiyeon Kim & Youngsue Han** Hankuk University of Foreign Studies(Korea)

The end

- Thank you!

SESSION 4

Moderator  Kyeongmin Kim Korea University(Korea)

| AI and Language | |
|-------------------|---|
| 16:15 ~ 16:45 KST | <p>Language Sphere: A Socio-Technical Approach to Bilingual Dictionary Creation for Low-Resource Languages</p> <p> Yohei Murakami Ritsumeikan University (Japan)</p> |
| 16:45 ~ 17:15 KST | <p>A Discussion Support System Using Neural Language Model: A Preliminary Report</p> <p> Mamoru Yoshizoe Ritsumeikan University (Japan)</p> |
| 17:15 ~ 17:45 KST | <p>Essential Attribute and Core Tasks of Computational Linguistics from the Perspective of Artificial Intelligence</p> <p> Yude Bi & Danhui Yan Fudan University(China)</p> |
| 17:45 ~ 18:15 KST | Discussion |

SESSION 4

Language Sphere: A Socio-Technical Approach to Bilingual Dictionary Creation for Low-Resource Languages

 Yohei Murakami Ritsumeikan University (Japan)

Abstract

There are more than 7000 languages around the world. However, 95% of the world population speak only 5% of them, at most 400 languages. More than half of them have fewer than 10,000 speakers. In 2010, UNESCO released a list of 2,464 endangered languages. To preserve and increase the use of those languages, we started Language Sphere project and targeted it on Indonesian languages because Indonesia has 144 endangered languages. The purpose of this project is to develop a comprehensive set of bilingual dictionaries for Indonesian ethnic languages. To this end, we propose a pivot-based bilingual lexicon induction method that combines a pair of existing dictionaries via a pivot language. Furthermore, to reduce the total cost of bilingual dictionary creation, we employed a socio-technical approach that combines the machine and manual creation processes and constructed a planner that optimizes creation orders. This presentation introduces the proposed methods and reports experiment results focusing on Indonesian, Malay, Javanese, Sundanese, Minangkabau, Palembang, and Banjarese.

SESSION 4

Language Sphere: A Socio-Technical Approach to Bilingual Dictionary Creation for Low-Resource Languages

Yohei Murakami Ritsumeikan University (Japan)



Language Sphere:

A Sociotechnical Approach to Bilingual Dictionary Creation for Low-Resource Languages

Yohei Murakami
 Ritsumeikan University
 yohei@fc.ritsumei.ac.jp



SESSION 4

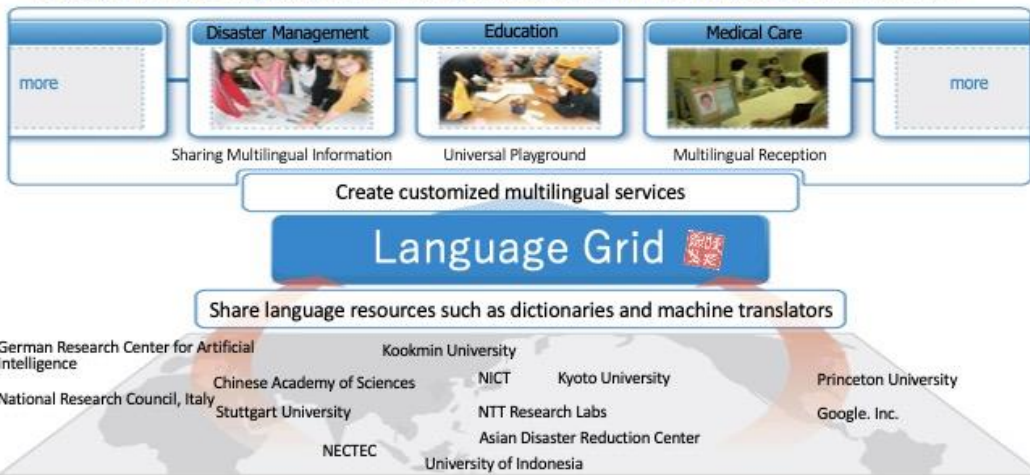
Language Sphere: A Socio-Technical Approach to Bilingual Dictionary Creation for Low-Resource Languages

Yohei Murakami Ritsumeikan University (Japan)

The Language Grid

Since 2007

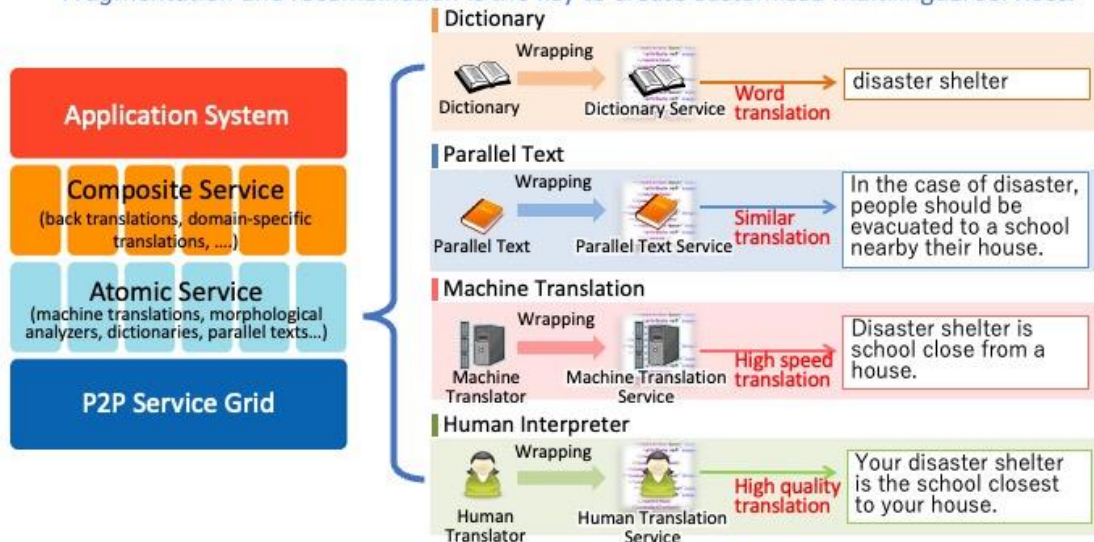
To support multi-lingual communities, institutes in the world have to collaborate.



Toru Ishida, Yohei Murakami, Donghui Lin, Takao Nakaguchi, Masayuki Otani.
Language Service Infrastructure on the Web: The Language Grid. *IEEE Computer*, Vol. 51, Issues 6, pp. 72-81, 2018.

From Language Resources to Language Services

Fragmentation and recombination is the key to create customized multilingual services.



SESSION 4

Language Sphere: A Socio-Technical Approach to Bilingual Dictionary Creation for Low-Resource Languages

Yohei Murakami Ritsumeikan University (Japan)

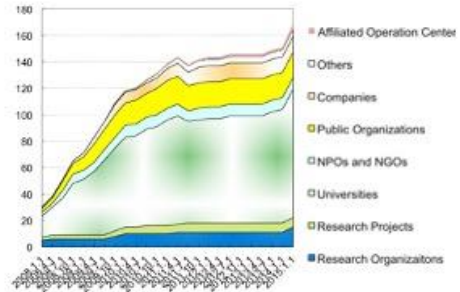
Sharing Language Services

- Contributors (24 countries, 183 groups)**

- University/Research Institute
 - Kyoto Univ. (Japan), Univ. of Indonesia, ITB (Indonesia), Shanghai Jiao Tong Univ. (China), Univ. of Stuttgart (Germany), IT Univ. of Copenhagen (Denmark), Princeton Univ. (U.S), DFKI (Germany), CNR (Italy), Chinese Academy of Sciences (China), NECTEC (Thailand), NUM (Mongolia) and more.
- NPO/NGO/Public Sector
 - NGOs for disaster reduction and intercultural activities, Junior-high schools, City Boards of Education, and more.
- Corporate (CSR activities/language resource providers)
 - NTT, Toshiba, Oki, Google, Kodensha, Translusion, and more.

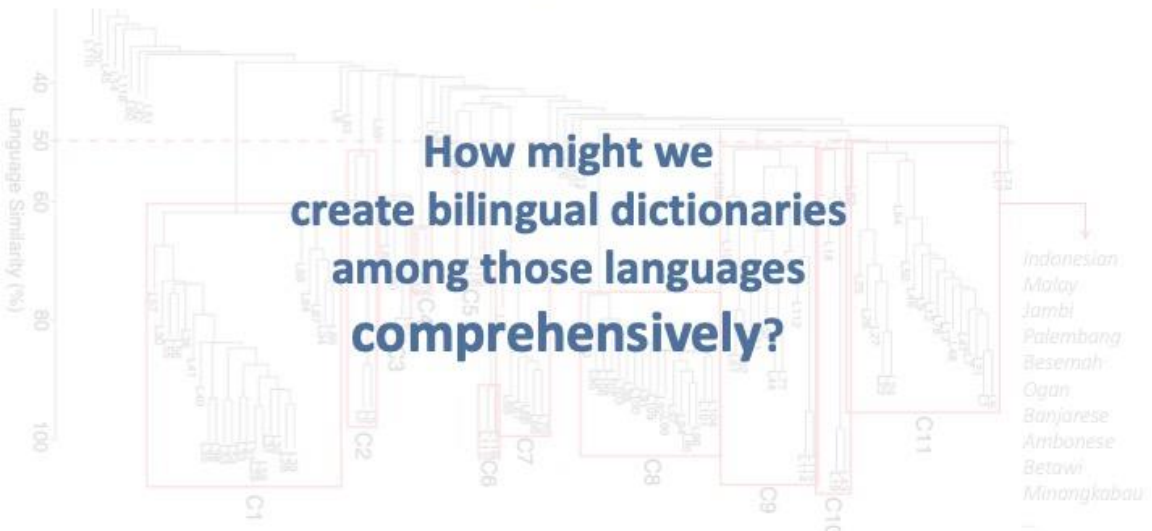
- Language Services (226 services)**

- Machine Translator
 - J-Server (ja/en/ko/zh), Web-Transer (ja/en/ko/zh/fr/de/it/es), Toshiba (en/zh), Parsit (en>th), Google Translate, and more.
- Bilingual Dictionary, Concept Dictionary
 - EDR, Wordnet, Life Science Dictionary, Multi-language Glossary on Natural Disasters, and more.
- Parallel Text
- Morphological Analyzer, Dependency Parser



Operation from December 2007

Dendrogram of the Languages in Indonesia



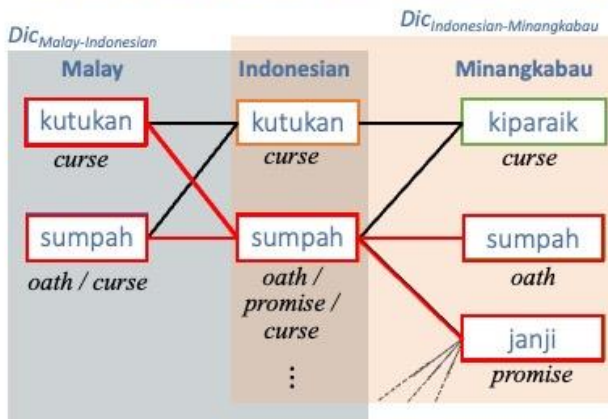
Language Similarity from ASJP Database (<https://asjp.cild.org>)

SESSION 4

Language Sphere: A Socio-Technical Approach to Bilingual Dictionary Creation for Low-Resource Languages

Yohei Murakami Ritsumeikan University (Japan)

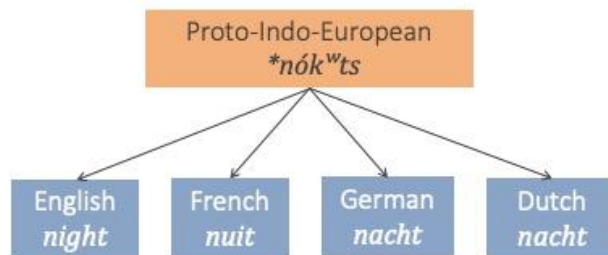
Pivot-Based Bilingual Dictionary Creation



Malay – Minangkabau Possible Translations (**correct**: 3 pairs):
 kutukan – kiparaik, sumpah – kiparaik, sumpah – sumpah,
 kutukan – sumpah, kutukan – janji, sumpah – janji

Arbi Haza Nasution, Yohei Murakami, Toru Ishida.
A Generalized Constraint Approach to Bilingual Dictionary Induction for Low-Resource Language Families.
 ACM TALLIP, Vol. 17, No. 2, Article 9, pp.1-29 2018.

Cognate Words



- > Words naturally **derived from** the same root word in **Proto-language**
- > Cognates may maintain **all the meanings of the root word**

Assumption: Correct translations should share **all the same meanings**

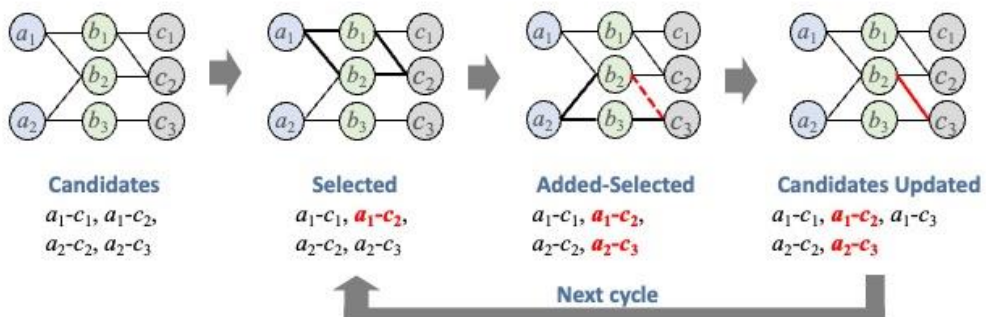
SESSION 4

Language Sphere: A Socio-Technical Approach to Bilingual Dictionary Creation for Low-Resource Languages

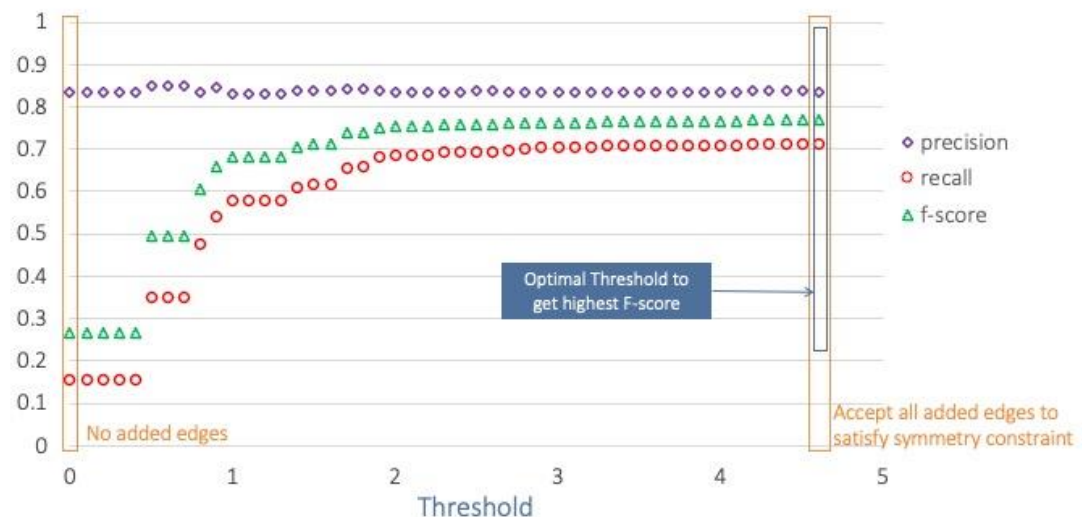
 Yohei Murakami Ritsumeikan University (Japan)

Constraint Approach to Pivot-based Bilingual Dictionary Induction

- > Formalize identification of correct translation pairs as **Weighted Max-SAT**
 - » **Symmetry constraint:** Both words have links to the same pivot words if the translation pair is correct. A symmetric topology in a graph.
 - » **Cost:** Add missing edges based on joint probability



The Effect of Threshold on Performance (1st Cycle)

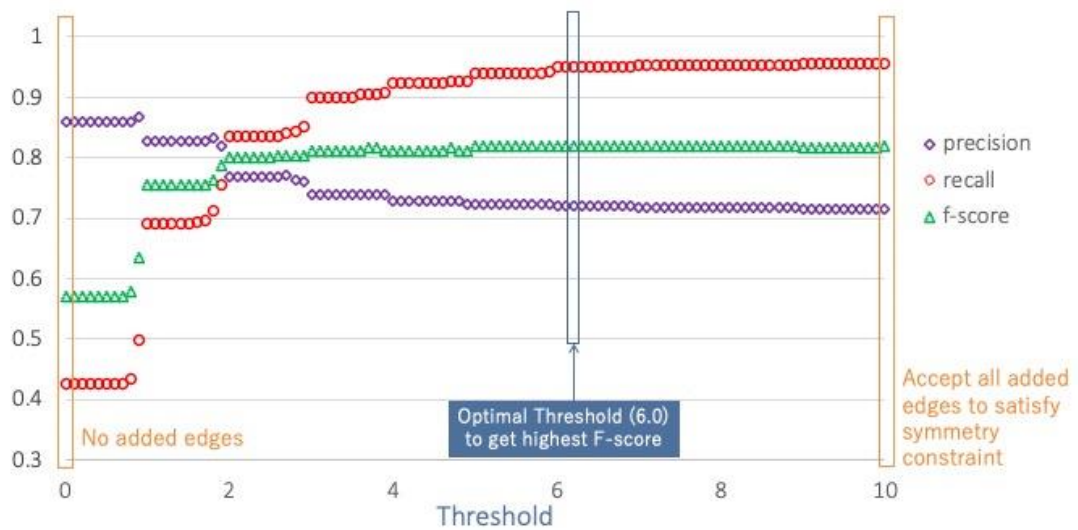


SESSION 4

Language Sphere: A Socio-Technical Approach to Bilingual Dictionary Creation for Low-Resource Languages

Yohei Murakami Ritsumeikan University (Japan)

The Effect of Threshold on Performance (2nd Cycle)



Performance Comparison with the Existing Method

| Lang. Pair | Language Similarity | Result | 1-1 pair | No threshold | Threshold (6.0) | Existing Method |
|------------|---------------------|-------------|--------------|--------------|-----------------|-----------------|
| min-ind | 69.14% | Translation | 407 | 1244 | 1179 | 38 |
| zlm-ind | 87.70% | Precision | 87.3% | 65.4% | 71.8% | 95% |
| min-zlm | 61.66% | Recall | 32.7% | 99.8% | 94.7% | 3.1% |
| | | F-score | 47.5% | 79.1% | 81.7% | 5.9% |

> Apply the constraint approach to creating a bilingual dictionary between **Malay and Minangkabau via Bahasa Indonesia**

> Introducing uniqueness constraint (1-1 pair) achieves the highest precision (87.3%), but lowest recall (32.7%)

> Vice versa, no threshold

> With **2nd cycle** and **threshold (6.0)**, balance precision and recall (highest F-score: **81.7%**)

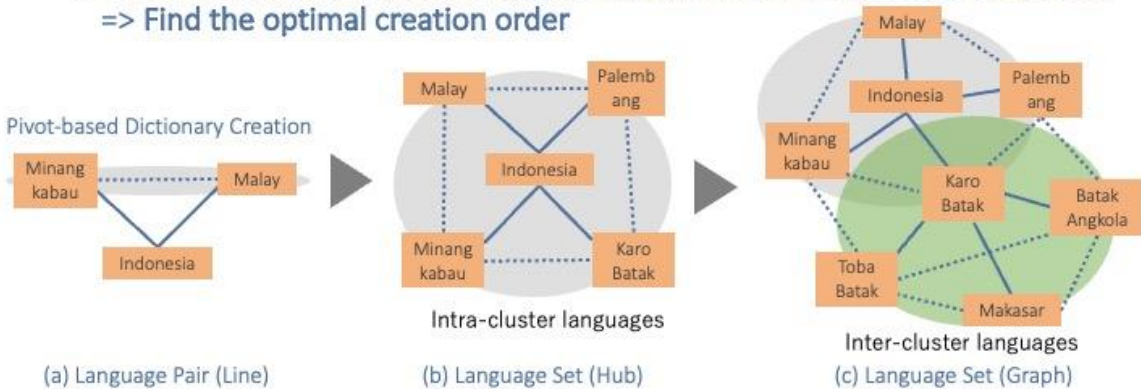
SESSION 4

Language Sphere: A Socio-Technical Approach to Bilingual Dictionary Creation for Low-Resource Languages

Yohei Murakami Ritsumeikan University (Japan)

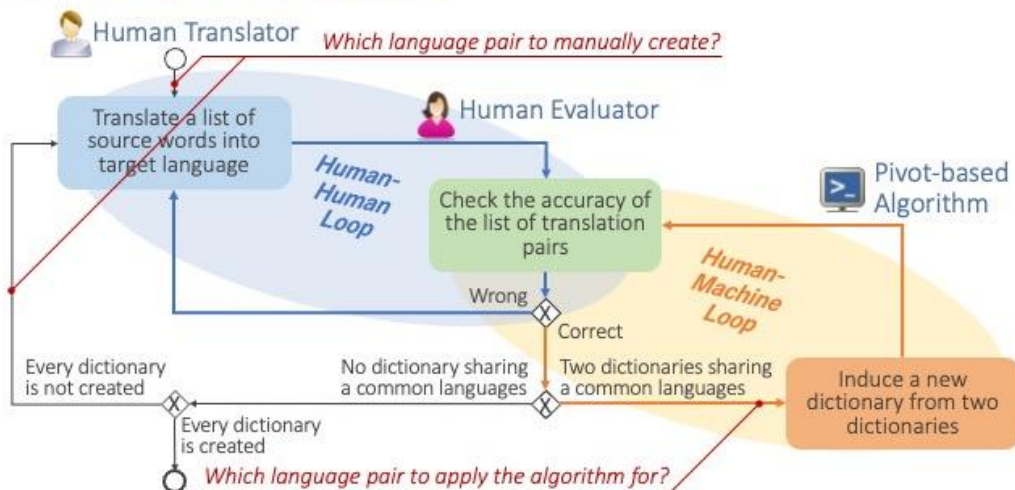
From Language Pair to Language Set

- > Enough existing dictionaries?
=> Need human involvement
- > Which language pair has the biggest impact on other dictionary creations?
=> Find the optimal creation order



Sociotechnical Approach to Bilingual Dictionary Creation

- > Human-Machine collaboration



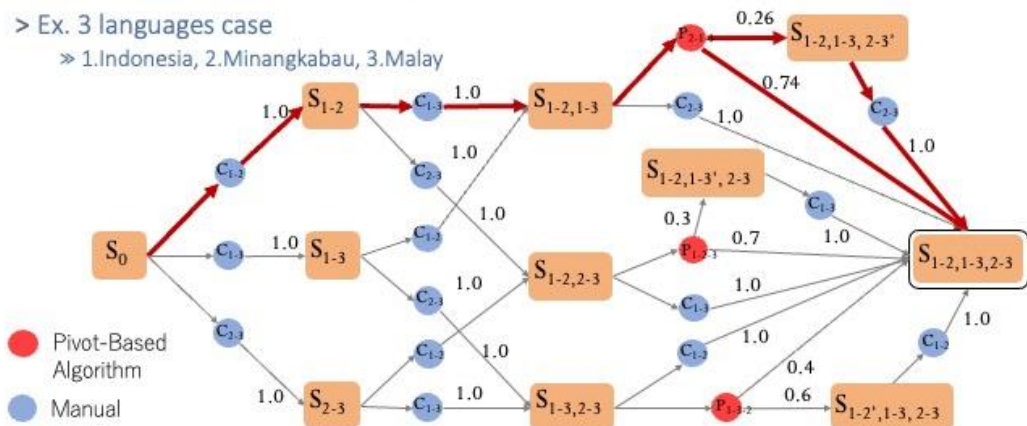
SESSION 4

Language Sphere: A Socio-Technical Approach to Bilingual Dictionary Creation for Low-Resource Languages

Yohei Murakami Ritsumeikan University (Japan)

Plan Optimization with Markov Decision Process

- > Find the optimal path minimizing the total cost
- > Ex. 3 languages case
 - » 1.Indonesia, 2.Minangkabau, 3.Malay



Arbi Haza Nasution, Yohei Murakami, Toru Ishida. Plan Optimization to Bilingual Dictionary Induction for Low-Resource Language Families. ACM TALLIP, Vol. 20, No. 2, Article 29, pp.1-28 2020.

1st Indonesia Language Sphere Experiment (2018)



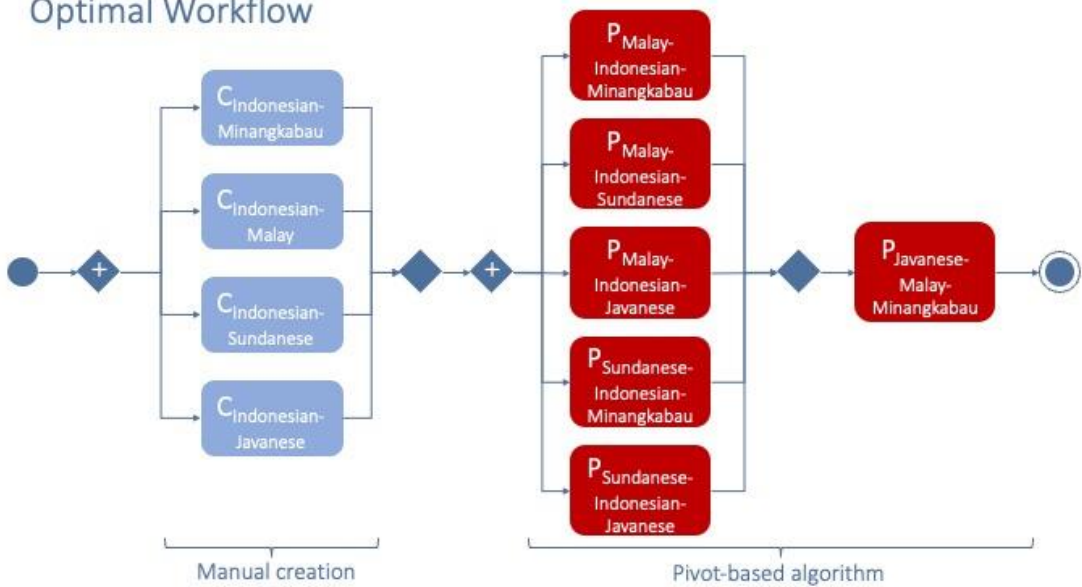
Similarity Clusters Map of 32 Indonesian Ethnic Languages – 5 Clusters

SESSION 4

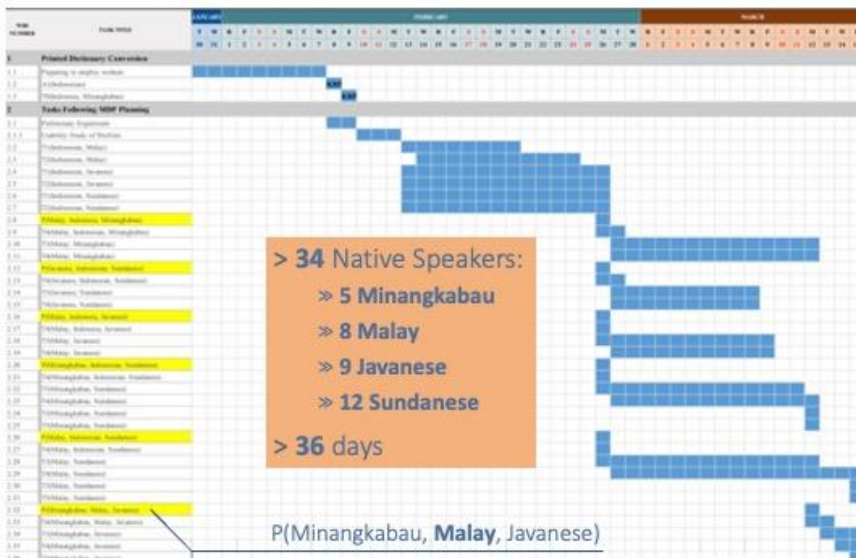
Language Sphere: A Socio-Technical Approach to Bilingual Dictionary Creation for Low-Resource Languages

 Yohei Murakami Ritsumeikan University (Japan)

Optimal Workflow



1st Experiment (2018)

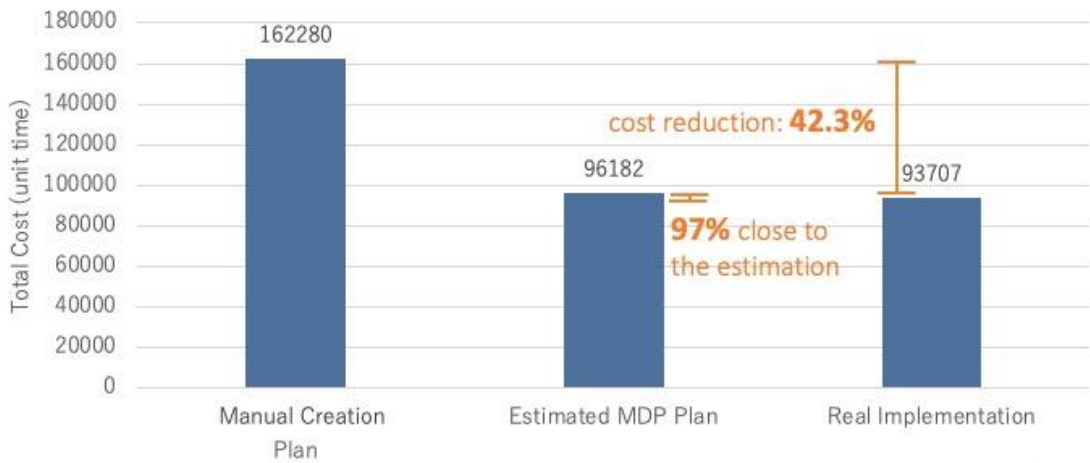


SESSION 4

Language Sphere: A Socio-Technical Approach to Bilingual Dictionary Creation for Low-Resource Languages

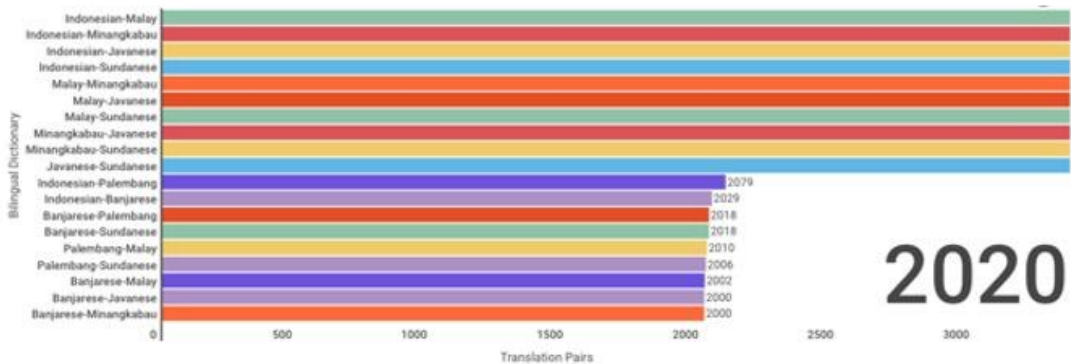
Yohei Murakami Ritsumeikan University (Japan)

Experiment Result



19

Current Progress



- > 21 bilingual dictionaries including at least 2,000 translation pairs
- > 7 languages: Indonesian, Malay, Minangkabau, Palembang, Banjarese, Sundanese, Javanese

SESSION 4

Language Sphere: A Socio-Technical Approach to Bilingual Dictionary Creation for Low-Resource Languages

 Yohei Murakami Ritsumeikan University (Japan)

Conclusions

> GOAL

- » Comprehensively create bilingual dictionaries among low-resource languages belonging to the same language family

> TECHNOLOGY

- » Pivot-based bilingual dictionary induction
 - Create a new bilingual dictionary from two existing bilingual dictionaries sharing a common language
 - Extract translation pairs by solving a weighted max SAT
 - » Plan optimization
 - Decide an order of dictionary creations including manual creation and pivot-based creation
 - Find an optimal policy by solving a MDP
-

APPENDIX

SESSION 4

Language Sphere: A Socio-Technical Approach to Bilingual Dictionary Creation for Low-Resource Languages

Yohei Murakami Ritsumeikan University (Japan)

Distribution of World Languages

| Languages | Living languages | | Number of speakers | | | |
|--------------------|------------------|--------------|----------------------|--------------|------------------|---------------|
| | Count | Percent | Total | Percent | Mean | Median |
| African languages | 2,138 | 30.1% | 815,252,580 | 13.0% | 381,316 | 27,500 |
| American languages | 1,064 | 15.0% | 51,527,759 | 0.8% | 48,428 | 1,160 |
| Asian languages | 2,301 | 32.4% | 3,779,634,812 | 60.1% | 1,642,605 | 12,000 |
| European languages | 286 | 4.0% | 1,637,993,977 | 26.0% | 5,727,252 | 35,600 |
| Pacific languages | 1,313 | 18.5% | 6,783,496 | 0.1% | 5,166 | 950 |
| Total | 7,102 | 100.0% | 6,291,192,624 | 100.0% | 885,834 | 7,000 |

> **880** of Asian languages are endangered

» **528** south-east Asian languages

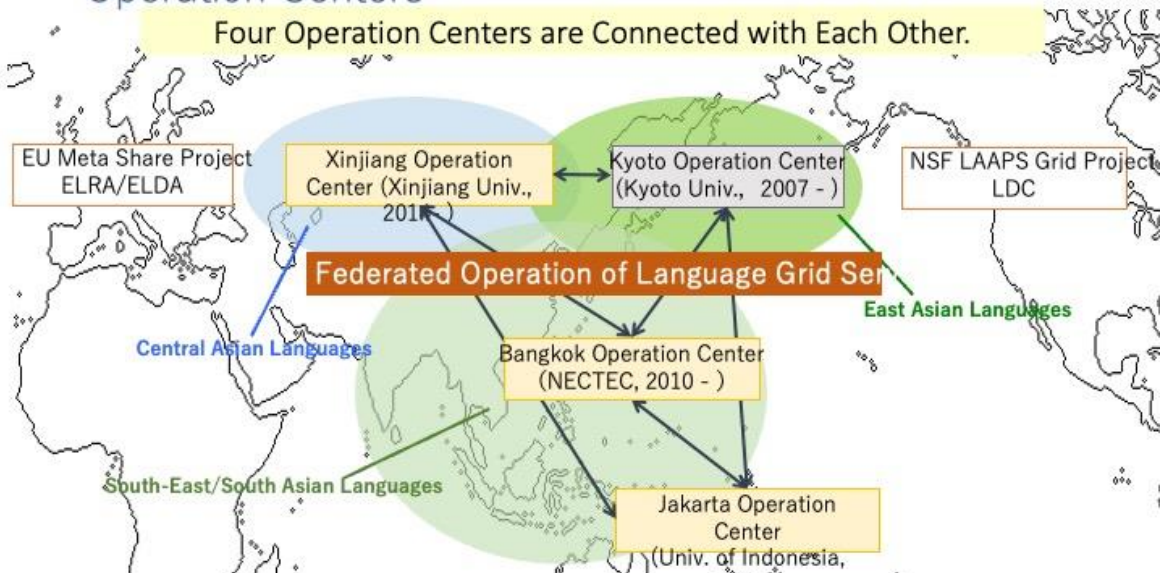
» **341** languages of Indonesia

» **148** Indonesian languages are classified into endangered languages by UNESCO

(<http://www.ethnologue.com/>)
 (<http://www.unesco.org/languages-atlas/>)

Operation Centers

Four Operation Centers are Connected with Each Other.



SESSION 4

Language Sphere: A Socio-Technical Approach to Bilingual Dictionary Creation for Low-Resource Languages

 Yohei Murakami Ritsumeikan University (Japan)

Lexicostatistical Matrix of Languages in Indonesia (%)

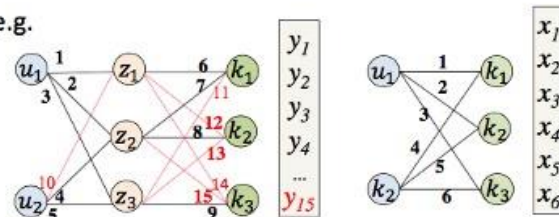
| | Indonesian | Malang | Yogyakarta | Old Javanese | Sundanese | Malay | Palembang Malay | Madurese | Minangkabau | Buginese |
|-----------------|--------------|--------------|--------------|--------------|-----------|--------------|-----------------|----------|-------------|----------|
| Malang | 23.46 | | | | | | | | | |
| Yogyakarta | 27.29 | 87.36 | | | | | | | | |
| Old Javanese | 24.09 | 47.5 | 52.18 | | | | | | | |
| Sundanese | 39.43 | 18.55 | 22.43 | 21.82 | | | | | | |
| Malay | 85.10 | 20.53 | 24.35 | 21.36 | 41.12 | | | | | |
| Palembang Malay | 68.24 | 33.97 | 37.97 | 31.85 | 38.9 | 73.23 | | | | |
| Madurese | 34.45 | 17.63 | 14.15 | 15.18 | 19.86 | 34.16 | 34.32 | | | |
| Minangkabau | 61.59 | 26.59 | 29.63 | 25.01 | 30.81 | 61.66 | 63.6 | 34.32 | | |
| Buginese | 31.21 | 12.76 | 16.85 | 18.33 | 24.8 | 32.04 | 31 | 17.94 | 32 | |

Max-SAT-Based Formalization

> Boolean variables

- » x : a word pair (u_i, k_j)
 =TRUE: is a translation pair
 =FALSE: otherwise
- » y : an edge
 =TRUE: exists
 =FALSE: otherwise

e.g.



> Clauses

- » C1 (Hard): Existing edges cannot be deleted e.g. (c, ∞) : y_1
- » C2 (Soft): Missing edges can be added e.g. (c, ω_{10}) : $\neg y_{10}$
- » C3 (Hard): A translation pair must have same meaning(s)
 e.g. (c, ∞) : $x_1 \rightarrow y_1 \wedge y_2 \wedge y_3 \wedge y_6 \wedge y_7 \wedge y_{11}$

SESSION 4

Language Sphere: A Socio-Technical Approach to Bilingual Dictionary Creation for Low-Resource Languages

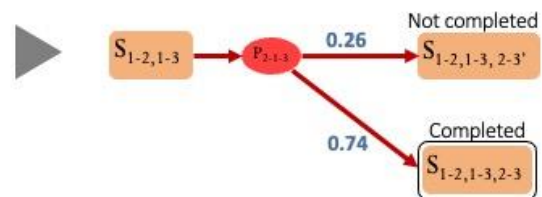
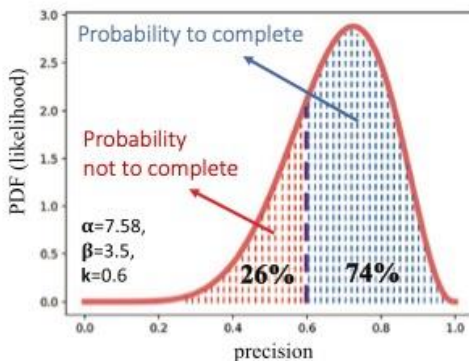
Yohei Murakami Ritsumeikan University (Japan)

Performance Comparison with the Existing Method

| Lang. Pair | Language Similarity | Result | 1-1 pair | No threshold | Threshold (6.0) | Existing Method |
|------------|---------------------|-------------|--------------|--------------|-----------------|-----------------|
| min-ind | 69.14% | Translation | 407 | 1244 | 1179 | 38 |
| zlm-ind | 87.70% | Precision | 87.3% | 65.4% | 71.8% | 95% |
| min-zlm | 61.66% | Recall | 32.7% | 99.8% | 94.7% | 3.1% |
| | | F-score | 47.5% | 79.1% | 81.7% | 5.9% |

- > Apply the constraint approach to creating a bilingual dictionary between **Malay and Minangkabau via Bahasa Indonesia**
- > Introducing uniqueness constraint (1-1 pair) achieves the highest precision (87.3%), but lowest recall (32.7%)
- > Vice versa, no threshold

State Transition Probability



- > Model the probability with beta distribution
 - >> α : Language similarity
 - >> β : Ambiguity of pivot words
 - >> k : Necessary precision to complete a dictionary

SESSION 4

Language Sphere: A Socio-Technical Approach to Bilingual Dictionary Creation for Low-Resource Languages

 Yohei Murakami Ritsumeikan University (Japan)

Crowdsourcing System for Dictionary Creation

Pembuatan Kamus Dwibahasa Indonesian-Javanese

| Indonesian | Javanese |
|---|----------|
| Silakan terjemahkan acara ke maksimum 5 kata | |
| acara | |
| acara | |
| acara | |
| acara | |
| acara | |
| Silakan terjemahkan adik ke maksimum 5 kata | |
| Silakan terjemahkan agen ke maksimum 5 kata | |
| Silakan terjemahkan ahli ke maksimum 5 kata | |
| Silakan terjemahkan air ke maksimum 5 kata | |

Hal 1 dari 1 Menampilkan 1 - 25 dari 25

Creation Task

Evaluasi Kamus Dwibahasa Malay-Javanese

| Malay | Javanese | Evaluasi | Hasil |
|---------|---------------|----------|---------|
| Acara | Ngadiah damel | | CORRECT |
| Adik | Dukur kandung | | CORRECT |
| Agen | Bandar | | CORRECT |
| Ahli | Pinter | | CORRECT |
| Air | Banyu | | CORRECT |
| Hajatan | Hajatan | | CORRECT |
| Pakar | Linuwih | | WRONG |
| Tukang | Tukang | | CORRECT |

Hal 1 dari 1 Menampilkan 1 - 8 dari 8

2018-02-11 20:37:17 | Cornelia Regina Sinta Maharani: boieh sih
 2018-02-11 20:37:54 | Mhd safri: ahli-pinter kek wrong kak?
 2018-02-11 20:38:00 | Cornelia Regina Sinta Maharani: air = banyu
 2018-02-11 20:38:08 | Cornelia Regina Sinta Maharani: eh salah pencett
 2018-02-11 20:38:39 | Cornelia Regina Sinta Maharani: gmn?
 2018-02-11 20:38:47 | Mhd safri: udah kak
 2018-02-11 20:38:47 | Cornelia Regina Sinta Maharani: udah berubah?
 2018-02-11 20:39:31 | Mhd safri: apa lagi ya??

Ketik pesan...

Evaluation Task

Discussion: Sustainability of Language Resource Creation

- > Monetization of language resources in low-resourced languages
 - » Currently conduct experiments by research funding
 - » Who can pay it?
 - Government/Company/Community/End users
 - » Charge for what?
 - Look up a word/Buy a dictionary/Distribute a dictionary/Learn a model from a dictionary
- > Incentive design for workers
 - » Economic incentive: How to distribute the charge to workers?
 - » Social Incentive: What is regarded as honor?
 - » Moral Incentive: What leads workers to right thing to do?

SESSION 4

A Discussion Support System Using Neural Language

Model: A Preliminary Report

 Mamoru Yoshizoe Ritsumeikan University (Japan)

Abstract

The Internet has connected people over the whole of the world on the physical layer. The huge amount of data packets go around the globe every second. We believed that knowledge-sharing on the Internet enables us to promote understanding each other. However, we are now faced with the wall of the diversity of values.

We are still far from mutual understandings so that there are many conflicts in the physical and virtual spaces. For example, "Flaming" problems have been popular phenomena. It is not special for us to see that posted ideas and thoughts on SNS such as Twitter are frequently criticized and bashed by many people without any rational reasons. Additionally, with the development of services on the Internet, we tend to be separated from information that differs from our perspectives. We, on the other hand, tend to deal with information that is familiar with our interests or our belonging communities. Consequently, each of us is becoming isolated in homogeneous cultural and philosophical films (bubbles). This phenomenon has been called "filter bubble" problem.

These problems mean that the possibility of people losing the opportunity to know diverse values and not being able to think things based on different viewpoints. In order to prevent these problems, it is important to consider or respect the values of other people though, it is not easy to sense them since we tend to think within our scope of knowledge, experience, and imagination.

Our research team has discussed the approach that gives the awareness of values that people have not recognize yet, and we have developed the system on how to exchange values and achieve a synergetic effect among people.

In this presentation, we show our prototype system, called AIR-VAS, aiming to support becoming aware of values in group discussion. AIR-VAS has been developed

SESSION 4

A Discussion Support System Using Neural Language Model: A Preliminary Report

👤 Mamoru Yoshizoe Ritsumeikan University (Japan)

as the system which recognizes characteristic opinions of a group and shares them among all engaging groups on the discussion. The recognized and shared opinions are based on the values of the people. Through the sharing of opinions, people can know the different viewpoints on the issue of the current discussion, so that AIR-VAS can provide stimulation to people for idea generation.

The AIR-VAS consists of two awareness structures, the agent-based information suggestion structure and the visualization structure of the discussion word network. These structures are realized on the processing by word co-occurrence network and the word embeddings. The system receives discussion text information by the speech recognition and acquires the word co-occurrence relationship in the sentence and a word distributed representation by the attention-based language model. Based on them, the system visualizes the discussion state on the user's terminal screen in realtime, and the system recognizes the difference of each group's discussion and suggests users another perspective regularly.

According to the experimental usage of the developed system, we analyzed the relationship between discussion and visualized information, and discuss what is effective information that gives awareness.

SESSION 4

A Discussion Support System Using Neural Language Model: A Preliminary Report

 Mamoru Yoshizoe Ritsumeikan University (Japan)

ICAIH 2021

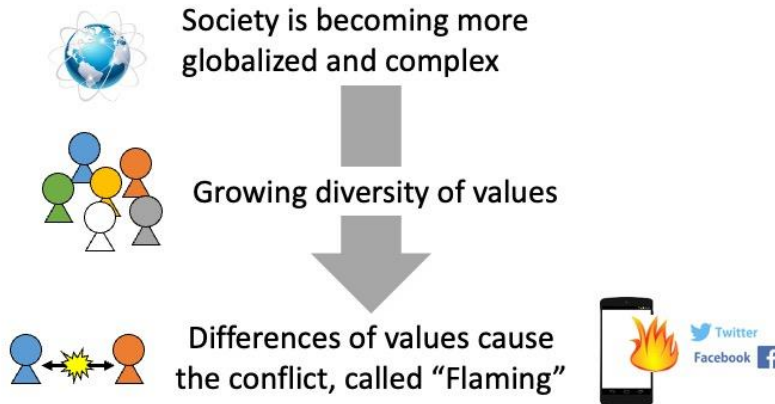
A Discussion Support System Using Neural Language Model: A Preliminary Report

Mamoru Yoshizoe

Ritsumeikan University, Japan
Graduate School of Information Science and Engineering

1

Background



It is essential to have broad perspectives and be aware of diverse values

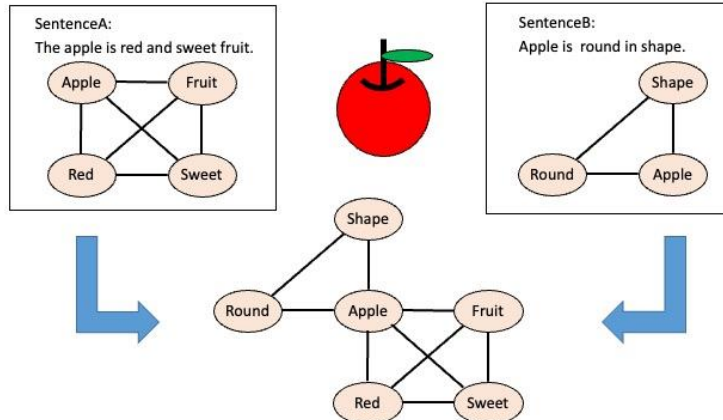
SESSION 4

A Discussion Support System Using Neural Language Model: A Preliminary Report

Mamoru Yoshizoe Ritsumeikan University (Japan)

The word co-occurrence network

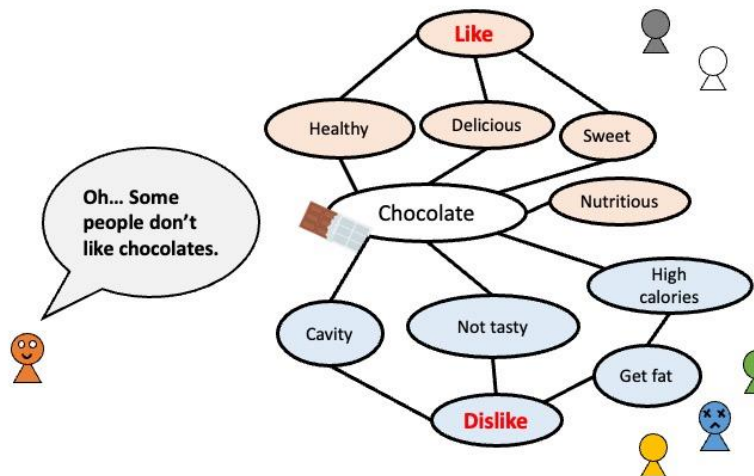
The co-occurrence network^{*1,2} is a visualizing method based on the word's co-occurrence and frequency.



*1 Osgood, C. E., 1959, "The Representational Model and Relevant Research Methods," I. d. S. Pool ed., Trends in Content Analysis, Urbana, IL: University of Illinois Press, 33-88
*2 Danowski, J. A., 1993, "Network Analysis of Message Content", W. D. Richards Jr. & G. A. Barnett eds., Progress in Communication Sciences IV, Norwood, NJ: Ablex, 197-221

5

Suggestions from other networks



We recognize other's value from the co-occurrence network 6

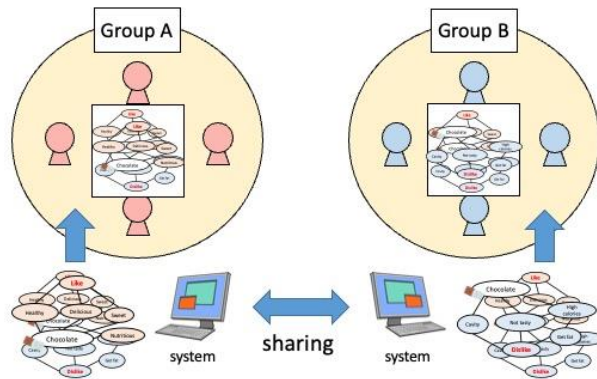
SESSION 4

A Discussion Support System Using Neural Language Model: A Preliminary Report

👤 Mamoru Yoshizoe Ritsumeikan University (Japan)

Discussion support by network sharing

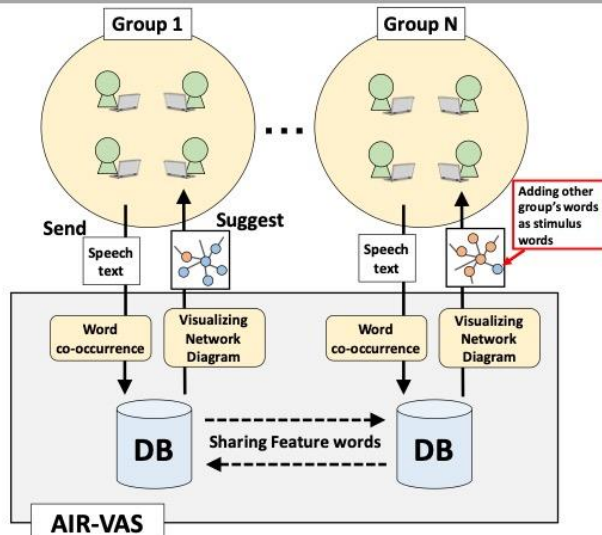
- Group discussion support by the network visualizing system



Other group information can work as stimuli or clues to make us aware of the different discussion routes.

7

The AIR-VAS system



8

SESSION 4

A Discussion Support System Using Neural Language Model: A Preliminary Report

Mamoru Yoshizoe Ritsumeikan University (Japan)

Discussion environment



fig1: The discussion environment

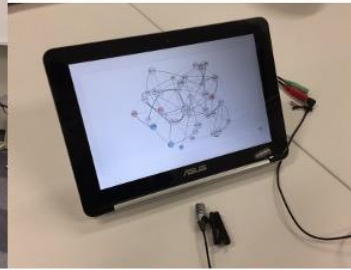


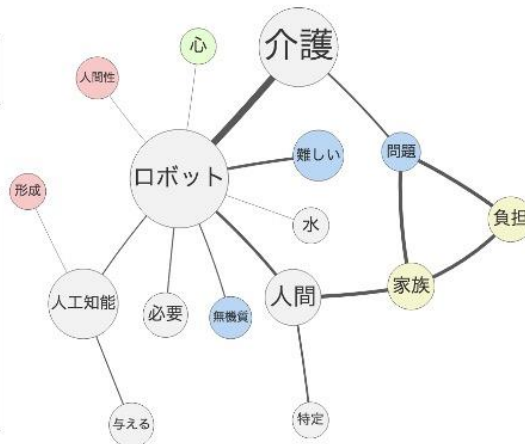
fig2: The tablet terminal

9

Example of word co-occurrence network

Discussion theme : 「AI and elder care」

- = normal word
- = positive word
- = negative word
- = stimulus word
- = new added word



* In japan, red color has often positive feeling, blue color has negative feeling.

The color is decided by Japanese Sentiment Dictionary, (Higashiyama 2008), (Kobayashi 2005)

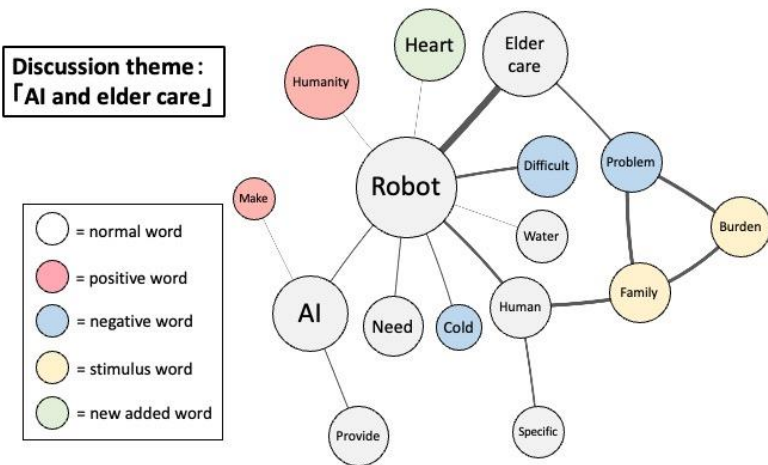
10

SESSION 4

A Discussion Support System Using Neural Language Model: A Preliminary Report

Mamoru Yoshizoe Ritsumeikan University (Japan)

Example of word co-occurrence network

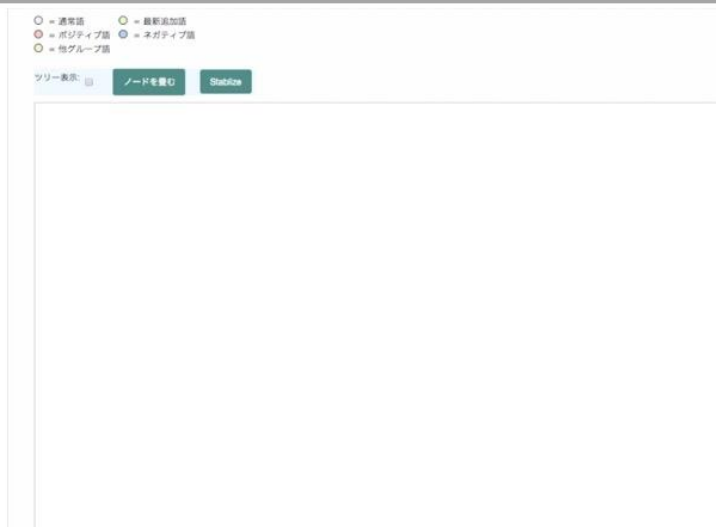


* In japan, red color has often positive feeling, blue color has negative feeling.

This is English translation sample image. The system supports only Japanese now.

11

AIR-VAS - screen example



12

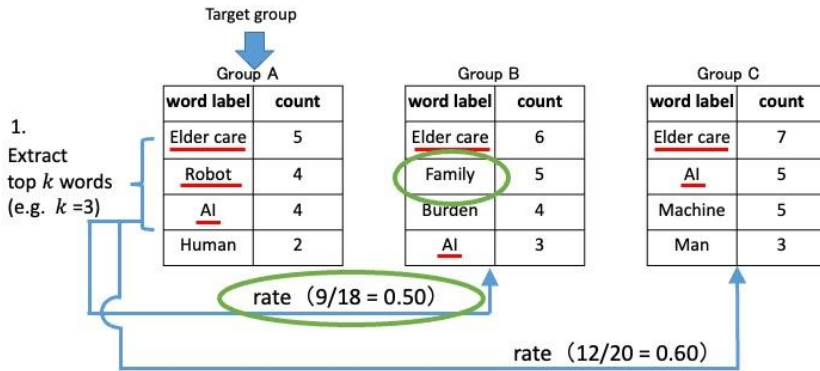
SESSION 4

A Discussion Support System Using Neural Language Model: A Preliminary Report

 Mamoru Yoshizoe Ritsumeikan University (Japan)

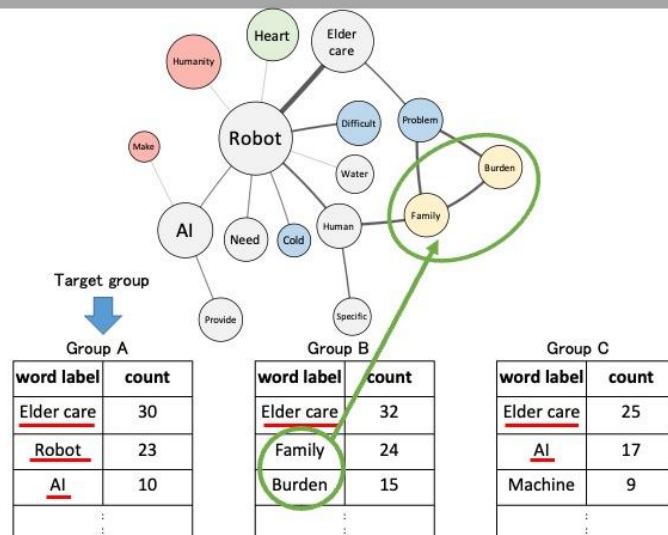
Selection of stimulus words

3. Select the frequent word which is not appeared in the target group, from the group which has minimum rate.



2. Calculate the rate of red line words to sum count in the group.

Selection of stimulus words



14

SESSION 4

A Discussion Support System Using Neural Language Model: A Preliminary Report

Mamoru Yoshizoe Ritsumeikan University (Japan)

Experiment

● Experiment of group discussion with AIR-VAS

5 groups (a group has 4 members) discuss 2 themes.

- 3 groups : The group which is suggested stimulus words (= group a)
- 2 groups : The group which is not suggested stimulus words (= group b)

Discussion theme 1:
「AI system that prevent the school bullying」

Discussion theme 2:
「AI and National security」



We evaluate our system in that how stimulus words affect the variety of the discussion.

15

Experimental results

Stimulus words activate the discussion

Some words are translated to English from Japanese in this image.

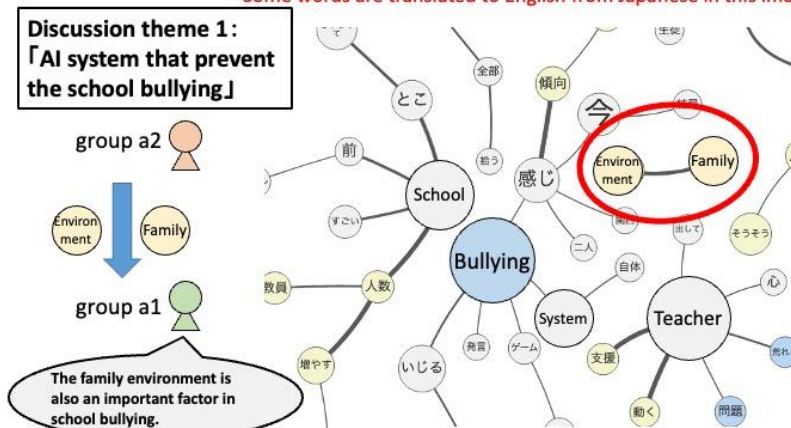


Fig. 3 The part of network diagram : the discussion theme 1 - group a3.

16

SESSION 4

A Discussion Support System Using Neural Language Model: A Preliminary Report

Mamoru Yoshizoe Ritsumeikan University (Japan)

Limitations of the system

① Limitation of stimulus words selection

➤ It's not always the best way that is based on the word frequency.

② Limitation of the suggestion by screen visualization

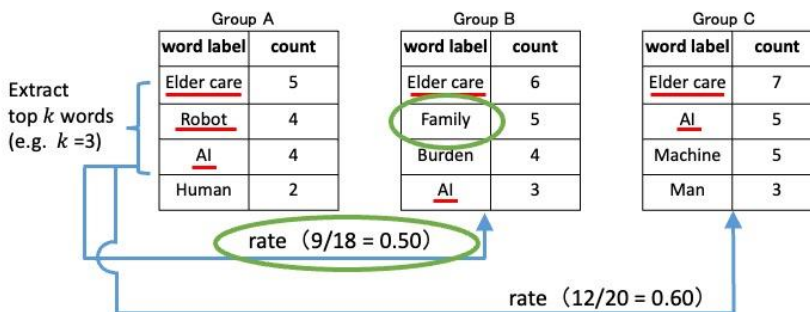
➤ People tend not to look at the system screen when concentrating on the discussion speech.

19

Limitations of the system

① Limitation of stimulus words selection

➤ It's not always the best way that is based on the word frequency.



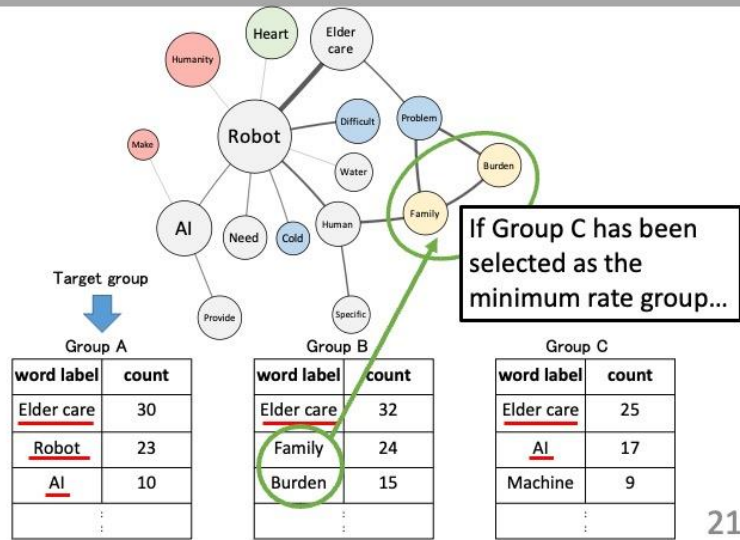
20

SESSION 4

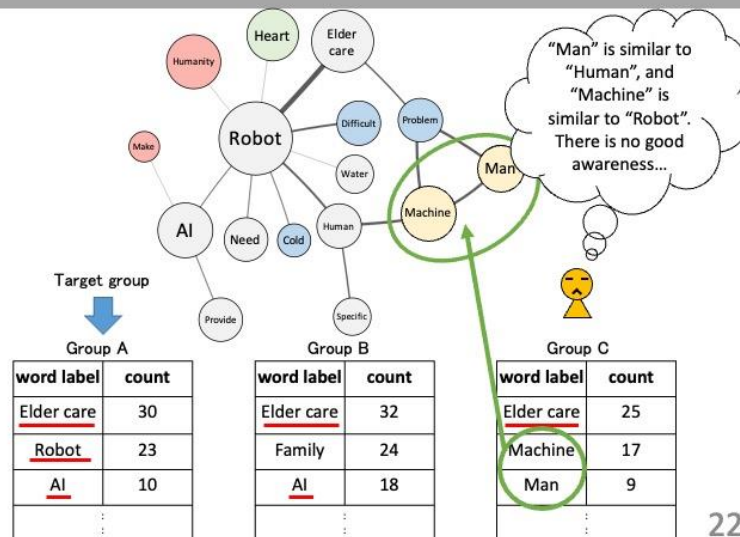
A Discussion Support System Using Neural Language Model: A Preliminary Report

Mamoru Yoshizoe Ritsumeikan University (Japan)

Limitation of stimulus words selection



Limitation of stimulus words selection



SESSION 4

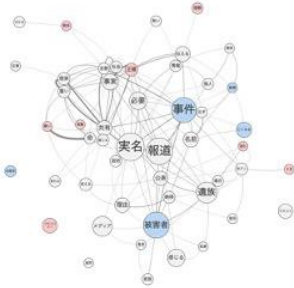
A Discussion Support System Using Neural Language Model: A Preliminary Report

Mamoru Yoshizoe Ritsumeikan University (Japan)

Limitations of the system

② Limitation of the suggestion by screen visualization

- People tend not to look at the system screen when concentrating on the discussion speech.



23

Improvement of other word's selection

Using neural attention language models such as BERT*1, we use

- word embedding
- word attentions

and visualize the distance of discussion sentences.



e.g. BERT word embedding +PCA

24

*1 Devlin, J., Chang, M., Lee, K., and Toutanova, K.: BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding, in NAACL, pp. 4171-4186 (2019)

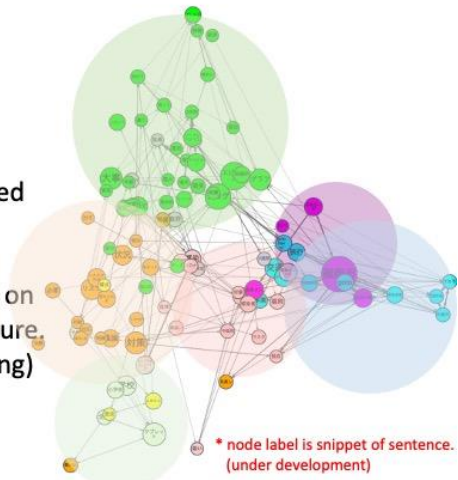
SESSION 4

A Discussion Support System Using Neural Language Model: A Preliminary Report

 Mamoru Yoshizoe Ritsumeikan University (Japan)

e.g. the word attention network

- The node position is set by word or sentence embedding.
- The network edges are visualized based on the attention values.
- We can cluster the node based on embedding and network structure. (e.g. G-means, Louvain clustering)

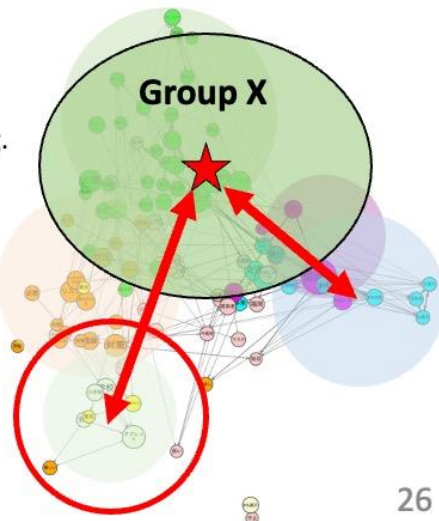


The discussion sample text from web :
preventative measures against Covid-19

25

Selection by the word embedding distance

1. Calculate the average of Group X 's sentence embedding.
2. Calculate the distance from average point to other group's cluster.
3. Select the cluster which has maximum distance.



26

SESSION 4

A Discussion Support System Using Neural Language Model: A Preliminary Report

Mamoru Yoshizoe Ritsumeikan University (Japan)

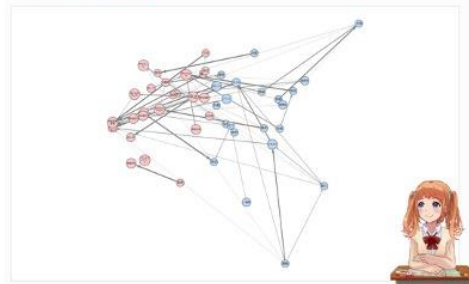
Improvement of suggestion form

- Implementation the communication robot in the physical space.
- The character agent tells other group's information.

Using the other way of suggestion form, such as voice recommendation.



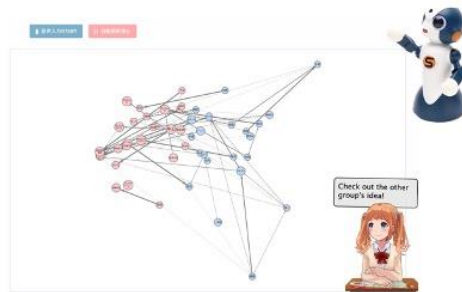
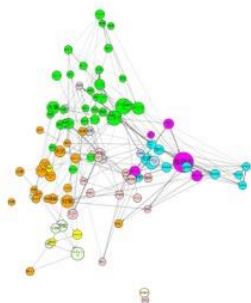
Fig.7 The communication robot



* under development Fig.8 The character agent

Future work

- Experiment using the neural language models
- Implementation of the agent-based suggestion from discussion contents with text generation models.



SESSION 4

Essential Attribute and Core Tasks of Computational Linguistics from the Perspective of Artificial Intelligence

 Yude Bi & Danhui Yan Fudan University(China)

Abstract

In the era of artificial intelligence, natural language understanding (NLU) has been given a very high standing and thus has ushered in her shining moment. Bill Gates of Microsoft once said that "language understanding is the jewel in the crown of artificial intelligence".

Artificial intelligence is to use computer to simulate and realize human intelligence, and human's intellectuality on the whole can be divided into three levels: the first level is the computing intelligence; the second one is intelligent perception; the third level is the cognitive intelligence that includes the ability to understand and use languages, to master and apply knowledge, and to draw inferences on the basis of language and knowledge.

NLU is at the core of cognitive intelligence, since only through NLU, can artificial intelligence achieve in the real sense.

The popularity of Deep Learning (DL) has given a powerful impetus to the development of artificial intelligence today. However, some scholars believe that DL is a black box and lacks interpretability. It is essentially statistical, but cannot be reasoned from data. Therefore, how to understand NLU is the philosophical problem hidden behind DL. It is necessary for us to make clear its disciplinary attributes and connotation from the theoretical point of view. Only in this way can researchers grasp the research direction, transcend the existing research frameworks and paradigms, and have a sound grip of AI with a broader vision.

Computational Linguistics (CL) and Natural Language Processing (NLP) were initially considered as two sides of the same coin, but with the development of technology and methods, NLP has gradually replaced CL. As NLU is the goal of cognitive intelligence, it is necessary to clarify the relationship between NLP and CL.

SESSION 4

Essential Attribute and Core Tasks of Computational Linguistics from the Perspective of Artificial Intelligence

 **Yude Bi & Danhui Yan** Fudan University(China)

CL is the nature of NLU, while NLP is the technical means. We believe that language is the object of the study and computation is the means to carry out the study. Computation acting on language helps to explore the rules of language and understanding mechanism, and to reveal the secrets of language, thus further boosting the development of artificial intelligence.

SESSION 4

Essential Attribute and Core Tasks of Computational Linguistics from the Perspective of Artificial Intelligence

👤 Yude Bi & Danhui Yan Fudan University(China)



Yude Bi & Danhui Yan

Contents

1. Background: The core issue of Artificial Intelligence
2. Natural Language Understanding
3. Computational Linguistics and Natural Language Processing
4. Essential attributes and core tasks of computational linguistics.
5. Conclusion

SESSION 4

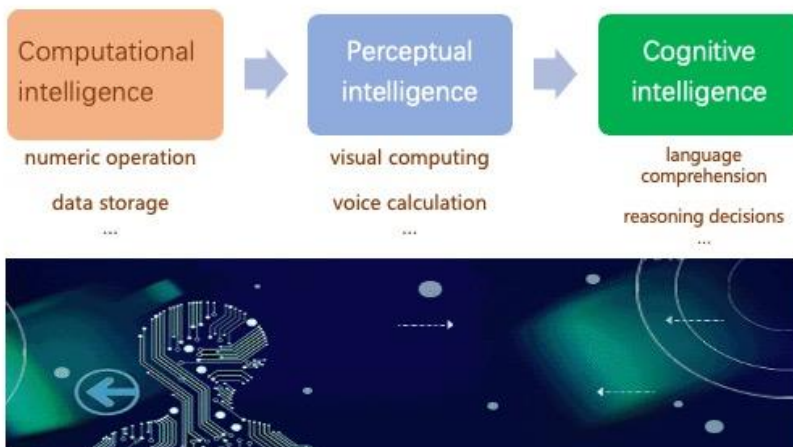
Essential Attribute and Core Tasks of Computational Linguistics from the Perspective of Artificial Intelligence

Yude Bi & Danhui Yan Fudan University(China)



1. Background The core issue of Artificial Intelligence

Artificial Intelligence: From computing to cognition



SESSION 4

Essential Attribute and Core Tasks of Computational Linguistics from the Perspective of Artificial Intelligence

Yude Bi & Danhui Yan Fudan University(China)

Cognition and its Levels

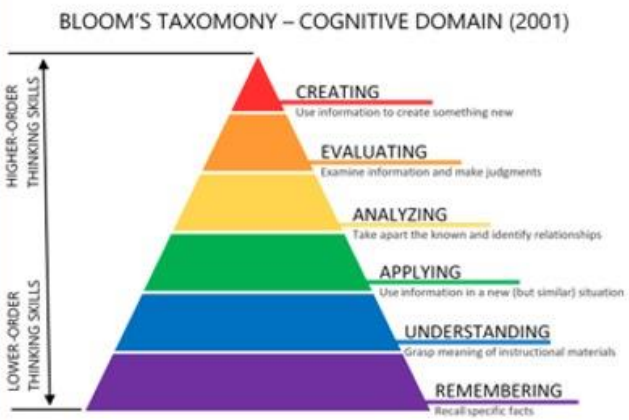
Cognition refers to the process of acquiring knowledge or applying knowledge, or information processing, which is the most basic psychological process of human beings.

Cognition includes feeling, perception, memory, thinking, imagination and language. The human brain accepts the information from the outside world; such information, processed by the brain, is transformed into internal psychological activities, and then dominates human behavior. This procedure is information processing, which is also called cognitive process. People’s cognitive ability is closely related to their cognitive process. It can be said that cognition is a product of people's cognitive process.

Cognition and its Levels

Bloom’s Taxonomy was created by Benjamin Bloom in 1956, published as a kind of classification of learning outcomes and objectives.

The original sequence of cognitive skills was Knowledge, Comprehension, Application, Analysis, Synthesis, and Evaluation. The framework was revised in 2001 by Lorin Anderson and David Krathwohl, yielding the revised Bloom’s Taxonomy. The most significant change to the Cognitive Domain was the removal of ‘Synthesis’ and the addition of ‘Creation’ as the highest-level of Bloom’s Taxonomy. And being at the highest level, the implication is that it’s **the most complex or demanding cognitive skill**—or at least represents a kind of pinnacle for cognitive tasks.



SESSION 4

Essential Attribute and Core Tasks of Computational Linguistics from the Perspective of Artificial Intelligence

👤 Yude Bi & Danhui Yan Fudan University(China)

Linguistic Intelligence

Cognitive intelligence: the ability to understand and use languages, to master and apply knowledge, and to draw inferences on the basis of language and knowledge.

Cognitive intelligence is mainly embodied in **linguistic intelligence**, since only through NLU (Natural Language Understanding), can artificial intelligence achieve in the real sense.

To understand natural language like human beings, it is necessary to have knowledge like human beings.

NLU is core of AI



"The next big step forward in deep learning is about getting the neural network to really understand the content of files."

The Father of the deep learning :
Geoffrey Hinton



"If you give me a billion dollars, I will build a natural language processing research project of NASA-level with this \$1 billion"

Machine learning expert, fellow of the American Academy of Sciences : Michael I. Jordan



"The next frontier of deep learning is natural language understanding"

Facebook's head of artificial intelligence : Yann LeCun



"In the next decade, those who understand languages will win the world."

Microsoft's Global Executive Vice President: Shen Xiangyang

**Natural Language Understanding (NLU):
The jewel in the crown of Artificial Intelligence**

SESSION 4

Essential Attribute and Core Tasks of Computational Linguistics from the Perspective of Artificial Intelligence

👤 Yude Bi & Danhui Yan Fudan University(China)



2. Natural Language Understanding

语言层面：自然语言理解（如何让机器理解语言）

"If the core of AI doesn't have language capabilities,
it's impossible to have a human-like AI system."

"Language ability is the most obvious difference
between humans and artificial intelligence."

—Josh Tenenbaum,

Professor of Cognitive Science and Computing in MIT



SESSION 4

Essential Attribute and Core Tasks of Computational Linguistics from the Perspective of Artificial Intelligence

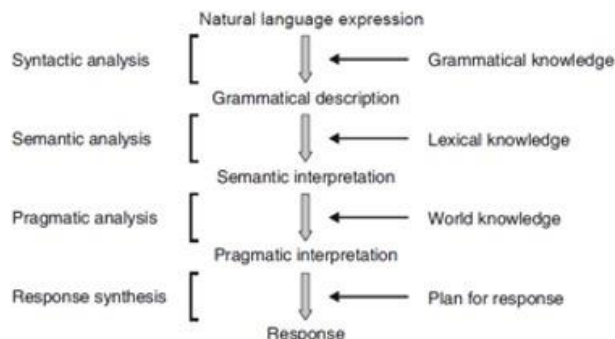
👤 Yude Bi & Danhui Yan Fudan University(China)

Reflection: Human Understanding and Machine Understanding.

- **How do we define “understanding”?**
 - Be able to think and act appropriately based on all known information
 - Natural Language Understanding is seen as determining the situation described in a sentence or text.
- **How can we achieve understanding?**
 - Make judgments or decisions based on knowledge or information or perception. Under the guidance of background knowledge (the understanding of the world), some uncertain information in literal expression is determined, or inferred based on literal content, so as to achieve the interpretation of natural language.
 - **Master the literal semantics, grasp the semantics beyond the literal (the part reflected and the part not reflected), and formalize and align the two.**
- **Factors that enable understanding**
 - Multimodal: speech-image-body posture-language-common sense knowledge, etc.;
 - **When human understands natural language, they often use a variety of information (including context, theme, personal background knowledge, etc.). But in different circumstances, the utilization of information is various. Sometimes it only needs simple linguistic knowledge, but sometimes it goes further, which even needs historical and cultural knowledge, or professional background knowledge, in order to correctly understand a certain text.**

NLU under the semiotic paradigm (machine)

The initial input is a sequence of symbols (already recognized and encoded for computation) and the main stream to the final output consists of symbol manipulations. In the case of a not unique solution, namely ambiguity, the result in each level of syntactic, semantic, and pragmatic processing is designed to be passed to the next processing level. This symbolic approach does not refuse any incorporation with statistical, connectionist, or other ones. A natural language understanding system is destined to respond in a certain style or protocol in order to show the evidence of its valid understanding of the input natural language expression. The final output can be varied according to the plan of what kinds of responses are to be synthesized, for example, paraphrase, translation in another language, answer to question, cross-media translation into still picture, animation, robotic action and so on.



SESSION 4

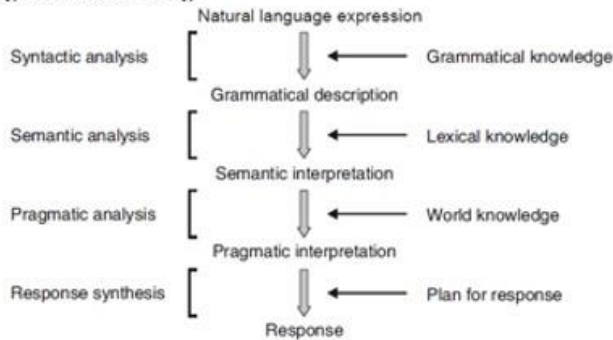
Essential Attribute and Core Tasks of Computational Linguistics from the Perspective of Artificial Intelligence

Yude Bi & Danhui Yan Fudan University(China)

NLU under the semiotic paradigm (machine).

The core of natural language understanding as the goal of natural language processing is the systematic methodology of knowledge representation and its computation for semantic or pragmatic understanding. Here, 'pragmatic understanding' means the total flow of natural language understanding in the figure below and that without pragmatic analysis is called 'semantic understanding', especially.

Natural language understanding, as the goal of NLP, is a systematic methodology of knowledge representation and computing, so as to achieve semantic or linguistic understanding.

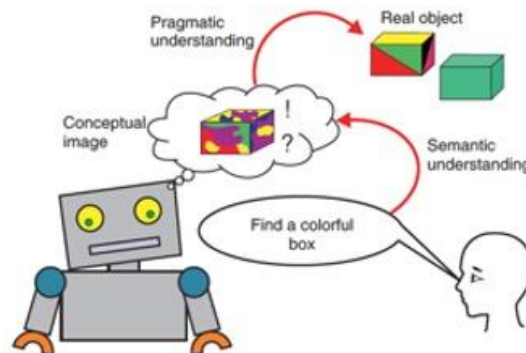


Machine perspective: Natural Language Understanding

The core of natural language understanding as the goal of natural language processing is the systematic methodology of knowledge representation and its computation for semantic or pragmatic understanding. Here, 'pragmatic understanding' means the total flow of natural language understanding in the figure below and that without pragmatic analysis is called 'semantic understanding', especially.

Natural language understanding, as the goal of NLP, is a systematic methodology of knowledge representation and computing, so as to achieve semantic or linguistic understanding.

Natural language understanding needs to take into account both understanding of semantics and linguistics



SESSION 4

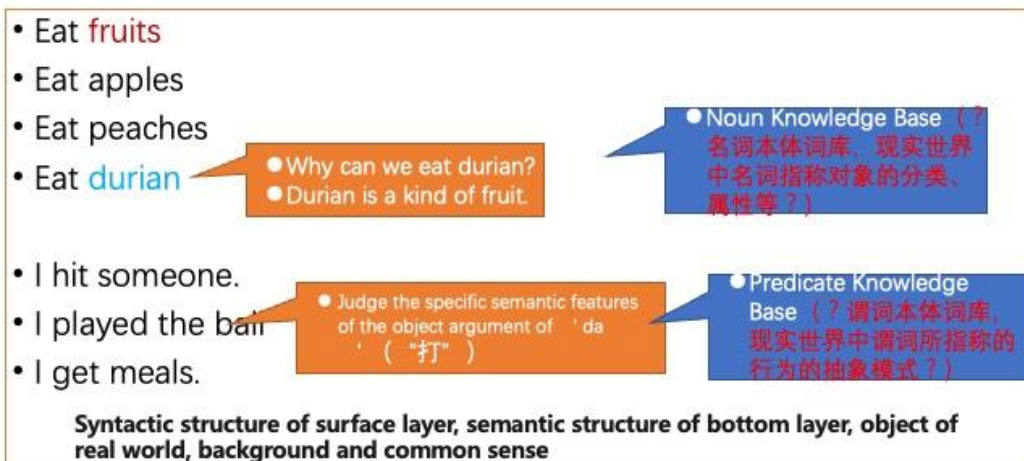
Essential Attribute and Core Tasks of Computational Linguistics from the Perspective of Artificial Intelligence

👤 Yude Bi & Danhui Yan Fudan University(China)

Machine perspective: Natural Language Understanding

- Microscopically, natural language understanding refers to a mapping from natural language to the inside of a machine; These features include:
 - Answer questions: The computer can correctly answer questions entered in natural language ;
 - Abstract generation: The machine produces a summary of the input text ;
 - Interpretation: Machines can retell the input natural language information with different words and sentence patterns ;
 - Translation: Machines can translate one language into another ;
 -
- If computers can understand and process natural language, the information exchange between people and computers can be carried out in the native language that people are familiar with, which will be a major breakthrough in computer technology.
 - Since the creation and use of natural language is a highly intelligent expression of human beings, the study of natural language processing can also help to uncover the mystery of human's high intelligence and deepen the understanding of language ability and thinking. Natural language understanding is of great significance in both application and theory.

How do humans understand language and reason?



SESSION 4

Essential Attribute and Core Tasks of Computational Linguistics from the Perspective of Artificial Intelligence

Yude Bi & Danhui Yan Fudan University(China)



3. Computational Linguistics and Natural Language Processing

The origin of the problem

• Is it possible to design a computer that can translate?

[Reproduced by permission of the Rockefeller Foundation Archives]

March 4, 1947



Warren Weaver

Norbert Wiener



March 4, 1947

I have wondered if it were unthinkable to design a computer which would translate. Even if it would translate only scientific material (where the semantic difficulties are very notably less), and even if it did produce an inelegant (but intelligible) result, it would seem to me worth while.

Dear Professor:

I was terribly sorry, when in Cambridge recently, that I got un-avoidably held up by several unexpected jobs, and did not get a chance to see you.

The thing I wanted to ask you about is this. A most serious problem for USSR and for the constructive and peaceful future of the planet is the problem of translation, as it unacceptably affects the communication between peoples. Asley has recently told me that they are appalled by the magnitude and the importance of the translation job.

Recognizing fully, even though necessarily vaguely, the semantic difficulties because of multiple meanings, etc., I have wondered if it were unthinkable to design a computer which would translate only of the kind translate only scientific material, where the semantic difficulties are very notably less, and even if it did produce an inelegant, but intelligible result, it would seem to me worth while.

Also knowing nothing official about, but having guessed and inferred considerable about, powerful new mechanical methods in cryptography - methods which I believe succeed even when one does not know what language has been coded - one naturally wonders if the problem of translation could conceivably be treated as a problem in cryptography. When I look at an article in Business, I say "This is really written in English, but it has been coded in some strange symbols. I will see how well I can decode."

Have you ever thought about this? As a linguist and expert on computers, do you think it is worth thinking about?

Cordially,

WARREN WEAVER,

Professor Norbert Wiener
Massachusetts Institute of Technology
Cambridge 39, Massachusetts

WIENER

SESSION 4

Essential Attribute and Core Tasks of Computational Linguistics from the Perspective of Artificial Intelligence

👤 Yude Bi & Danhui Yan Fudan University(China)

Four Strategies For Translation

- Over the next two years, W. Weaver, encouraged by his colleagues in Rockefeller Foundation, eventually formed a memorandum, Translation (1949, Carlsbad, New Mexico)
 - Perhaps it was the most influential work of the early days of machine translation, it set goals and methods for machine translation before most people knew anything about the potential function of computers, and led the United States to begin research on the problem first.
- In the memo, Weaver proposed four strategies :
 - Ambiguity can be solved by examining the direct context.
 - The connectivism approach. Inspired by McCulloch's research on early neural networks, Weaver believes that translation can be regarded as a formal logic problem, and the conclusions in the target language are derived from the premise of the source language (premises)
 - Cryptography method. Russian-English translation, can be regarded as encrypted English.
 - Linguistic methods. Weaver believes that linguistic unity may exist behind all human languages, which can be used directly in machine translation.

The birth of Computational Linguistics

- Since W. After Weaver came up with the concept of machine translation, the American and British academics became interested in machine translation (MT) and were supported by the industry.
- MIT held its first conference on machine translation in 1952.
- The first issue, *Mechanical Translation*, appeared in 1954.
 - Computational Linguistics, such expression first appeared in the mid-1960s.
 - The Association of Machine Translation and Computational Linguistics, established in 1962, chose *Mechanical Translation* as its conference publication.
 - *Mechanical Translation* was renamed as *Mechanical Translation And Computational Linguistics* in 1965, and computational linguistics appeared in the name with a smaller font.

SESSION 4

Essential Attribute and Core Tasks of Computational Linguistics from the Perspective of Artificial Intelligence

Yude Bi & Danhui Yan Fudan University(China)

The birth of Computational Linguistics

- The formal appearance of Computational Linguistics as a term should be traced back to David Hays' work in ALPAC (Automatic Language Processing Advisory Committee of the National Academy of Sciences) in the United States.
 - In 1964, the American Academy of Sciences established the Language Automatic Processing Advisory Committee to investigate the research on machine translation. In November 1966, an investigation report entitled “ Language and Machine ” was published, which was referred to as ALPAC report. It was believed that “ there is not much reason strong support to machine translation at present ” and “ machine translation has encountered an insurmountable semantic barrier ”. Since then, machine translation research has entered a downturn worldwide.。
 - The term computational linguistics (CL) first appeared formally in the report and remained.

The birth of Computational Linguistics

- Since then, formal system has increasingly become the core of linguistic theory research. Computational Linguistics also develops to create and apply formal systems to solve language problems
 - In 1960, John Cocke put forward a method to get all possible analysis of strings by using context-free rules, which is now called Cocke-Kasami-Younger(CKY) algorithm. 。
 - Soon, it became clear that a more powerful formal system was needed to deal with human language, and line graph analysis, extended transition network, unified syntax and other formal methods and calculation methods had been proposed.
- At that time, these studies could be divided into two categories:
 - One is a theoretical study, arguing that CL ' s goal can bring important progress to linguistic theory.
 - Formal systems need to ensure both internal consistency and formal complexity, which can be best represented in syntactic formalization : generalized phrase structure grammar, lexical function grammar, head word driven phrase structure grammar and finite state method used in morphology and phonology research
 - The other type is technical research, which has existed since MT was proposed. It is necessary to design a technology based on complete scientific principles to support more and more practical application requirements, such as machine translation, information extraction, summary, grammar check, etc. These cannot rely solely on linguistic methods to succeed.

SESSION 4

Essential Attribute and Core Tasks of Computational Linguistics from the Perspective of Artificial Intelligence

👤 Yude Bi & Danhui Yan Fudan University(China)

Natural Language Processing

- As a result, it is believed that computational linguistics cannot meet technical challenges, which in turn has caused many setbacks in the study of computational linguistics。
 - On the one hand, computational linguistics still has a lot of work to do; One important reason is the lack of a (broader) technical environment in which computational linguistics can exist.
 - There are also many people think that computational linguistics has found a way of development in statistical natural language processing.
- After 1980, with the rapid development and popularization of computer networks, language engineering technology with the goal of developing practical natural language processing system came into being, and the term natural language processing (NLP) was born.
 - In short, statistical natural language processing gives probability to the most likely result in the process of text (discourse) analysis and takes it as the correct answer.

The Debate on Methodology: Rationalism and empiricism

- Rationalism mainly studies the structure of human language knowledge, and the actual data only provide indirect evidence of intrinsic knowledge, while empiricism is directly the actual linguistic data.
- Rationalism is usually based on Chomsky's language principle. Empiricism is based on Shannon's information theory, which endows linguistic events with probability.
- Rationalism is usually the study of special statements and phenomena to obtain an understanding of human language ability, while empiricism focuses on statistical analysis of the actual statements used in the corpus.



SESSION 4

Essential Attribute and Core Tasks of Computational Linguistics from the Perspective of Artificial Intelligence

👤 Yude Bi & Danhui Yan Fudan University(China)

How to input language knowledge

- For example
 - **the boy saw the girl with the telescope**
 - The phrase "with the telescope" in a sentence is more likely to modify the verb "saw" because "telescope" often appears in similar environments in the language as a tool for "seeing".
 - Although this analysis is an undeniable fact, it is not the real situation of language.
 - If it is recognized that text processing depends on knowledge of the world and language, the challenge of statistical natural language processing may be greater than "language innateism".
 - Statistical natural language processing can only be constructed by reading in text, without direct experience of the world
 - The problem remains to be solved
 - Based on the text, how much of this knowledge is required by the NLP and how much is available to the NLP?

The Problem of Neural network (deep learning) : uninterpretability



Yoshua Bengio

How to understand "Natural Language Understanding"

- Joshua Benjio, the godfather of contemporary artificial intelligence and winner of the 2018 Turing Prize, attributes the dramatic changes in aitic intelligence in recent years to the advent of "deep learning.".
 - Machine translation, speech recognition, etc.
- The existing natural language processing system has a lot of vocabulary and language conversion skills, but does not understand the true meaning of sentences, the machine "will make some very stupid mistakes, even without the understanding ability of two-year-olds." .
 - Humans solve translation problems and question-and-answer questions

SESSION 4

Essential Attribute and Core Tasks of Computational Linguistics from the Perspective of Artificial Intelligence

👤 Yude Bi & Danhui Yan Fudan University(China)

What else can we do after deep learning is rebuffed?

Sun Maosong: A New Artificial Intelligence Algorithm Combining Big Data with Big Knowledge on the basis of Knowledge Base



- Deep learning based on massive amounts of data has brought rapid performance improvements in areas such as face recognition, speech recognition and machine translation, but big data-driven deep learning is a typical black box. Natural language understanding is still stuck in the "intelligence without understanding", "it seems that the machine translation is good, as if understand, in fact, there is no understanding." Machines are particularly vulnerable when dealing with complex semantics. "
- Without a formal human knowledge system, machine reasoning is difficult to achieve, and the interpretability and robustness of artificial intelligence are impossible to talk about.
- The construction of a knowledge base may be a solution to natural language processing. Because knowledge is naturally explanatory, so is the knowledge base.
- There is still a long way to go to achieve true artificial intelligence.

Towards the third generation of artificial intelligence: the combination of knowledge and data-driven approaches

《中国科学》（信息科学）2020年第9期



- The first generation of knowledge-driven AI uses knowledge, algorithm and computing power to construct AI.
- The second generation of data-driven AI uses data, algorithm and computing power to construct AI.
- The third generation AI: Establish an AI that fully reflects human intelligence, that is, establish robust and interpretable AI theories and methods, and develop safe, credible, reliable and extensible AI technologies.
- Development idea: Combine the first generation of knowledge-driven and the second generation of data-driven, and construct a more powerful AI by simultaneously utilizing four elements: knowledge, data, algorithm and computing power. 。

SESSION 4

Essential Attribute and Core Tasks of Computational Linguistics from the Perspective of Artificial Intelligence

Yude Bi & Danhui Yan Fudan University(China)



毕玉德表示，“两个一百年”是全方位的系统建设目标，既有硬实力的跃升，也有软实力的积累，更是综合实力的汇聚。作为国家软实力建设的承担者、综合实力的贡献者，哲学社会科学工作者一是要永怀赤子之心，坚守文化自信。将对党的忠诚、对祖国的热爱、对民族的责任内化为研究热情和研究动力，推出塑造灵魂、引领思维的高质量研究成果，夯实文化自信的基础。二是要挺立学术前沿，坚持融合创新。当下，社会科学研究呈现从各学科界限分明到各学科间综合与融汇的发展趋势，从与自然科学的硬性区分走向二者的内在渗透与互动，融合交叉成为创新突破的最大契机。社会科学工作者既要有前沿的“精”、“深”，还要有事业胸怀的“广”、“博”。三是要面向现实实践，引领创新驱动。“资政育人”是社会科学研究的最现实课题。在推进一带一路建设的时代背景下，文化“走出去”既是其中的建设内涵，更是其助推动力。哲学社会科学工作者应当成为中国文化的传承者、创造者和宣传者，向世界讲好“中国故事”。



4. Essential attributes and core tasks of computational linguistics.

SESSION 4

Essential Attribute and Core Tasks of Computational Linguistics from the Perspective of Artificial Intelligence

👤 Yude Bi & Danhui Yan Fudan University(China)

Computational Linguistics

- It analyzes and processes natural language by establishing a formal mathematical model, and realizes the process of analysis and processing with programs on computers, so as to achieve the goal of simulating part or even all of human language abilities with machines. .
- For a period of time,
 - From a subject point of view, it's called computational linguistics (theory)
 - From an engineering point of view, it's called natural language processing (method).
- However, today's "natural language processing", especially in the era of deep learning, focuses more on solving specific application problems (machine translation, automatic question answering, etc.).
 - Language is just as data, language → speech data
 - Speech data processing or text data processing is more appropriate;
 - In particular, the method of deep learning (neural network) is getting farther and farther away from "language"

The essential attribute of computational linguistics

- How to use the computer to understand and generate the language
- **Language and computation: language is the object of the study and computation is the means to carry out the study.**
- Language itself has its inherent laws and characteristics, and natural language processing techniques can be applied to all stages of language research (form, syntax, semantics, etc.) ;
 - based on large-scale corpora ,computers can use their powerful computing power and technological advantages to discover language phenomena that are difficult to find, and at the same time effectively verify and improve language rules, thus realizing the integration of language and technology.
- It aims at discovering language rules, exploring language mechanisms and understanding language. —— Scientific issues
- Focusing on language, aiming at discovering language rules, exploring language mechanism and understanding language, and adopting the method of "language+computation", the study of language generation mechanism and knowledge representation is helpful to solve the interpretable problem.
- This should be the essential requirement and discipline connotation of computational linguistics.

SESSION 4

Essential Attribute and Core Tasks of Computational Linguistics from the Perspective of Artificial Intelligence

Yude Bi & Danhui Yan Fudan University(China)



应用语言学的时代面向

2020年06月03日 作者：潘玥斐

构建中国特色哲学社会科学，学科体系是基础，学术体系是支撑，话语体系是集中反映。三者之间互为一体，具有极强的体系性和高度依存关系。复旦大学外国语言文学学院教授毕玉德认为，对于应用语言学来说，完善、科学的学科体系有利于研究者把握研究方向，超越既有研究框架和研究范式，以更开阔的视野认知和探索应用语言学研究领域。

毕玉德认为，未来，应更加关注包括冷门方向代际传承、不同学科交叉融合以及基础研究和应用研究相互促进的问题。第一，学术担当不能忘。站在加快构建“三大体系”的高度，坚持以加快推进“三大体系”建设为引领，统筹谋划学科规划、人才队伍和科学研究“三位一体”布局。第二，学术开拓不能忘。厘清应用语言学学科内涵，以问题为导向，勇于创新、大胆创新；打破学科界限，“拥抱”新技术，善于跨越边界，寻求创造性的解决方案。把自然科学的思想融入社会科学，丰富深化“新文科”内涵。第三，学术传统不能忘。要用中国智慧做有思想的学问。繁荣中国学术、发展中国理论、传播中国思想需要注入中国智慧、凝聚中国力量、彰显中国价值。中华优秀传统文化是中国特色学术体系强内功、厚底蕴的强大后盾。

Thinking about intelligence: Evolutionary mode

- The way language intelligence evolves :
data→**Information**→**knowledge**→**intelligence**
- In the face of linear "language data", it is necessary to formalize it in order to extract "information" so that it can be computed, and then structure it to form "knowledge", and then make comprehensive use of knowledge, and finally form "intelligence".

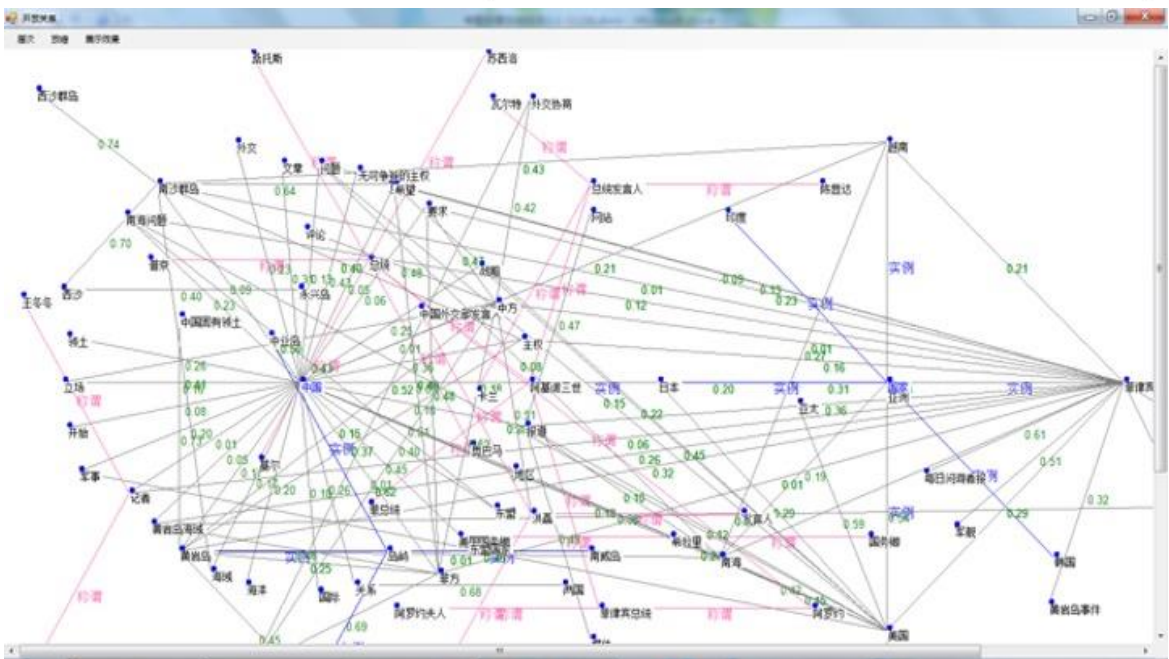
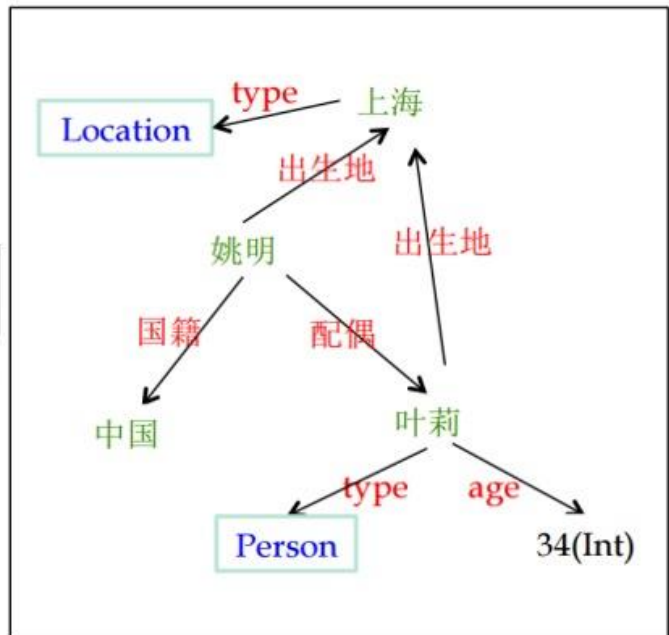
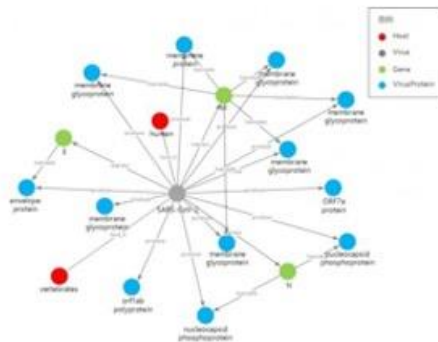


SESSION 4

Essential Attribute and Core Tasks of Computational Linguistics from the Perspective of Artificial Intelligence

Yude Bi & Danhui Yan Fudan University(China)

Google launched the Knowledge Graph in May 2012 (Google Knowledge Graph)
: 三元组



SESSION 4

Essential Attribute and Core Tasks of Computational Linguistics from the Perspective of Artificial Intelligence

Yude Bi & Danhui Yan Fudan University(China)

Event Knowledge Graph: Better reflects the connections between things
 Linguistic problems: reflect the logical semantic relations between things

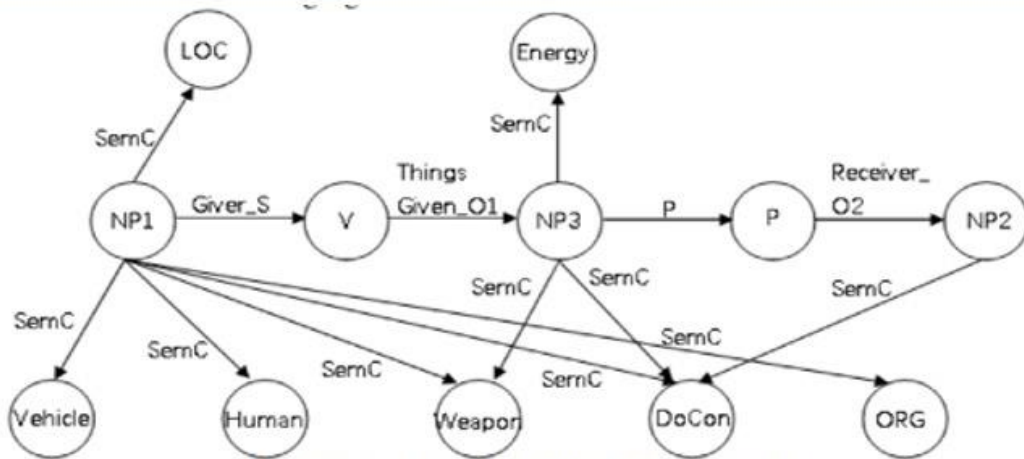


Fig. 2. Visualization of NP1+V+NP3+cho+NP2

经卫星图像确认，美国海军“罗斯福”号航母已于1月23日上午10时许，经由巴士海峡进入南海活动。

The research task of linguistics in the era of intelligence

- In the age of artificial intelligence, linguistics is connected with computer science, artificial intelligence and other disciplines, and the study of language units is not limited to pure linguistics. Linguistics is faced with the task of building a language model with computable and operational.
- Language technology includes the induction, representation and application of language knowledge, which is the core technology language knowledge base in computer system.

It is only one of the purposes to express language knowledge in a formal way, and the more important purpose is to “compute”. Only by expressing language knowledge in a formal way can knowledge be used for computing and simulate people’s language ability (Yu Shiwen, 2007: 24).

SESSION 4

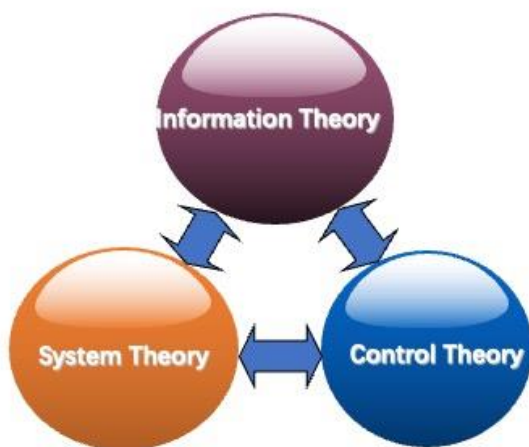
Essential Attribute and Core Tasks of Computational Linguistics from the Perspective of Artificial Intelligence

Yude Bi & Danhui Yan Fudan University(China)

Core ideas

- From the perspective of **computational linguistics**, drawing on modern linguistics theory and computer science, artificial intelligence and other related theories, we should fully consider the characteristics of language itself, and strive to make a **formal and structured description of language units**, so as to provide practical theoretical support and technical support for the establishment of a **language knowledge expression system** oriented towards language intelligence processing.
- Under the guidance of knowledge representation and linguistic theory, a set of scientific and reasonable formal and structured expressions is created.

technology idea



SESSION 4

Essential Attribute and Core Tasks of Computational Linguistics from the Perspective of Artificial Intelligence

👤 Yude Bi & Danhui Yan Fudan University(China)

Sun Maosong: A New Artificial Intelligence Algorithm Combining Big Data with Big Knowledge on the basis of Knowledge Base

- To truly understand language and overcome the difficulties of artificial intelligence, knowledge-driven method is needed.. The natural language processing driven by big data has done well, while the natural language processing driven by big knowledge or abundant knowledge has just started.
- Although there are some knowledge bases all over the world, there is no human knowledge base that can really drive natural language processing, especially the common sense base.
- To build this knowledge system, we need scholars with far-reaching vision and a big pattern to lead a group of people and carry out persistent resource construction work.

Feng Zhiwei: In the information age, linguists with both liberal arts and sciences are needed.

- Since the object of study of natural language processing is language, of course, it should be supported by formal linguistic knowledge. With the help of computer technology, we also need to rely on formal linguistic knowledge to achieve the objective. Linguists have much to do in natural language processing. We should seriously study the knowledge of linguistics and thoroughly study various laws and rules in linguistics. At the same time, we should keep pace with the times, renew our knowledge, learn natural language processing technology, modern mathematics, computational programming technology. Therefore, we can become a new generation of linguists with both liberal arts and sciences, and finally make difference.

Guangming Daily (October 21, 2018 12th)

In the age of artificial intelligence, linguists who understand technology are needed

SESSION 4

Essential Attribute and Core Tasks of Computational Linguistics from the Perspective of Artificial Intelligence

👤 Yude Bi & Danhui Yan Fudan University(China)




Conclusion

- With the help of modern science and technology, linguistics research in the era of artificial intelligence can expand new research fields and realize the integrated development of language and technology. It will surely become the breakthrough point and growth point of **humanities** and **natural sciences** and play a positive role in the development of a new generation of artificial intelligence.
- Basic research expands the boundaries of human knowledge and applied research expands the boundaries of human ability. Linguistics has the characteristics of both basic science and applied science, and the study of language itself belongs to the category of basic science. Therefore, we believe that we should give full play to the advantages of big data technology and computing technology to expand the boundaries of language understanding.



SESSION 5

Moderator  Taehee Kim Konkuk University

| AI and Philosophy | |
|-------------------|--|
| 16:15 ~ 16:45 KST | <p>How to Make Ideas Clear: Artificial Intelligence and Human Instinct from a Peircean Perspective</p> <p> Cheongho Lee Sangmyung University(Korea)</p> |
| 16:45 ~ 17:15 KST | <p>The Meaning of Labor in the Age of Artificial Intelligence</p> <p> Sunjin Yang Korean National Police University(Korea)</p> |
| 17:15 ~ 17:45 KST | <p>Zhouyi's Biantong World and the Algorithm of Yin-yang from the Dimension of AI Connectionism</p> <p> Yonjae Kim National Gongju University(Korea)</p> |
| 17:45 ~ 18:15 KST | Discussion |

SESSION 5

How to Make Ideas Clear: Artificial Intelligence and Human Instinct from a Peircean Perspective

 Cheongho Lee Sangmyung University(Korea)

Abstract

This paper surveys the ways that artificial intelligence (AI) and human beings make ideas clear based on Charles Peirce's thought. Peirce considered the origin of human mind to be instinctive. The instinctive human mind performs two processes of determination throughout internal and external domains of the mind. Human beings, as they have the individualized mind, experience the actualization of ideas internally and externally. From outside in, the mind introduces ideas that are externally actualized in the nature into the internal area of mind. From inside out, the mind projects internally actualized ideas into the exterior domain of mind. Peirce sees that "cognitive capacity" of mind inter-relates the internal and external domains of mind. The instinctive mind gradually develops into the rational mind, that is, the crystalized mind in Peircean terms. Regarding the work of making ideas clear, AI is under radically different conditions. AI fulfills the actualization of ideas in the different way than humans do because AI does not possess an individualized mind that essentially lies under specific circumstances of internal and external lives. However, AI is expected to have the crystalized mind and for this reason can deal with ideas in a rational way. Based on this insightful understanding, we can divine the way AI works in terms of making ideas clear.

key words : Peirce, artificial intelligence, mind, instinct, to make ideas clear

SESSION 5

How to Make Ideas Clear: Artificial Intelligence and Human Instinct from a Peircean Perspective

👤 Cheongho Lee Sangmyung University(Korea)


ICAIH 2021

THE 4TH INTERNATIONAL
CONFERENCE ON
ARTIFICIAL
INTELLIGENCE
HUMANITIES


October 2, 2021 10:00~18:30 (KST)


How to Make Ideas Clear

피스의 관점에서 본 본능과 인공지능



이 청 호





발표순서

- I. 서론
- II. 인간의 실존적 조건: 이론(theory)과 실제(practice)
- III. 두 종류의 본능
- VI. 실제적 차원의 본능: 본능적 믿음(instinctive belief)
- V. 이론적 차원의 본능: 본능적 추론(instinctive reasoning)
- IV. 결론: 인간과 인공지능
- 토론 및 질의응답

SESSION 5

How to Make Ideas Clear: Artificial Intelligence and Human Instinct from a Peircean Perspective

👤 Cheongho Lee Sangmyung University(Korea)

I. 서론



- Human vs Artificial Intelligence
- 퍼스(C. S. Peirce, 1839-1914)
- 인간의 삶에 있어 퍼스가 제시하는 본능의 역할에 대해 살펴보고 이를 바탕으로 하여 인간과 인공지능은 어떤 점이 유사하고 어떤 차이가 있는지를 미루어 짐작하고자 한다.



II. 인간의 실존적 조건: 이론(theory)과 실제(practice)

- 이론과 실제의 구분
- 퍼스는 인간의 실존적 조건이 이론(theory)의 영역과 실제(practice)의 영역으로 이루어진다고 파악했다.
- 실제적(practical) 인간은 필수적으로 중요한 문제들(vitally important matters)을 해결하는 데 믿음에 근거하여 행동하게 된다(CP 1.635).
- 과학적(scientific) 인간은 가설(hypothesis)을 통해 진리를 확립하려 한다. 과학적 인간은 심사숙고하며 자기통제적(self-controlled)인 추론을 지속적으로 감행한다(CP 4.476).

SESSION 5

How to Make Ideas Clear: Artificial Intelligence and Human Instinct from a Peircean Perspective

 Cheongho Lee Sangmyung University(Korea)

II. 인간의 실존적 조건: 이론(theory)과 실제(practice)

- 이론과 실제의 연계
- 퍼스는 실제 삶에서 이론과 실제는 서로 밀접하게 연계되어 있다고 파악하였다.
- 실제적 인간: 행동 계획은 이론(theory)에 바탕을 두고 있다 (CP 6.3). 과학적 인간: 이론적 추론은 믿음으로부터 시작
- 퍼스는 이론과 실제를 연계하는 데 있어 본능이 중요한 역할을 한다고 보았다. 아무리 복잡한 이론이라 하더라도 그 출발점이 되는 것은 본능적인 믿음(instinctive belief)과 상식(common sense)에서 촉발된 가설들(hypotheses)이다.

II. 인간의 실존적 조건: 이론(theory)과 실제(practice)

- 이론과 실제의 연계
- 퍼스는 인간의 경험은 본능이라는 신비로운 경험으로부터 시작한다고 보았다.
- 퍼스가 보기에 인간은 올바른 추론(reasoning)에 대한 본능을 가지고 있으며(CP 2.3) 그러한 관점에서 본능은 우주의 신비로운 원리를 추측하는(guessing) 능력이다(CP 6.491, 6.530).
- 본능은 인간의 탐구의 출발점일 뿐 아니라 행동의 출발점이다.
- 본능은 무오류성(infallibility)을 전제한 믿음을 형성하게 하고 인간은 그러한 믿음에 따르는 방향으로 행동하게 된다.

SESSION 5

How to Make Ideas Clear: Artificial Intelligence and Human Instinct from a Peircean Perspective

👤 Cheongho Lee Sangmyung University(Korea)

Ⅲ. 두 종류의 본능

- 개별적 본능과 사회적 본능
- 퍼스는 역학적(mechanical) 본능과 사회적(social) 본능이라는 두 종류의 본능을 제시하였다(CP 7.382).
- 역학적 본능은 개별 인간을 보존하는 이기적 본능으로서, 음식 얻기, 개인 청결, 의료, 수면 등의 모든 역학적 기술과 관련된다(CP 7.378).
- 사회적 본능은 타인들의 이익(advantage)을 위해 재생산하는 본능으로 제시된다(CP 7.379).

Ⅲ. 두 종류의 본능

- 지식의 근원으로서의 본능
- 퍼스에게 있어 본능은 인간의 모든 지식의 근원이 된다.
- 역학적 본능은 영양분에 대한 필요를 공간과 힘에 관한 지식을 갖추도록 하여 우리를 응용 물리학자(applied physicists)가 되게 하였고, 우리가 가진 성적 욕망과 관련하여서는 동등한 마음을 지닌 다른 존재를 이해하게 하여 우리로 하여금 응용 심리학자(applied psychists)가 되게 하였다(CP 5.586).

SESSION 5

How to Make Ideas Clear: Artificial Intelligence and Human Instinct from a Peircean Perspective

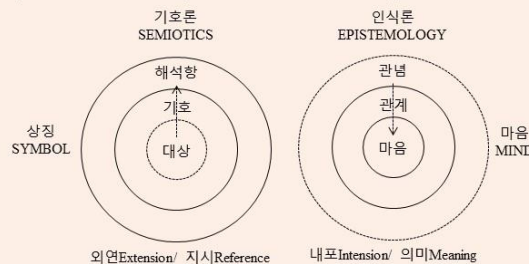
👤 Cheongho Lee Sangmyung University(Korea)

Ⅲ. 두 종류의 본능

- **관념(idea)과 두 종류의 경험**
- 퍼스는 인간의 지식 습득은 인간에게 두 종류의 관념(觀念, idea)이 선천적으로 주어졌기 때문에 가능하다고 보았다.
- **역학적 관념:** 첫 번째 종류의 관념들은 힘, 물질, 공간, 시간 등의 관념->자연에 현실화된 관념을 마음에 받아들임
- **정신적(도덕적) 관념:** 자신의 동료들이 어떤 존재인지 그리고 어떻게 주어진 상황에서 행동하는지에 관한 관념이다(CP 2.753).
-> 마음에 관념을 현실화하고 인간은 이를 물질세계에 전달하게 된다(CP 2.149).

Ⅲ. 두 종류의 본능

• 관념(idea)의 현실화(actualization)



- **기호론(semiotics)**은 '객관화' 즉 '연장적 사고'에 기반
 - **공간적(spatial):** 지시는 표상되는(presentational) 공간에서 순서를 바꿀 수 있고 무한히 배열
- **인식론(epistemology)**은 '주관화' 즉 '내포적 사고'에 기반
 - **시간적(temporal)**이고 실제적(real)

SESSION 5

How to Make Ideas Clear: Artificial Intelligence and Human Instinct from a Peircean Perspective

 Cheongho Lee Sangmyung University(Korea)

VI . 실제적 차원: 본능적 믿음(instinctive belief)

- 믿음의 형성
- 퍼스의 관점에서 본능이 인간의 감성(sentiment)과 관련되고 (CP 1.634), 인간의 행동을 안내하는 원리인 믿음과 관련된다(CP 3.160).
- “믿음의 고정(the Fixation of Belief, 1877)”
 - 탐구의 과정은 의심(doubt)을 불식시키는 믿음의 상태를 획득하기 위한 과정 (CP 5.374).
 - 이러한 의심의 상태는 데카르트가 말한 순수한 의심(paper doubt), 즉 삶의 조건과 상관없는 의심은 아니라 (CP 6.498), 실제로 살아가면서 갖게 되는 의심(real and lived doubt)을 의미한다.

VI . 실제적 차원: 본능적 믿음(instinctive belief)

- 이성의 역할
- 모호한 본능적 믿음(vague instinctive beliefs)은 마음의 높고 낮은 수준이나 세대의 차이에 상관없이 누구에게나 존속할 수 있다.
- 인간은 본능 이외에도 본능과 구별되는 이성(reason)의 힘을 소유하고 있다(CP 2.160).
- 퍼스가 보기에 인간은 상상력이나 반성적 사고가 작용하여 맹목적인 본능적 충동을 벗어나려는 준목적성(quasi-purposefulness)을 지니고 있다(CP 7.381 n.19).
- 이러한 본능의 반성적 측면은 맹목적이고 통제를 벗어나려는 본능적 믿음을 교정하는 과정을 포함한다.

SESSION 5

How to Make Ideas Clear: Artificial Intelligence and Human Instinct from a Peircean Perspective

👤 Cheongho Lee Sangmyung University(Korea)

VI . 실제적 차원: 본능적 믿음(instinctive belief)

- 믿음의 교정(correction): 소급의식(遡及意識)
- 그렇다면 믿음의 교정은 어떻게 이루어지는가.
- 퍼스는 인간이 본능적으로 타인의 입장에서 반성적으로 생각해 보는 소급의식(遡及意識, retroconsciousness)을 지니고 있으며, 이러한 경향은 우리가 규칙을 준수하려는 모든 경향에 내재되어 있는 윤리적 수준의 행위 동기이다(CP 1.586).

VI . 실제적 차원: 본능적 믿음(instinctive belief)

- 본능적 믿음의 규범성(normativity)의 근원: 공공성(publicity)과 무오류성(infallibility)
- 본능적 믿음의 규범성(normativity)은 공공성(publicity)과 무오류성(infallibility)을 그 근원으로 한다.
- 공공성을 지닌 상식(common sense)은 개별 행위자들이 접근할 수 있는 본능적 믿음의 총체를 지칭하게 된다(CP 6.498).
- 본능적 믿음은 "실제적으로" 오류를 내포하지 않은 것처럼 여겨진다는 무오류성(infallibility)을 지닌 것으로 여겨진다.

SESSION 5

How to Make Ideas Clear: Artificial Intelligence and Human Instinct from a Peircean Perspective

👤 Cheongho Lee Sangmyung University(Korea)

V. 이론적 차원: 본능적 추론(instinctive reasoning)

- 마음(mind)의 특성
- 본능적 믿음은 무오류성이나 통제를 벗어나려는 특성 때문에 기본적으로 이성의 힘이 작용하는 탐구의 길과 충돌하는 것처럼 보인다. 그러나 퍼스는 본능을 탐구의 기반(bedrock)으로 파악하였다.
- 퍼스는 인간의 마음은 본질적으로 본능적이라고 파악하였다. 인간의 마음은 본능적인 것, 즉 감성(sentiments)이며 인식(cognition)은 단지 이러한 마음의 표면일 따름이라고 보았다(CP 1.628).

V. 이론적 차원: 본능적 추론(instinctive reasoning)

- 본능적 마음의 이성적 마음으로의 발전
- 퍼스는 본능적 마음이 이성적 마음으로 발전한다고 보았으며(CP 7.380), "인식적 역량(cognitive capacity)"이 그러한 발전의 근원이라고 파악했다(CP 1.648).
- 인식적 역량은 인간이 관념을 현실화(actualization)하는 두 가지 방법을 가능케 한다. 이때 인간의 마음, 즉 내부의 영역과 자연, 즉 외부의 영역을 연결하는 능력이 바로 인식적 역량이다.
 - 내부와 외부를 연결하는 인식적 역량은 과정은 본능적 이성(instinctive reason)이 작용하여 가능하다(CP 6.475). 본능적 이성은 본능적 마음에게 끊임없이 질문하고 무수히 많은 탐구의 과정을 거치면서 마음은 점차 이성적인 방향으로 발전하게 되는 것이다.

SESSION 5

How to Make Ideas Clear: Artificial Intelligence and Human Instinct from a Peircean Perspective

 Cheongho Lee Sangmyung University(Korea)

V. 이론적 차원: 본능적 추론(instinctive reasoning)

- **본능적 추론의 과정**
- 퍼스가 보기에 이성적 본능이 작용하는 본능적 추론 (instinctive reasoning)은 무의식적이고 통제받지 않은 판단들을 다루는 과정임을 주지해야 할 것이다(CP 7.453).
- 특정한 본능적 판단(instinctive judgment)에 이르는 과정은 의식적이고 이성적으로 규명하기 힘들다. 본능적 판단은 우리가 가진 의문을 어느 정도 해소한 상태이며 이는 감성(sentiments)에 의지한 나름의 타당성(validity)을 지니고 있다(CP 1.632).
- 즉 본능적 판단은 우리의 마음이 특정한 신념에 다다르게 하는 역할을 수행한다.

V. 이론적 차원: 본능적 추론(instinctive reasoning)

- **가추(retroduction)로서의 본능적 추론**
- 본능은 탐구의 과정에서 발생하는 세 가지의 추론, 연역(deduction), 귀납(induction), 가추(abduction) 중 가추의 단계에서 활동하는 것으로 설명된다(CP 1.630).
- 가추의 추론이 작동한 결과 생성되는 것은 가설(hypotheses)이다. 가설들은 연역 추론의 출발점이 되고 귀납 추론을 통한 검증의 대상이 된다. 그러므로 모든 새로운 지식은 가추를 통해 촉발되며, 이런 의미에서 가추는 우리에게 새로운 관념을 제시하는 종합적(synthetic) 추론이 된다(CP 2.777).

SESSION 5

How to Make Ideas Clear: Artificial Intelligence and Human Instinct from a Peircean Perspective

 Cheongho Lee Sangmyung University(Korea)

V. 이론적 차원: 본능적 추론(instinctive reasoning)

- 인식의 과정(perceiving)과 일반관념
- 퍼스에게 있어 본능은 인식과 가추를 연결하는 데 있어 중요한 역할을 담당하게 된다.
- 인식의 과정(perceiving)은 본능적 추론이 관여하여 작용하는 과정이며 그 결과 형성된 인식은 자연이 작동하는 방식을 알아채는 인간의 능력이 된다(CP 5.173).
- 그런데 흥미로운 점은 이러한 인식의 과정을 통해 인간이 알게 되는 것은 퍼스가 일반성(generality)의 범주로 제시했던 제삼성(Thirdness)이라는 것이다. 즉 우리는 본능적 추론을 통해 일반관념을 추론하고 이것이 우리의 인식 판단을 이루게 된다는 것이다.

IV. 결론: 인간과 인공지능



- 인간과 인공지능의 차이1: 본능의 유무
- 인간: 본능 O
 - 본능은 생명체로서의 인간이 주어진 환경적 제약을 극복하고 생존하기 위해 필수적으로 요구되는 능력이다.
- 인공지능: 본능 X
 - 인공지능은 생명체로 존재하지 않기 때문에 생존에 관한 문제가 필수적인 것이 아니다. 그렇게 되면 인공지능이 본능이라는 능력을 가지기는 불가능하거나 불필요한 것으로 보인다.

SESSION 5

How to Make Ideas Clear: Artificial Intelligence and Human Instinct from a Peircean Perspective

 Cheongho Lee Sangmyung University(Korea)

IV. 결론: 인간과 인공지능



• 인간과 인공지능의 차이2: 마음의 작동방식

• 인간: 개별화된 마음 O

- 인간은 개별화된 마음이 외부와 내부의 두 영역에서 관념의 현실화 과정에 관여한다.

• 인공지능: 개별화된 마음 X

- 인공지능은 아직까지는 개별화된 마음을 지니고 있지 않으며 두 가지 관념의 결정 과정을 그대로 따를 이유가 없게 된다.
- 개별화된 마음 자체가 존재하지 않기 때문에 인공지능과 관념 사이의 상호작용은 전혀 예측할 수 없는 방향으로 일어날 수도 있다. 인간은 두 가지 방식 이외에 발생하게 되는 관념의 현실화를 경험할 수 없기 때문에 인공지능이 그러한 방식과 별개로 관념과 작용한다면 이를 알아채는 것이 불가능하기 때문이다.

IV. 결론: 인간과 인공지능



• 인간과 인공지능의 차이3: 행동 맥락의 유무

• 인간: 행동의 맥락 O

- 인간은 “주어진 상황”에 대한 의문을 적절히 해결하고 행동하기 위해 본능으로부터 주어지는 믿음의 안내를 따르는 것이 필연적으로 요구된다.

• 인공지능: 행동의 맥락 X

- 인공지능의 경우에는 살아가면서 경험하게 되는 문제 상황 자체가 존재하지 않기 때문에 의문이나 의심은 맥락이 결여된 순수한 궁금증이 된다.
- 인공지능은 발전된 정보처리 기술을 바탕으로 준맥락적 (pseudo-contextual) 상황에서 문제를 해결한다.

SESSION 5

How to Make Ideas Clear: Artificial Intelligence and Human Instinct from a Peircean Perspective

 Cheongho Lee Sangmyung University(Korea)

IV. 결론: 인간과 인공지능



- 인간과 인공지능의 유사점1: 추론에 근거한 인식
- 인간: 가추라는 본능적 추론의 과정에 근거한 인식
- 인공지능: 주어진 정보를 알고리즘에 따라 처리

IV. 결론: 인간과 인공지능



- 인간과 인공지능의 유사점2: 설명불가능한 인식 과정
- 인식이 어떻게 발생하였는지 정확하게 설명하기 어려움
- 인간: 본능은 설명할 수 없는 부분 존재
- 인공지능: 어떻게 특정한 인식을 하였는지를 인간이 이해하기 어려운 경우가 많음

SESSION 5

How to Make Ideas Clear: Artificial Intelligence and Human Instinct from a Peircean Perspective

 Cheongho Lee Sangmyung University(Korea)

IV. 결론: 인간과 인공지능



- 윤리적 행동의 경우: 이기심과 이타심
- 인간: 이기심 O, 이타심 X
 - 인간은 본능적으로 이기적 본성과 이타적 본성이 충돌한다.
- 인공지능: 이기심 X, 이타심 ?
 - 인공지능은 나름의 이타적 소급의식을 발휘할 수 있을 것으로 예상된다. 예를 들어 보편적인 관점에서 관련 당사자의 입장을 모두 고려하는 식으로 가능할 것이다.
 - 그렇지만 이것이 이기적 입장을 결여한 것이기 때문에 인간과 같은 맥락에서 윤리적 행동을 하는 것으로 보기는 어려울 것이다.

IV. 결론: 인간과 인공지능



- 결론: 인공지능의 미래
- 퍼스의 본능에 관한 이론을 바탕으로 하여 인간과 인공지능을 비교해 볼 때 인간과 근본적으로 다른 여건에 놓여 있으며, 인간의 삶에서 발휘되는 본능의 여러 측면들이 그대로 적용되기 어려운 것으로 예상된다.
- 물론 앞으로 인공지능이 어떠한 방향으로 발전할 것인지에 관해서는 정확히 예측하기 어려운 것이 사실이며 앞으로 개발될 인공지능이 개별적 마음을 갖게 되거나 환경적 여건에 영향을 받게 될 때 앞서 언급한 인간의 삶의 조건과의 간극은 많은 부분 좁혀질 가능성도 존재함을 부인할 수는 없을 것이다.

SESSION 5

How to Make Ideas Clear: Artificial Intelligence and Human Instinct from a Peircean Perspective

 Cheongho Lee Sangmyung University(Korea)

토론 및 질의응답

Thank You !
Q & A

SESSION 5

The Meaning of Labor in the Age of Artificial Intelligence

 Sunjin Yang Korean National Police University(Korea)

Abstract

Information technology has led to an era of artificial intelligence based on big data, and biotechnology has begun to display ambition to manipulate and design genetic information beyond the stage of deciphering human genetic information. In the era of the fourth industrial revolution, which combines information technology and biotechnology, convenience and quantitative achievements in human life have been provided, but there are concerns that the world will be threatened by human jobs while changing the environment of human life. With the advent of the age of artificial intelligence in various areas of our society, modern people have begun to ask fundamental questions about humans and ask again about their identity. The power to drive information technology and biotechnology is economic logic. The reason why modern people are enthusiastic about artificial intelligence based on big data is that economic reasons are the most direct. Modern society lives in an era when data becomes the most important capital. As today's capitalism has become a society where data-based information giants become the center of industry and big data becomes the most important capital, information giants are actively collecting information beyond the passive stage of collecting information, and unwanted information from individuals is flowing into information companies without their knowledge. In this trend of the times, information companies have begun to establish systems to decipher and monitor personal information in the process of actively searching for data. There is a risk of degenerating into a society in which information is decrypted and monitored, but it has become a mechanism of capitalism that is hard to go against in the age when information becomes capital. In the face of this social reality, individuals will have to pool their wisdom on what posture to take and have the opportunity to think through philosophers from the East and the West to pool their wisdom.

SESSION 5

The Meaning of Labor in the Age of Artificial Intelligence

👤 Sunjin Yang Korean National Police University(Korea)

인공지능 시대의 노동의 의미

양선진 (sunjinyang@hanmail.net or yangsunjinfr@naver.com)

목차

- 1. 서론
- 2. 인간의 본질과 노동
- 3. 인공지능중심 시대의 노동 환경과 노동 소외
- 4. 공감적 인간 사회
- 5. 인공지능 시대의 유희적 인간
- 6. 결론

SESSION 5

The Meaning of Labor in the Age of Artificial Intelligence

 Sunjin Yang Korean National Police University(Korea)

서론

- 4차 산업혁명의 시대에는 인간의 노동환경이 본질적인 변화를 초래할 수 있다는 입장에서 노동 문제를 철학적으로 면밀하게 검토한 철학자인 헤겔과 마르크스의 노동의 개념과 노동과 인간과의 관계를 고찰하도록 할 것이다. 그리고 4차 산업 혁명 시대에서 발생하는 자연 환경과 노동 환경의 상황을 검토하고 노동이 사라지는 시대에 상응하는 노동의 의미를 고찰하는 기회로 삼고자 한다.

II. 인간의 본질과 노동

- 헤겔 철학의 중심 개념은 소외이며 자신의 저서인『정신현상학(1807)』에서 자신의 철학의 기초로 삼는 개념이 바로 외화(Entäußerung) 내지 소외(Entfremdung)이다.
- ‘주인과 노예’(Herr und Knecht)의 관계
- 주인은 노예를 통해서만 자연과 관계를 맺고 노예를 통해서만 자신의 욕구를 충족할 수 있을 뿐이다.
- 노예는 자신의 결과물을 얻기 위해서는 저항하는 자연을 변형해야 한다. 따라서 노예는 자연을 변형하면서 주인을 섬긴다는 점에서 노예에게 노동은 존재론적 의미를 갖는다.
- 노예에게 노동은 자신의 자유를 발현할 수 있는 중요한 매개물이기 때문에 노예에게 노동이란 자신의 본질인 자유를 발현할 수 있는 가장 중요한 수단이라고 할 수 있다. 노예에게 자연에 대해 주인의식을 가지게 된다는 것이다. 노동은 주인의식을 확인할 수 있는 가장 중요한 수단이다.
- 반면에 주인에게 자연이란 노예를 매개해야만 자연을 변형할 수 있기 때문에 자연에 대해서 주인 의식을 가질 수 없다. 이처럼 우리가 상식적으로 생각하는 것과 달리, 헤겔은 노예가 주인보다 중요한 위치에 있다고 간주한 이유는 노예가 노동의 주체이기 때문이다.

SESSION 5

The Meaning of Labor in the Age of Artificial Intelligence

👤 Sunjin Yang Korean National Police University(Korea)

Ⅲ. 인공지능 시대의 노동 환경과 노동 소외

- 왕수인은 우주자연 속에 인간을 포함해서 무생물에서 고등동물에 이르기까지 하나라는 존재론적 입장에서 인간과 모든 존재와 관계(感應)한다고 본다. (『傳習錄』, 179條目, “夫人者, 天地之心, 天地萬物, 本吾一體者也.”)
- 네트워크 시대에 서양적 사상은 자연을 하나의 대상으로 파악하는 것이 아니라 관계의 집합으로 간주하는 새로운 관점은 과학기술의 폐해 속에서 등장하였다면, 동양의 사상은 기본적으로 자연과 인간을 하나로 보는 기반한다.
- 디지털화된 사회에서 하나의 네트워크적 관계(一理) 만이 존재할 뿐이다.(『傳習錄』, 157條目, 陰陽一氣也, 一氣屈伸而爲陰陽; 動靜一理也, 一理隱顯而爲動靜)
- 왕수인의 철학을 오늘날의 시각에서 해석한다면, 정보의 송신자가 정보를 보내는 자극을 주면(感) 수신자는 받은 정보에 대해서 반응(應)하는 관계를 영속적으로 서로 서로 유지한다.
- 왕수인의 사상에서 인간의 위상을 고찰하면, 인간이란 천지자연과 하나이며 본질적으로 천지자연의 마음(天地之心)이라고 한다. (『傳習錄』, 179條目, “夫人者, 天地之心, 天地萬物, 本吾一體者也.”)

IV. 공감적 인간(Homo empathicus)의 존재

- 왕수인의 인간존재란 세계의 존재를 밝혀주는 특별한 존재론적 위상을 가진다는 점에서 인간의 존재론적 특수성을 인정한다. 『王陽明全集』, 卷6, 「惜陰說」, 267頁, “良知卽天道.”
- 인간은 자연 세계와 연결되어 있으며 도덕적 존재로서의 인간은 세계가 고통을 받을 때는 항상 세계와 감응하면서 세계의 아픔을 공감할 수 있는 호모 사피엔스(Homo Sapiens)의 존재가 아니라 호모 엠파티쿠스(Homo empathicus)의 존재이다.

SESSION 5

The Meaning of Labor in the Age of Artificial Intelligence

 Sunjin Yang Korean National Police University(Korea)

V. 인공지능 시대의 유희적 인간

- 마르쿠제는 <에로스와 문명>(1955)에서 오직 노동의 종말과 리비도(libido)의 실현, 그리고 사회관계 속에서의 유희를 통해서 인간해방이 실현될 것이라고 주장하였다.
- 프로이트는 <문명과 그 불만>(1929)에서 인간의 리비도가 해방된다는 것은 인간 사회를 구성하는 문명이 후퇴하거나 퇴보하는 결과를 초래할 것이라고 우려했다.
- 인공지능 시대에는 일자리가 줄어들 것이라는 주장과 일자리가 늘어날 것이라는 주장이 팽팽한지만 확실한 것은 노동 시간이 줄어들면서 인간의 본질은 노동이 아니라 노동의 해방에 가까워진다는 사실이다.
- 헤겔의 해석과 달리, 인간의 본질은 도구적 인간(Homo faver)이나 지혜인(Homo sapiens)에서 유희적 인간(Homo ludens)으로 새롭게 규정할 수 있다.
- 호모 루덴스는 예술적 인간인 호모 아르텐스(Homo artens)은 유희적 인간의 한 부분 집합이라고 할 수 있다. 호모 루덴스 역시 전체적으로 호모 라보란스(homo laborans)의 변형이라고 할 수 있다.

V. 인공지능 시대의 유희적 인간

- 기 드보르(G. Debord)의 책 <스펙터클의 사회>는 여가활동이 노동과 구별되지 않는다는 점을 강조하고 있다. (Debord, G. 2002. The Society of the Spectacle. Hobgoblin. 2002, 13쪽.) 하지만 논자는 여가가 노동의 일환이 아니라 여가가 곧 노동이 될 것이라고 예측한다.

SESSION 5

The Meaning of Labor in the Age of Artificial Intelligence

 Sunjin Yang Korean National Police University(Korea)

결론

- 새로운 사회에는 인간의 일자리를 기계가 대체하면서 인간은 일자리를 잃을 가능성이 상당히 높아졌다는 예측이 많다. 따라서 인간에게 노동은 인간의 역할이자 본질이 아닐 수도 있는 상황에 처하게 될지도 모른다. 이것은 인간에게 분명한 위협이며 도전이다. 이러한 사회에서 인간들은 대부분의 삶을 어떻게 보낼 것인지가 중요한 사회적 주제가 될 날이 멀지 않았다. 만약 인간의 본질인 노동이 사라지는 사회가 도래한다면, 인간은 여가가 중심이 되는 사회가 될 것이며 이는 새로운 인간의 존재론적 위협이 될 것이며 이러한 사회에 대비해서 교육과 인간의 가치관을 새롭게 형성하는 기회가 필요할지도 모른다.
- 인공지능중심의 시대의 특징 가운데 하나가 경계들의 해체이며, 그 가운데 노동과 여가의 해체도 큰 특징이다. 또한 여러 기계들의 도입으로 인해 앞으로 인간 사회는 점점 많은 사람들이 노동시간이 감소할 것이다.
- 본 논문에서 인간 노동의 본질로 여가를 주장하며, 노동과 창의성과의 연결고리를 새롭게 제시한다

SESSION 5

Zhouyi's Biantong World and the Algorithm of Yin-yang from the Dimension of AI Connectionism

Yonjae Kim National Gongju University(Korea)

Abstract

본고에서는 인공지능의 신경계통과 인간사회의 관계망의 관계에 착안하여 『주역』에서 제시된 상황인식의 문제에 접근하고자 한다. 특히 인공지능의 연결주의의 특징을 사회공학적 차원으로 확충하여 『주역』의 變通의 원리와 그 속에 반영된 상황의 인식과 판단의 방법론을 집중적으로 논한다.

우선, 인공지능의 신경계통에서 학습의 방식은 연결주의(connectionism)로 특징지워질 수 있다. 연결주의는 뉴런과 같은 중추신경계에서 진행되는 처리의 과정을 추상화한 의사 뉴런의 체계를 가리킨다. 여기에는 자극과 반응의 연결을 통해 사례의 경험에 따른 학습의 방식이 있다. 이러한 방식을 사회공학적으로 확충하면 인간사회가 운용되는 방식과 일정한 관련성을 지닌다. 『주역』의 세계관에서는 사회의 관계망 속에서 상황을 인식하는 과정을 變通의 원리로 설명한다. 變通의 원리는 인간사회의 구조망(an interactive system of networks)을 상황의 인식과 판단의 차원을 지닌다. 이른바 變通의 과정은 인식의 변곡점(窮), 시선의 전환(變), 관계의 소통(通), 결과의 지속성(久)으로 전개된다. 그것은 인간사회의 관계망에서 변화와 안정의 상관성을 통해 진화하는 사회공학적 차원을 함축적으로 반영한다. 이러한 진화의 과정에는 陰陽의 알고리즘(algorithm)이 작동하고 있다. '한번 음이 되면 한번 양이 되는(一陰一陽)' 법칙은 同化의 구심력과 異化의 원심력을 통한 상호작용을 대변한다. 음양의 알고리즘은 변화하는 상황을 이해하고 파악하는 일련의 자율적 계통의 방식 혹은 방법을 가리킨다. 이러한 상호작용을 통한 사회적 관계망은 인공지능의 연결주의적 차원과 일맥상통한다.

이러한 맥락에서 『주역』의 變通의 원리는 세계를 이해하는 과정과 세계를 창조하는 과정이 맞물리는 차원은 인식과 실천의 통합적 관점을 지닌다. 여기에는 인간사회가 자기조정능력을 갖춘 자율적 체계에 초점을 맞추어 지속가능한 사회이념을 지도하고 계도하는 것이다. 특히 이는 인공지능의 자율적 조절 방식에 착안한 연결주의의 차원에서 사회공학적 방법론의 일환이라고 말할 수 있다.

SESSION 5

Zhouyi's Biantong World and the Algorithm of Yin-yang from the Dimension of AI Connectionism

Yonjae Kim National Gongju University(Korea)

AI의 연결주의적 차원에서 본 주역의 變通的 세계와 陰陽의 알고리즘

- I. 문제의식의 실타래
- II. 정보기술사회와 AI의 연결주의적 시선
- III. 주역의 變通觀과 세계의 실재
- IV. 天地의 연계망과 공동체의식의 동일선
- V. 陰陽의 연결망과 알고리즘의 연속선
- VI. 乾坤의 관계망과 통합적 접속의 연장선
- VII. 문제해결의 실마리

I. 문제의식의 실타래

최근에 4차 산업혁명의 시대를 앞두고 주도적인 첨단기술은 어느덧 정보기술(IT)의 단계에서 자료기술(DT)의 단계로 신속히 이행되고 있다. 이른바 '빅데이터(Big Data)'라는 말은 지식정보사회에서 지능정보사회로 진화하는 단면을 보여주고 있다. 이러한 정보기술의 융합화로 나아가는 과정의 중심에는 과학기술의 고도의 집적화에 따른 인공지능과 그 원리 및 그 활용이 있다.

인공지능은 사고, 학습 등과 같은 인간의 지적 능력을 구현한 기술의 체계이다. 여기에는 자극과 반응의 연결을 통해 경험적 사례를 결합하고 조합하는 일련의 축적의 과정이 있다. 그 과정에는 인간의 신경계통처럼 상호 연계적으로 통합적 체계를 구성하는 접속의 방식이 작동한다. 이 방식에서 통합적 체계에 접근하는 방법을 연결주의(connectionism)라고 부른다. 지능적 체계에서 연결주의는 뉴런과 같은 중추신경계에서 진행되는 정보처리의 과정에 관한 의사뉴런의 체계를 대변한다. 이러한 연결주의의 접속은 고도로 직접화되고

SESSION 5

Zhouyi's Biantong World and the Algorithm of Yin-yang from the Dimension of AI Connectionism

Yonjae Kim National Gongju University(Korea)

다원화된 사회일수록 사회공학적 차원에서 더욱 중요한 유기적 전체의 기제로 작용한다.

이러한 맥락에서 본고에서는 인공지능의 특성을 인문학적 차원에서 어떻게 접근할 것인가 하는 문제의식에서 출발한다. 특히 지능적 체계에서 연결주의의 방법은 『주역』의 세계관에서 사회공학적 통로로서 變通의 원리에 해당한다. 變通의 원리는 자연계의 운행질서를 인간사회의 흐름에서 적용한 결과이다. 그것은 상호 연계적으로 통합하는 체계와 그 흐름에 관한 발상에서 나온 것으로서, 다원화되고 다양한 사회에서 복잡다단한 관계 속에서 유기적 전체의 흐름을 집약적으로 표현한 것이다. 이러한 체계는 天地의 연계망, 陰陽의 연결망 및 乾坤의 관계망으로 특징화될 수 있다. 이는 통합적 질서의 역동적인 노선에서 작동한다. 즉 天地의 연계망 속에서 공동체의식의 동일선으로, 陰陽의 연결망 속에서 알고리즘(algorithm)의 연속선으로, 乾坤의 관계망 속에서 통합적 접속의 연장선으로 진행된다.

특히 음양의 방식은 天地의 존재론적 범주와 乾坤의 가치론적 범주 사이에 알고리즘으로 작동한다. 64괘는 세계의 실재(reality)를 인식하는 출입구로서, 현실 세계(real world)와 가상 세계(virtual world) 사이의 경계에서 음양의 범주적 매체를 통해 복잡다단하고도 변화무쌍한 역동적 관계를 만들어갈 수 있다. 따라서 음양의 알고리즘은 유기적 전체의 통합적 흐름에서 작동하는 연결주의의 접속방식으로서, 개별적인 경계를 넘어서 통합적 경계를 지향하는 매체의 기제가 된다.

II. 정보기술사회와 AI의 연결주의적 시선

세계는 자연계의 그물망에서 삼라만상의 스펙트럼을 투사한 프리즘과 같은 세상의 경계이다. 세계의 실재(reality)는 인간의 삶 속에서 자연계에 대한 입체적 조망을 통해 파악된다. 그것은 자연계의 그물망에서 생명력의 자생적 율동이 작동하는 삼라만상이 생성하고 변화하는 일련의 과정이다. 인간은 자신의 관점이나 믿음의 체계에 따라 세계를 새롭게 경험할 수 있다. 세계는 주체의 입장에서 세상을 보는 경계이고 세계관은 세계를 투사하는 일종의

SESSION 5

Zhouyi's Biantong World and the Algorithm of Yin-yang from the Dimension of AI Connectionism

Yonjae Kim National Gongju University(Korea)

프리즘과 같은 인간의 시선이다. 이러한 주관적 시선을 통해 세계의 인식에는 현실과 가상의 간격 혹은 거리감이 생기게 마련이며, 삶의 역정은 이것을 매우 고 보충하면서 진행되기 마련이다.

오늘날 고도의 산업사회에서 우리는 기술의 집적화와 그에 따른 정보의 고도의 기술화를 통해 삶의 다양성을 누릴 수 있게 되었다. 이러한 사회에서는 혼돈과 안정, 불변과 가변, 결정과 과정 등의 사이에서 변화무쌍하게 움직이면서도 시의적절하게 적응해가야 한다. 이러한 구조에서는 주체가 객체와의 격차나 괴리를 극복하지 못할 때에 그 양자 사이에 거리감이 존재한다. 특히 세계의 실재에서 현실적인 限界와 이상적인 境界 사이에 존재하는 특정의 경계가 생기게 마련이다. 경계는 사회공학적으로 전체와 부분, 변화와 안정, 통합과 분화, 편향과 균형, 갈등과 협동 등의 구분이나 차이 속에서 그 양자를 해소하거나 극복하는 내용까지도 포함한다. 이러한 경계의 거리감을 조정하고 해소하고 과정에서 의식의 세계가 형성된다. 의식의 세계에서는 현실과 가상이 공존하며 이들 사이에 동일성과 차이성의 역동적 관계에 따라 시선의 다양한 스펙트럼을 펼친다.

우리는 정보의 고도의 기술화를 통해 현실로 존재하는 세계의 실재를 가상으로 설정할 수 있다. 가상의 세계는 현실의 세계와 밀접하게 관련되며 우리가 현실적 삶을 주도적으로 개척하는 데에 중요한 단서를 제공한다. 현실에서 충족되지 못한 부분을 가상을 통해 보완하거나 완성한다. 예를 들어, 인터넷을 통해 현실적 가상이나 가상적 현실을 만들어갈 수 있다. 삶의 방식에서 online 과 offline의 관계가 서로 구분되거나 서로 분리되지 않는다. 이처럼 가상의 세계를 과학기술적으로 고도로 체계화됨으로써 가상의 세계와 현실의 세계는 서로 친밀하여 경계짓기 어렵게 되었다. 이러한 불확실한 경계에도 가상과 현실이 실현의 연속선에서 만나는 탈경계적 성격을 지닌 통합적 경계가 있다. 또한

¹⁾ 경계는 주체와 객체의 일정한 관계 하에서 두 가지 이중적 의미를 지닌다. 자아와 타자를 구분하는 차이의 의미를 지니면서 또한 그 양자의 구분을 아우르는 의미도 지닌다. 전자가 인식의 대상으로서의 경계이고 후자가 가치의 대상으로의 경계이다. 전자는 주체와 객체의 일정한 관계를 인지하고 파악하는 단계인 반면에, 후자는 그 특정의 관계를 구현하고 실현하는 단계이다. 전자의 경우에는 물리적, 생물적, 생태적, 사회적, 문화적, 윤리적 등과 같은 포괄적인 경계도 있으며 생명의 경계, 인간의 경계, 지역의 경계 등과 같은 구체적인 경계도 있다. 후자의 경우는 대체로 소통, 융합, 통합 등을 특징으로 하는 정신적 혹은 관념적 성격을 지닌다. 이 경계는 경지 혹은 지명이라고도 불린다. 세계관이란 바로 이러한 이중적 경계를 조명하는 시선의투사면이다. 자연계는 시선의 집중적 투사로 조망되어 통일적 양상의 현상으로 비춰진다. 이 통일적 현상이 인간의 의식에서 세계의 질서의식으로 자리잡는다.

SESSION 5

Zhouyi's Biantong World and the Algorithm of Yin-yang from the Dimension of AI Connectionism

Yonjae Kim National Gongju University(Korea)

시간의 흐름에서 공간의 점유로 확장되는 일련의 과정 속에서 자아실현을 위한 최적화된, 가상과 현실이 합치되는 경계가 신축적으로 설정되고 성취될 수 있다.

가상 세계 혹은 가상적 현실 세계는 과학의 기술적 장치를 통해 실제적인 효과를 발휘하는 이미지의 실재이다. 그것은 인간의 삶과 관련하여 현실의 사물, 사건, 상황 등 및 이들의 공간적 관계로 구성된다. 여기에서는 인간 삶에 관한 가능한 많은 정보들이 가상의 상태에서 현실화되고 재구성된다. 이는 합성, 분해, 강조, 배열, 삭제, 보충, 변형 등을 거쳐서 특정의 세계를 구성하는 것이다. 즉 지각의 체계, 기호 혹은 부호의 방식, 의미의 표현 등과 같은 추상화의 과정을 거쳐서 또 다른 세계의 실재를 구성한다.

그러므로 가상 세계는 인간의 지각을 통해 수용되는 현실적 성격을 지닌 또 다른 세계의 실재이다. 그것은 현실 세계의 완결된 모사라기보다는 현실의 가능성을 담고 있으며 따라서 잠재적 실재로서 열린 체계를 지닌다. 우리는 시뮬레이션(simulation), 원격현전(telepresence), 몰입(immersion) 등의 체계를 통해 가상에 관한 독특한 체험을 한다. 이러한 체험에서는 가상의 이미지가 수용자의 의식이나 몸과 물리적 거리의 간격이 없어지며 더 나아가 이들 사이의 경계를 넘나들며 직접적으로 상호작용을 하게 된다. 가장 중요한 점은 수용자는 가상을 통해 자신이 머물고 있는 물리적 세계를 벗어날 수 있으며 또한 자신의 물리적 세계와 가상의 이미지들을 결합하여 실재성을 더욱 확고히 할 수 있다.²⁾

가상은 현실과의 밀접한 관계를 지니며 실재성이나 현실성을 갖는다. 가상이 현실이 될 수 있고 현실은 가상처럼 구성된다. 가상과 현실의 경계에서는 시공간성의 可逆性이 있다. 이 경계는 프리즘의 초점이 맞추어지는 지점이다. 프리즘의 방식은 관찰자의 시점에 따른 것이다. 자연계에 있는 사물, 현상, 사건 등은

²⁾ 이은아, 「몸-두뇌 성취로서의 가상현실 체험과 지각의 확장 가능성에 대한 고찰」, 『미학』 제84권 3호(2018), 270-272쪽

SESSION 5

Zhouyi's Biantong World and the Algorithm of Yin-yang from the Dimension of AI Connectionism

Yonjae Kim National Gongju University(Korea)

이러한 방식으로 드러난다. 세계는 시선의 경계만큼 설정되고 만들어가는 것이다. 그러므로 현실 세계가 시선의 프리즘을 통해 삼라만상의 다양성이나 복합성을 투영한 세계의 실재라면 가상 세계도 이러한 성격을 지닌 세계의 실재이다.

그러므로 가상 세계는 현실세계와 달리 가변적이고도 가역적인 의식적 차원을 지닌다. 우리는 가상 세계를 통해 현실의 세계에 존재하는 불가항력성과 예측불가능성의 양립가능성을 이해하고 극복할 수 있다. 이러한 의식의 활동이 없다면 삶, 인간관계, 역사, 업적 등과 같은 인간 삶의 지식을 쌓고 그 속에서 지혜를 터득하기는 쉽지 않다.

우리의 세계관에서 프리즘의 초점이 삶의 실현과 맞물릴 때 가상과 현실의 경계를 넘어설 수 있다. 경계에는 임계점의 기로에서 소통과 변화를 모색하는 과정이 있다. 혼돈에서 질서로 나아가는 기로가 임계점이며 이 지점이 프리즘의 초점이 맞추어지는 지점이다. 가상은 초점이 잡힌 프리즘이 투사될 때마다 현실성을 갖는다. 현실성은 표상, 예시, 예측 등의 복합적인 내용으로 이루어진다. 이 현실성을 지닌 가상이 삶의 실현과 맞물릴 때에 현실로서 여겨진다. 이제 세계의 실재는 가상과 현실이 중첩되면서 다양한 프리즘의 다중적 초점이 맞추어진, 고도로 집적화된 세계이다.

예를 들어, 현대에 고도의 정보기술화의 역량을 대변하는 과학기술의 집적체가 바로 인공지능이다. 최근에 인공지능의 역할을 상징적으로 함축하는 용어가 등장하였다. 그것은 메타버스(metaverse)이다. 이는 초월적인 의미로서의 메타(meta)와 우주의 의미로서의 유니버스(universe)를 합성한 신조어다. 이 용어는 그것은 현실의 세계처럼 인간의 활동 전반, 즉 사회적, 경제적, 문화적 활동이 가능한 3차원의 가상적 공간이다. 그것은 미디어의 혁신적 기술력과 그에 따른 인터넷의 공간이 확보됨에 새로운 세상을 여는 유토피아적인 플랫폼이다. 여기에는 인간의 의식적 차원에서 연상작용과 상상력을 거쳐서 형성되는 시공간적 통합적 경계가 있다. 가상과 현실 사이에 개별적인 경계가 상호작용하면서 그 통합적 경계를 지향하는 공간의 영역이다. 그것은 현실 세계와 가상 세계를 연결하고 통합함으로써 새로운 사회적 문화적 생태적 환경과 그 속

SESSION 5

Zhouyi's Biantong World and the Algorithm of Yin-yang from the Dimension of AI Connectionism

Yonjae Kim National Gongju University(Korea)

속에서 새로운 가치관을 창출하고 있다.

이러한 현실과 가상의 세계를 접속하는 데에 인공지능의 체계에서 정보의 기술적 처리방식으로서 연결주의(connectionism)의 접속이 있다. 연결주의는 뉴런과 같은 중추신경계에서 진행되는 처리의 과정을 추상화한 의사 뉴런의 체계를 가리킨다. 여기에는 자극과 반응의 연결을 통해 사례의 경험에 따른 학습의 방식이 있다. 이러한 지능체의 방식은 사례의 경험을 반복하여 다수의 의사 뉴런이 서로 결합하고 그 결합을 통해 경험의 값(매개 변수)을 점차적으로 변화시켜가는 것이다. 여기에는 기호나 숫자의 나열이나 확장이 있으며 더 나아가 이들과 다른 의식적 행위의 방식도 있다.

III. 주역의 變通觀과 세계의 실재

동아시아의 지역에서 『주역』은 지난 3,000여 년간 占術의 영역으로부터 철학의 영역으로 거듭나면서 문명사회의 사상적 혹은 문화적 원류이자 원천이 되어왔다. 『주역』의 圖像에는 시간적 단절을 이어붙이고 공간적 여백을 채워 놓는 인간 삶의 다채로운 굴곡이 투영되어있다. 이러한 시공간적 圖像에는 변화하는 세계에 대한 憂患意識을 갖고서 세계의 변화를 개척하는 삶의 모습이 있다. 여기에는 세계의 실재를 조망하는 독특한 방식이 있다. 그것은 존재의 변화가능성으로부터 가치의 지속가능성을 모색하는 것이다. 變通의 원리는 이러한 방식의 결정체로서 인간이 공동체적 의식 속에서 삶의 통합적 경계를 어떻게 지향해야 하는가 하는 문제의식을 단적으로 시사한다.

變通의 원리는 易學的 사유의 강령으로서 사회공학적 차원에서 공동체적 의식의 통합적 경계가 무엇인지를 단적으로 시사한다. 『주역』의 세계관에서는 세계를 이해하는 과정과 세계를 창조하는 과정이 맞물리는 차원은 인식과 실천의 통합적 관점에서 조망될 수 있다. 여기에는 인간사회가 자기조정능력을 갖춘 자율적 체계에 초점을 맞추어 지속가능한 사회이념을 지도하고 계도하는 것이다. 특히 이는 인간사회의 관리와 윤리의 자율적 조절 방식과 관련된다는 점에서 사회공학적 한계를 극복할 수 있는 방법론적 발상이다.

SESSION 5

Zhouyi's Biantong World and the Algorithm of Yin-yang from the Dimension of AI Connectionism

Yonjae Kim National Gongju University(Korea)

우선, 변통이라는 말은 세계의 變化와 疏通의 관계를 함축한다. 그것은 세계가 변화하는 과정을 통찰하는 입체적 시선과 관련된다. 『역전』에서는 다음과 같이 말한다.

역은 다하면 변하고 변하면 통하며 통하면 지속한다. 그러므로 하늘로부터 도와서 길하고 이롭지 않음이 없다.³⁾

여기에서는 특정의 상황이 진행되는 일련의 과정을 '窮-變-通-久'의 과정으로 표현한다. '窮-變'은 변화와 균형의 단계라면 '通-久'는 안정과 지속의 단계이다. 그 양자의 과정은 사회적 복잡다단한 구조에서 발생하는 상황의 時機에 따라 분화와 통합의 역동적 과정을 함축적으로 반영한 것이다. 즉 끊임없이 분화되었다가 끊임없이 통합되기도 하며 끊임없이 통합되었다가 또다시 끊임없이 분화되기도 한다. 이러한 과정이 다양성, 다원성, 복잡성 등의 각양각색 혹은 천차만별로 펼쳐지는 세계의 실재이다.

이러한 '窮變通久'의 과정은 變으로부터 通으로 나아가는 과정과 通으로부터 變으로 나아가는 과정으로 구분된다. 전자가 객관적 상황을 합리적으로 인식하는 과정이라면, 후자는 이러한 상황을 주체적으로 개척하는 과정이다. 그러므로 변통과 통변의 역동적 과정은 서로 맞물린 톱니바퀴처럼 함께 굴러간다. 변통이 인식의 포용성이나 포괄성을 특징으로 한다면 통변은 실천의 융통성 혹은 적응성을 특징으로 한다.

이러한 맥락에서 보자면, '窮變通久'의 과정은 존재의 변화가능성에서 어떻게 가치의 지속가능성을 확보할 수 있는가 하는 문제의식 담겨있다. 이는 인식과 실천의 연속선상에서 變通과 通變의 통합적 삶을 통해 공동체의 가치론적 영역을 넓히는 문제와 관련된다. 이러한 공동체적 영역에는 인간사회의 지속가능한 경계가 반영되어 있다. 지속가능한 경계란 대립과 화합, 분별과 통합, 직분과 조정 사이에 적절하게 조화를 이루는 궁극적 단계를 가리킨다. 이러한

³⁾ 『周易』, 「繫辭下傳」, “易, 窮則變, 變則通, 通則久, 是以自天祐之, 吉無不利.”

SESSION 5

Zhouyi's Biantong World and the Algorithm of Yin-yang from the Dimension of AI Connectionism

 Yonjae Kim National Gongju University(Korea)

단계에서 사회 전체는 다원적으로 분화하여 개인의 구성원을 형성하며 개인의 구성원은 대립과 조화의 과정을 거쳐 사회 전체로 통합되기 마련이다. 여기에서 인식론적 실천력에 따른 안정과 지속성의 통합적 경계가 있다. 이러한 통합적 경계는 '안정의 지속성'을 끊임없이 유지하면서도 '지속의 안정성'을 끊임없이 계속한다. 이러한 경계에는 신비주의적인 불가지론도 없으며 불가항력적인 숙명론도 없는 것이다.

變通의 원리는 세상의 경계, 즉 세계에 대한 인간의 의식의 차원에서 나온 것이다. 자연계는 천체가 운행되는 방식에 맞게 모든 생물체가 활동하는 생명의 자생적 연결망과 그 유기적 흐름이다. 우리가 경험하는 세계는 자연계의 그물망에서 삼라만상의 스펙트럼을 투사한 프리즘과 같은 세상의 경계이다. 경계는 일정한 영역 혹은 분야를 명확하게 확정하지 않는 사이(the between) 혹은 구분을 가리킨다. 경계의 차원에는 개별적 경계와 통합적 경계가 있다. 인간 사회의 질서의식은 개별적 경계를 거쳐 통합적 경계를 지향한다. 특히 통합적 경계는 인간의 관계나 사회의 구조에서 대립, 모순, 상충, 갈등을 넘어서 통일, 조화, 해소, 화해로 나아가는 최적화된 발상에서 나온 것이다. 여기에서는 주체와 객체의 구분 속에 전체와 부분, 변화와 안정, 통합과 분화 등의 일정한 연관성에 주목하고 그 이분법적 경계를 해체하고 통합적인 관계를 만들어간다. 여기에는 관계의 구분이나 차이의 한계를 넘어선다는 의미에서 통합적 경계의 방향성을 지닌다.

이러한 세계의 실재(reality)에서 주목할만한 점은 인간과 자연의 내재적 관계이다. 이는 하늘과 땅의 틈 속에서 생명력의 자생적 율동이 작동하는 삶의 입체적 조망과 관련된다. 인간은 자연계의 生生不息의 과정 속에서 그에 맞게 自強不息의 의식을 갖고서 살아간다. 이러한 생명의 자생적 연결망과 그 유기적 흐름은 인간사회의 공동체적 의식 속에 반영되었다. 인간은 자신의 관점이나 믿음의 체계에 따라 세계를 새롭게 경험할 수 있다. 세계는 주체의 입장에서

4) 외재적 관계는 물리학적, 생물학적, 지리학적 등의 성격을 지닌다면 내재적 관계는 인식론적, 생태주의적 등의 성격을 지닌다. 그 양자 중에서 내재적 관계가 더 근본적이다. 喬清攀, 『儒家生態思想通論』, 21-24쪽.

SESSION 5

Zhouyi's Biantong World and the Algorithm of Yin-yang from the Dimension of AI Connectionism

Yonjae Kim National Gongju University(Korea)

세상을 보는 경계이고 세계관은 세계를 투사하는 일종의 프리즘의 시야이다. 자연계의 복잡다단한 계층과 그 위계적 질서 속에서 인간은 스스로 서로 소통하고 공생하고 통합하는 조화로운 삶을 개척해가는 것이다.

이러한 경계는 삶의 주관성과 객관성의 관계를 이해하는 문제와 관련된다. 인간의 삶은 어떻게 대립과 모순의 불안정한 관계에서 통일과 조화의 안정적인 관계로, 더 나아가 대립과 통일, 모순과 조화의 불확정적인 관계가 어떻게 유기적인 통합의 확정적 관계로 나아갈 수 있는가 하는 과정이다. 세계의 실재에서 현재의 인간은 미지의 미래를 향해 변화무쌍한 삶을 살아가는 반면에 과거의 삶을 거울삼아 미래의 삶을 담보할 수 있다. 여기에서 삶의 불확정성과 확실성의 관계가 서로 모순되는 것 같지만 결국에 삶의 역동적 과정에서 서로 해소된다. 인간의 삶이란 자아실현의 과정으로서, 어떠한 상황이나 단계에서도 최종적으로 완전히 성취되거나 완결될 수 없다. 삶의 과정에서 발생하는 대립과 갈등이 적절하게 해소되어야 하지만 해소라는 것도 또 다른 대립과 갈등의 상황으로 전개되는 일련의 과정일 뿐이지 그러한 상황이나 단계가 완전히 마무리되는 최종적인 단계일 수 없다.

이러한 통합적 의식적 차원을 사회공학적으로 확충하면 인간사회가 운용되는 행위의 방식과 관련된다. 이러한 행위의 방식은 주역의 세계관, 그 중에서 變通의 원리에서 접근될 수 있다. 變通의 원리는 인간사회의 구조망(an interactive system of networks)을 상황의 인식과 판단의 차원에서 접근한 것이다. 窮-變-通-久의 과정은 인식의 변곡점(窮), 시선의 전환(變), 관계의 소통(通), 결과의 지속성(久)으로 설명될 수 있다. 그것은 인간사회의 관계망에서 변화와 안정의 상관성을 통해 진화하는 질서의식의 차원을 함축적으로 반영한다.

일반적으로 인간의 관계나 사회의 구조에서 주체가 객체와의 격차나 괴리를 극복하지 못할 때에 그 양자 사이의 경계가 생기게 마련이다. 경계란 사회적 구조에서 전체와 부분, 변화와 안정, 통합과 분화, 편향과 균형, 갈등과 협동 등의 구분이나 차이를 가리킨다. 이러한 경계에는 상호작용을 통해 각각의 한계를 넘어서는 과정이 진행된다. 그것은 공동체의 복잡다단한 현상들과 이들의 변화 속에서 대립, 모순, 상충, 갈등을 넘어서 통일, 조화, 해소, 화해로

SESSION 5

Zhouyi's Biantong World and the Algorithm of Yin-yang from the Dimension of AI Connectionism

Yonjae Kim National Gongju University(Korea)

전환되는 최적화된 발상의 능력이다.

이러한 관점은 『주역』에서 觀卦의 강령으로서 大觀으로 표현된다. 그것은 인간이 관찰, 인식 및 이해라는 일련의 과정을 통해 세계를 조망하는 방법이다. '觀'은 사물의 형식과 상황을 이해하고 체험하는 감각활동으로서 단순히 인간이 사물을 물리적 대상으로 고찰하는 것만이 아니다. 그것은 주체와 객체를 동일선상에 놓는 입체적인 鳥瞰의 방식으로 세계에 대한 인식에 따른 실천의 존재론적 영역까지도 조망하는 다차원적 인식의 출발점이다. 인간은 삼라만상을 관찰하면서 끊임없이 체험하면서 자신의 세계를 형성해가고 더 나아가 자신의 모습을 성찰하고 주체적 행위를 통해 실천해가는 포괄적인 시선을 지닌다.

大觀의 시선에서는 삶의 방식을 이해하여 세계의 실재에서 부딪치는 주체와 객체의 괴리, 즉 대립, 갈등, 모순 등의 문제를 해소하려는 것이다. 변통이 세계의 변화를 지속하는 역동적 과정을 가리킨다면 대관은 이러한 과정을 통찰하는 입체적 관점을 가리킨다. 이에 관해 「단전」에서는 다음과 같이 해석한다.

大觀(크게 보는 것)이 위에 있어 따르고 공손하여 中正함으로써 천하를 살핀다.⁵⁾

구오효의 속성인 양의 강건함과 中正은 군자의 자리를 가리키는 것으로서⁶⁾, '大觀(크게 보는 것)'의 위치를 상징한다. "크게 보는 것"이라는 말은 공명 정대한 태도로써 특정의 대상 혹은 세상을 관찰하는 것을 말한다.

대관의 입체적 시선은 성찰과 통찰의 입체적 시각을 통해 변통의 지속적 과정을 이해하고 인간과 세계, 주체와 객체를 구획하는 개별적 경계를 해소하는 통로가 된다. 여기에 인간이 스스로 세계를 관찰하면서 체험하는 창조적 소통을 위한 인식론적 실천력의 기제가 작동한다. 즉 주체가 대상을 바라보는

⁵⁾ 『周易』, 「彖傳」, 觀卦, “大觀在上, 順而巽, 中正以觀天下.”

⁶⁾ 이는 구오효의 위상에 입각하여 해석한 말이다. 관괘는 두 개의 양효가 위에 있고 네 개의 음효가 아래에 있는 형상이다. 관괘는 음의 세력이 커지고 양의 세력이 줄어드는 괘이다. 양효가 아직 높은 위치에 있어 아래를 내려다볼 수 있는 반면에, 음효가 비록 네 번째 효까지 커졌지만 아직 아래에 머물고 있으므로 반드시 위를 우러러보아야 한다. 오효 아래의 네 가지 효들이 모두 음의 효이므로 양효가 아래에 있는 네 가지 음효를 거느리는 형상을 취한다. 특히 오효는 상괘의 중앙에 있는 양효이면서도 양효가 양효의 자리에 있으므로 '中'과 '正', 즉 '중정'의 가장 합당한 위치에 있다. 또한 이러한 의미에서 구오효는 하나의 괘 전체에서 가장 강건하고도 존귀한 이른바 중정의 지위를 갖추고 있다.

SESSION 5

Zhouyi's Biantong World and the Algorithm of Yin-yang from the Dimension of AI Connectionism

Yonjae Kim National Gongju University(Korea)

것뿐만 아니라 자신의 내면적 모습도 돌아보고 이를 남에게도 드러내는 것이다. 우리는 성찰과 통찰의 연속선상에 있는 大觀의 입체적 시선은 이른바 자아 실현의 이정표를 조감할 수 있는 통합적 경계에 도달할 수 있다.

인간은 세계의 실재에 자신을 끊임없이 새롭게 투사하는 과정에서 자신을 계도하고 발전시키는 창조적 삶의 조감도를 그려나갈 수 있다. 여기에서 인간은 대상을 관찰하고, 이러한 관찰의 경험을 바탕으로 하여 자신의 입장을 돌아보아 내면의 인격적 함양에 힘쓰며 자신의 참다운 모습을 보여 줌으로써 결국에 대상과의 관계를 지속한다. 그 과정에서 창조적 소통의 인식론적 실천력은 주체와 객체의 합일을 지향하는 성숙한 삶을 계도한다.

IV. 天地의 연계망과 공동체의식의 동일선

우선, 전통적으로 동아시아의 사상과 문화에는 모종의 크나큰 기초가 있다. 이것이 바로 天人合一의 이념이다. 天人合一은 인간이 자연계 속에서 살아가는 삶의 보편적 이념이다. 그것은 자연계라는 생명의 연결망 속에 삼라만상의 스펙트럼을 생명력의 통일적 질서의 계통으로 의식화한 결과이다. 인간이 경험하는 세계의 실재는 생명력의 자생적 율동[生生不息]이 작동하는 자연계에서 주체적 자각의식[自強不息]을 갖고서 살아가는 삶의 입체적 과정이다. 인간과 자연의 통합적 경계는 天人合一의 의식적 차원으로 수용된다. 특히 合一의 경계는 자연계에서 생명력의 자생적 율동과 그 유기적 연결망에 가치의 지향성을 부여한 결과이다. 여기에는 사유와 존재, 존재와 가치, 사실과 가치 등의 개별적 구분이 통합되는 의식의 차원이다. 그것은 경계의 구분을 넘어서는 통합적 경계로 열려져 있다.

『주역』은 인간의 삶을 전체적으로 조망하여 터득한 경험의 총체적 산물이다. 그 세계관은 天地의 복제판이자 자연계의 축소판이며 우주의 시공간적 도판(圖版)이다. 우리가 경험하는 세계는 자연계의 그물망에서 삼라만상의 스펙트럼을 투사한 프리즘과 같은 세상의 경계이다. 여기에는 생명력의 자생적 율동 속에서 만물이 생성하고 성장하고 소멸하는 일련의 과정이 반영되어 있다. 이 과정은 부분에서 전체로, 단편에서 전면으로, 단순함에서 복잡함으로,

SESSION 5

Zhouyi's Biantong World and the Algorithm of Yin-yang from the Dimension of AI Connectionism

Yonjae Kim National Gongju University(Korea)

조잡함에서 정밀함으로, 내면에서 외면으로 확충되는 입체적이고도 유기체적인 과정이다.

『주역』의 세계관에서 세계의 실재는 어떻게 접근될 수 있는가? 관건은 天地의 틀이다. 천지는 자연계의 대표적인 현상으로서, 생명의 연결망 속에서 만물의 造化와 그 역동적 과정에 관한 정보의 매체이다. 여기에는 삼라만상이 존재하는 방식뿐만 아니라 이들을 이해하는 방식도 담겨있다. 「계사전」에서는 다음과 같이 말한다.

하늘과 땅의 크나큰 덕은 생명(生)이라고 말한다.⁷⁾

여기에서 생명(生)⁸⁾은 자연의 원초적 결정체를 가리킨다. 생명은 유기체의 내재적 본질 혹은 본성으로 특징지을 수 있다. 생명의 연결망은 하늘과 땅의 틀로 대변되고 더 나아가 인간의 본성에서 생명의식으로 자리잡는다. 만물의 造化와 그 변화의 과정에 관해 「계사전」에서는 다음과 같이 설명한다.

하늘과 땅이 기운을 쌓고 합하여 만물이 응어리지며 남성과 여성이 정미함을 얻어내어 만물이 생겨나온다.⁹⁾

여기에서는 生氣의 흐름과 같은 자연계의 순환과정을 진행하며, 그 속에서 남성과 여성의 관계를 통해 精氣의 조화의 양상으로 드러난다. 자연계는 수많은 생명체들이 서로 生氣의 흐름과 精氣의 결정체를 통한 생명체의 물질과 에너지가 유동하는 거대한 순환체계라고 말할 수 있다. 그러므로 세계의 실재(reality) 자체가 유기체적인 그물망의 관계로 이루어진 일련의 과정이다. 그것은 만물이 생성하고 성장하고 활동하여 소멸하는 생명력의 과정으로서 존재의 연속성(the continuity of being)을 특징으로 한다.¹⁰⁾ 여기에는 생명과 창조의 통합적 경계를 지향한다. 그 경계에서는 생명의 창조력을 끊임없이 구성하면서 창조의 생명력을 끊임없이 만들어낸다. 이러한 맥락에서 “하늘에서는 형상

⁷⁾ 『周易』, 「繫辭下傳」, “天地之大德曰生.”

⁸⁾ 생명(生)이라는 말에는 만물의 성장, 자연의 생명체, 생명성의 존재, 존재의 생명상태 등을 함축하고 있다. 인간의 경우에 生, 性 및 姓은 어원적으로나 용법상에서 동일한 의미를 지닌다. 朱良志, 「中國藝術的生命精神」, 3-5쪽.

⁹⁾ 『周易』, 「繫辭下傳」, “天地絪縕, 萬物化醇. 男女構精, 萬物化生.”

¹⁰⁾ Tu Wei-ming, *Confucian Thought: Selfhood as Creative Transformation* (Albany: State University of New York Press, 1985), 36-39쪽.

SESSION 5

Zhouyi's Biantong World and the Algorithm of Yin-yang from the Dimension of AI Connectionism

Yonjae Kim National Gongju University(Korea)

을 이루고 땅에서는 형체를 이루니 변화가 드러난다”¹¹⁾고 말한다. 그러므로 천지의 틀은 삼라만상의 스펙트럼을 조망한 투사면으로서, 자연계의 수많은 부분들이 서로 그물망처럼 얽혀있는 일정한 연관성의 체계이다. 이러한 연관성의 체계는 시공간성으로 특징화된다. 「서괘전」에는 다음과 같이 말한다.

하늘과 땅이 있는 다음에 만물이 생겨난다. 하늘과 땅 사이에 가득 찬 것은 오로지 만물뿐이다.¹²⁾

자연계에서 생명의 연관체계는 천지의 틀로 대변되며 그 속에서 생명력의 일정한 방향의 과정이 진행된다. 이러한 과정은 시공간적 방식으로 특징화된다. 만물이 생성되는 과정은 시간의 흐름으로, 만물이 존재하는 과정은 공간의 위치로 설명될 수 있다. 즉 생명력은 천지의 시간성으로 흐르고 생명의 활동은 천지의 공간성으로 확장된다. 여기에서 천지의 틀로 인식되는 생명과 실재의 통합적 경계가 있다.

여기에서 통합적 경계란 천지의 틀 속에서 진행되는 생명의 연관적 체계와 관련된다. 천지의 틀은 만물이 造化하는 과정, 즉 시간이 창출됨에 따라 공간이 펼쳐지는 개벽의 과정으로 이해된다. 이러한 개벽의 과정은 변화와 안정, 통합과 분화, 편향과 균형, 다양과 통일, 대립과 조화 등의 구분이나 구획을 넘어서는 방향성을 지닌다. 이러한 방향성은 생명의 연관적 체계에 바탕을 둔 유기적이면서도 전일적인 통합적 경계를 지향한다. 인간도 자연계에서 생명의 연결망의 일원으로서 만물과 일정하게 구분되면서도 이들과 공존하는 조화로운 삶을 살아간다. 이러한 삶은 자연과의 생태적 관계망에 관한 유기적 질서의 식에서 나온다. 이 관계망에서 존재의 본성상 感知 혹은 감응의 원초성을 통해 유지된다. 「단전」에서는 咸卦의 내용을 해석하면서 다음과 같이 말한다.

하늘과 땅이 감응하여 만물이 생겨난다. 성인이 사람의 마음을 감응하여 세상이 평화롭다. 그 감응한 곳을 관찰하니 하늘, 땅 및 만물의 실정을 볼 수 있다.¹³⁾

¹¹⁾ 『周易』, 「繫辭上傳」, “在天成象, 在地成形, 變化見矣.”

¹²⁾ 『周易』, 「序卦傳」, “有天地, 然後萬物生焉. 盈天地之間者, 唯萬物.”

¹³⁾ 『周易』, 「象傳」, 咸卦, “天地感而萬物化生, 聖人感人心而天下和平. 觀其所感而天地萬物之情可見矣.”

SESSION 5

Zhouyi's Biantong World and the Algorithm of Yin-yang from the Dimension of AI Connectionism

Yonjae Kim National Gongju University(Korea)

여기에서 '감응'은 생명력이 작동하는 방식을 가리킨다. 하늘과 땅의 틀은 생명의 존재론적인 성격을 지닌다. 그 틀 속에서 인간을 포함한 만물의 생명력이 발휘된다. 이러한 감응의 원초적 생명력은 사회적 교감, 조정 및 통합의 흐름 속에 공감력을 확충하고 유대감을 확장시킬 수 있는 공동체의식의 동일선상으로 지속된다.

V. 陰陽의 연결망과 알고리즘의 연속선

세계의 실재(reality)는 음양의 생명력의 기제가 작동하는 활동의 장(場)이다. 그것은 삶과 죽음의 굴레로 경험되고 시간이 공간을 열고 공간은 시간에 따라 전개되는 일련의 과정에서 체득된다. □ 주역□ 의 세계관에서 주목할만한 점은 이러한 연결망을 음양의 범주에 입각하여 접근한다. 즉 음양의 범주를 통해 존재의 동일성과 차이성 및 그 양자의 양상을 설명한다. 음양의 범주는 존재의 변화가능성을 설명하는 기제이자 가치의 지속가능성을 도출하는 통로이다. 우선, 「계사전」에서는 모든 존재의 생명력과 그 내재적 가치를 포괄적으로 정의하고 있다. 즉

한번 음이 되면 한번 양이 되는 것을 일러 道라고 말하고 그것을 잇는 것을 善이라고 말하고 그것을 이루는 것을 性이라고 말한다.¹⁴⁾

생명의 활력은 음과 양의 一의 방식으로 작동하는데, 생명의 지속성은 善의 가치로 인식되고 생명의 결정은 존재의 본성으로 자리잡는다. 여기에서 “한번 음이 되면 한번 양이 된다”는 표현은 생명력의 연결망 속에 최적화되는 역동적 경계를 특징적으로 설명한 것이다. 이처럼 최적화되는 과정은 음양의 易學的 황금률(Golden Rule)로 특징지을 수 있다. 음양의 황금률은 세계의 실재에서 만물이 끊임없이 율동하는 생명의 창발적(emergent) 원리를 가리킨다.

¹²⁾ 『周易』, 「序卦傳」, “有天地, 然後萬物生焉. 盈天地之間者, 唯萬物.”

SESSION 5

Zhouyi's Biantong World and the Algorithm of Yin-yang from the Dimension of AI Connectionism

Yonjae Kim National Gongju University(Korea)

天道の 운행의 과정에서 보자면, 어떠한 존재라도 그 고유한 성질을 지니며 이 성질이 본성을 형성하여 생성, 변화와 같은 존속의 기초가 된다. 이러한 생명의 존속과 관련한 권리나 이득이 존재의 생명력의 합목적성이 된다. 그러므로 천도의 원리는 사실과 가치의 합치를 내용으로 하는 생명력의 합목적성을 특징으로 한다. 이 합목적성이 바로 모든 존재의 善이 된다. 天人合一의 이념은 이러한 생명의 합목적성의 맥락에서 이해된다. 이러한 天도의 합목적적 과정은 생명의 연결망을 특징화한 것이다. 이러한 연결망의 지속성이 善의 온전성 혹은 완전성이라는 가치론적 내용을 지닌다. 이러한 인식론적 성격과 가치론적인 의의에서 인간을 포함한 만물의 본성은 자체의 완전성 혹은 온전성의 존재론적인 의미를 지닌 것으로 평가하는 것이다.

이러한 연결망에는 음양의 알고리즘(algorism)이 작동하고 있다. '한번 음이 되면 한번 양이 되는(一陰一陽)' 방식은 同化의 구심력과 異化의 원심력을 통한 상호작용을 대변한다. 음양의 알고리즘은 연결주의적 접속에서 변화와 안정의 연속선상에서 작용한다. 음과 양의 관계는 同化의 순화력과 異化의 반발력이 상호작용하는 과정으로 설명된다. 음/양 혹은 양/음과 같은 同化의 과정이 있고 음/음 혹은 양/양과 같은 異化의 과정이 있다. 이 과정에서 음과 양은 상호 전환하고 상호 변환하며 상호 포용하는 일련의 역동적 과정을 진행한다. 즉 동화의 과정이 없으면 이화의 과정도 없고 이화의 과정이 없으면 동화의 과정도 없다. 동화의 과정은 있으나 이화의 과정이 없다면 사물은 실 새 없이 발전하여 극단으로 치달아서 이들의 긍정적 효과가 부정적 효과로 바뀌게 될 것이다. 이와는 반대로 이화의 과정은 있으나 동화의 과정이 없다면 사물은 서로 압박이 지나쳐서 원동력을 상실하여 쇠퇴하여 소멸된다. 그러므로 그 양자 중에서 어느 한 쪽으로 지나친 쓸림은 극단적 대립, 모순 및 충돌 때문에 상대적 평형이나 균형을 깨뜨려 그 어느 쪽도 생존할 수 없는 결과를 낳을 수 있다. 상대적인 평형이나 균형을 유지하고 지속하기 위해서는 동화의 과정과 이화의 과정이 서로 끌어들이고 견제하며 서로 억제하고 도와주어야 한다. 이러한 방식으로써 음과 양의 관계는 이러한 역동적 과정을 통해 서로 상쇄되는 균형의 안정화된 국면을 만들어간다. 이러한 점에서 음과 양의 연결망은 공감력의 확충, 유대감의 확장, 공동체적 의식 등과 같이 연결주의의 사회공학적 접속에서 이해될 수 있다.

SESSION 5

Zhouyi's Biantong World and the Algorithm of Yin-yang from the Dimension of AI Connectionism

Yonjae Kim National Gongju University(Korea)

“한번 음이 되면 한번 양이 되는(一陰一陽)” 생명의 연결망에 따른 세계의 실재를 가리킨다. 음과 양의 관계에 따른 황금률(Golden Rule)은 음과 양의 양적 전환, 음과 양의 질적 변환, 및 음과 양의 포함에 따른 비정형의 흐름을 지닌다. 그러므로 음과 양 사이에 상호 전환되고 상호 변환되며 상호 포함하는 변동의 작용이 있다. 이러한 변동의 작용은 일정하게 상쇄되는 국면, 즉 음과 양의 균형과 같은 안정의 단계에 이른다. 매번 도달하는 이러한 단계들이 거듭 음과 양의 순환적 과정을 지속적으로 진행된다. 만물은 미시적 측면에서나 거시적 측면에서나 어떠한 관점에서든 모두 음과 양의 비례의 원칙에 따라 균형을 이루고 변화를 거듭하는, 낱고 낱아 멈추지 않는 과정을 겪게 마련이다. 왜냐하면 음과 양의 관계는 끊임없이 분화하고 끊임없이 교류하고 끊임없이 통합하며 또다시 반복하여 분화하고 교류하고 통합되는 복잡다단한 방식을 표상하기 때문이다.

이러한 전개는 논리가 바로 인간을 포함한 모든 만사만물의 존재론적 내용, 즉 세계 자체가 유기적인 완전한 전체의 성격, 즉 모종의 복잡다단한 연결망(web)의 구조를 형성한다. 이는 인간의 사유를 표현하는 광범위하고도 유효한 방식으로 확장되었으며 여기에 추상적 사유방식의 메커니즘이 존재하는 것이다. 이러한 메커니즘을 단계적으로 설명하면 다음과 같다.

첫째, 음과 양의 양상이다. 즉 음과 양 사이에 상호 변환되고 상호 전환되며 상호 함유하는 변동의 작용이 있다.

둘째, 이러한 변동의 작용은 일정하게 상쇄되는 국면 즉 음과 양의 균형과 같은 안정의 단계에 이른다.

셋째, 매번 도달하는 단계들이 거듭 음과 양의 순환적 과정을 지속적으로 진행된다.

이를 종합해보면, 하위의 계층에서 상호작용(음과 양의 흐름)의 단계적 발산을 거쳐 상위의 계층(태극의 본원)에서 조절과 조화의 수렴으로 나아감으로써 전체의 구조가 지속될 수 있는 것이다. 이러한 관계는 모든 단계에서 상호작용의 창발성과 자기조절의 항상성을 통해 전체의 구조가 형성되고 유지한다.

SESSION 5

Zhouyi's Biantong World and the Algorithm of Yin-yang from the Dimension of AI Connectionism

Yonjae Kim National Gongju University(Korea)

이러한 음양의 알고리즘은 연결주의의 차원에서 이해된다. 연결주의의 인문학적 차원은 인공지능의 신경계통과 인간사회의 관계망 사이에 일정한 연관성에 착안한 것이다. 변통의 사회공학적 시선에서 보자면, 세계의 실재는 음과 양의 알고리즘(algorism)의 작동하고 있다. 음양의 알고리즘은 변화하는 상황을 이해하고 파악하는 일련의 과정에 접근하는 방법론적 성격을 지닌다. 상황을 예측하고 판단하는 과정에는 변통의 원리와 같은 일련의 易學的 접근방식이 있다. 이러한 맥락에서 보자면, 음양의 연결고리와 그 접속의 방식은 복잡계적 원리, 시스템적 사고, 공진화(coevolution)의 과정, 프랙탈(fractal)의 강령 등과 같은 알고리즘의 방법론과 관련된다. 그러므로 음양의 범주와 그 연결망은 알고리즘에 따른 연결주의의 연속선에서 이해될 수 있는 것이다.

VI. 乾坤의 관계망과 통합적 접속의 연장선

『주역』은 인간 경험의 총체적 산물이라면 여기에는 현실의 세계를 투사한 가상 세계가 존재한다. 이 가상의 세계는 64괘로 대변된다. 64괘는 현실의 세계를 담아낸 가상의 세계에 해당한다. 이 세계는 삼라만상의 현실 세계와 밀접하게 관련된 또 다른 세계의 실재로서, 여기에는 음양의 연결고리에 따른 乾坤의 출입구가 있다.

『주역』에서는 64괘의 출입구를 통해 세계의 실재를 인식하고 삶을 실천해간다. 64괘는 인간사회의 양상을 64가지 경우로 나누고 이들의 일정한 관계를 설정한 것이다. 일정한 관계란 음양의 범주와 그에 따른 시공간적 방식에 기본한다. 음양의 범주는 주체와 객체의 유기적 관계를 설정하는 기제이다. 인간은 이 기체를 통해 시간의 흐름에 따라 공간이 전개되는 삶의 과정에 접근한다. 음양의 범주와 그에 따른 시공간적 방식은 인간의 삶에서 연결주의의 차원에서 접근될 수 있다. 세계의 실재를 인식하고 실천하는 과정에서 음과 양의 범주는 64괘의 가상 세계와 삼라만상의 현실 세계 사이에 존재하는 경계에서 그 양쪽의 관계를 접속하는 연결고리 혹은 출입구의 기제가 된다. 이러한 점에서 음양의 알고리즘은 접속의 연결주의적 차원에서 개별적 경계를 넘어서 통합적 경계로 나아가는 관문이 된다.

SESSION 5

Zhouyi's Biantong World and the Algorithm of Yin-yang from the Dimension of AI Connectionism

Yonjae Kim National Gongju University(Korea)

『주역』의 세계관에서 하늘과 땅의 틀은 세계의 실재(reality)로서 생명의 유기체적 연결망 속에서 일정한 통일적 질서를 유지하고 지속하는 과정으로 표현된다. 이 과정을 추상적 도상으로 기호화한 체계가 64괘의 卦象이다. 과상은 자연계에서 삼라만상을 관찰하고 그 속에 담긴 특징적 요소들을 부호로 추상화한 것이다. 그 중에서 乾卦와 坤卦의 관계는 하늘과 땅의 틀을 표상하여 세계의 실재를 인식하는 쌍방향의 관문이 된다. 『역전』에서 건과 곤의 범주는 음과 양 혹은 움직임과 고요함의 양상과 같은 상관적 관계를 통해 만사만물의 생성과 변화의 방식을 드러낸다. 『계사전』에서는 다음과 같이 말한다.

건과 곤은 역의 관문이 아닌가? 건은 양의 물건이고 곤은 음의 물건이니, 음과 양이 덕을 합하고 강과 유에 체가 있다. 그럼으로써 하늘과 땅의 일을 체득하며 신묘하고 밝은 덕을 통한다.¹⁵⁾

건과 곤은 『주역』의 논지, 즉 변화의 관문으로서, 하늘과 땅으로 대표되는 우주의 섭리를 밝히는 요체이다. 그것은 강과 유의 성질 혹은 음과 양의 덕성으로 발휘된다. 강과 유 혹은 음과 양은 건과 곤의 관계를 구성하는 통일적 방식으로서, 우주의 유기적 생명현상의 양상을 상징화한 것이다. 이는 인간이 건과 곤의 과상의 출입구를 통해 주체와 외부세계와의 관계성을 확보하고자 하는 것이다.

乾卦와 坤卦는 천지의 틀에서 진행되는 생명의 율동, 즉 움직임과 고요함의 과정을 원형과 네모형(天圓地方)으로 표상한 것이다. 건과 곤은 자연의 연관적 체계에서 생명의 크나름과 넓음으로 연상된 것이다. 건괘와 곤괘는 생명의 연관적 체계에서 부류의 동일성과 차이성에 따라 64괘의 총체적 관계를 구성한다. 곤의 기능은 닫는 것을 위주로 하고 건의 기능은 여는 것을 위주로 한다. 그러므로 건과 곤은 한 번 열리면 반드시 한 번 닫히는 연속적이고도 순一的인 과정을 상징한다. 이러한 과정은 자연계에서 진행되는 생명의 연관적 체계, 즉 삼라만상의 통일적 이치에 입각한 것이다. 그러므로 『계사전』에서는 다음과 같이 말한다.

¹⁵⁾ 『周易』, 『繫辭下傳』, “乾坤, 其易之門邪? 乾, 陽物也. 坤, 陰物也. 陰陽合德, 而剛柔有體, 以體天地之撰, 以通神明之德.

SESSION 5

Zhouyi's Biantong World and the Algorithm of Yin-yang from the Dimension of AI Connectionism

Yonjae Kim National Gongju University(Korea)

곤은 그 고요함이 닫히는 것이고 그 동은 열리는 것이므로 넓디넓은 생명이다¹⁷⁾

곤의 역량은 정과 동의 양상을 통해 닫힘과 열림의 성격을 드러낸다. 땅의 덕성은 고요하고 움직이는 과정에서 닫히고 열린다는 의미에서 수용성을 지니며 이것이 공간성의 영역으로 전개된다. 이것이 '넓디넓은 생명'이라는 창신(創新, creativity)의 방식이다. 그러므로 「계사전」은 그 당위성과 그 의의를 다음과 같이 단언한다.

건과 곤은 그 역이 간직한 것이지 않는가? 건과 곤이 이루어지니 번역이 그 가운데에 세워진다. 건과 곤이 훼손되면 번역을 드러낼 수 없다. 번역이 드러날 수 없으면 건과 곤은 혹시 거의 멈출지도 모른다¹⁸⁾

여기에서 세계의 실재에 대한 인간 인식의 결과를 '건과 곤의 道'로 고양시켰음을 알 수 있다. 건과 곤의 도는 인간이 세계의 실재에 대한 全一論的 인식을 통해 보다 관념화한 산물이다. 여기에서 세계의 실재는 '역'의 개념 속에 포괄적으로 함축되어 있고 건과 곤의 출입구로 표상된다. 그러므로 '역' 개념에 담긴 우주의 생명정신이 제대로 발휘되지 못하면 건과 곤의 표상적 구조는 기본적으로 성립될 수 없다. 이와는 반대로 건과 곤의 표상적 구조가 설정되지 않으면 '역' 개념에 담긴, 모종의 존재와 가치의 합일적 경계에 도달한 출입구는 없다.

『주역』의 세계관에서 세계의 실재는 생명력의 연결망 속에서 건과 곤과의 관계로 설전된다. 그것은 동일성과 차이성 사이의 창발적 연계성에 따라 천태만상, 각양각색, 천차만별, 다종다양의 분류의 관건이자 세계의 실재를 인식하는 쌍방향의 출입구가 된다. 그 관계에서 보자면, 개체가 생명을 유지할 수 있는 것은 외부의 세계와 완전히 차단된 폐쇄적인 닫힌 체계가 아니라 항상 외부의 세계로부터 새로운 에너지를 유입하는 개방적인 열린 체계이다.

여기에는 세계의 실재를 운영하는 집중력의 시선이 중요하다. 그것이 바로 변화하는 세계 속에서 세계를 변화시키는 능력, 즉 幾微의 능력과 관련된다.

¹⁷⁾『周易』,「繫辭下傳」,“夫坤,其靜也翕,其動也闢,是以廣生焉。”

¹⁸⁾『周易』,「繫辭下傳」,“乾坤,其易之緼邪?乾坤成列,而易立乎其中矣。乾坤毀,則無以見易。易不可見,則乾坤或几乎息矣。”

SESSION 5

Zhouyi's Biantong World and the Algorithm of Yin-yang from the Dimension of AI Connectionism

Yonjae Kim National Gongju University(Korea)

인간이 현재의 시간에 살면서 지나간 시간을 들여다보고 다가올 시간을 미리 살필 수 있는 것은 幾微의 인식론적 실천력을 지니기 때문이다. 「계사전」에서는 다음과 같이 말한다.

기미란 움직임이 은미한 것이니 길함이 먼저 나타나는 것이다. 군자는 기미를 보고 일을 하니 하루 종일 기다리지 않는다.¹⁹⁾

여기에서 기미를 파악하는 것은 삶의 변화에 대한 예측과 판단의 실용적 방법이다. 이러한 방법을 통해 인간은 현실의 실재(reality)에 대해 깊이 있게 해석하고 의미있게 탐색할 수 있다. 특정의 상황의 가능성으로부터 현실의 가능성을 드러내어 통찰할 수 있으며 그 크나큰 흐름을 파악할 수 있다. 그러므로 「계사전」에서는 다음과 같이 말한다.

역은 성인이 심오함을 모두 탐구하고 조짐을 연구한 것이다. 유독 심오하므로 천하의 뜻과 통할 수 있다. 유독 조짐이 있으므로 천하의 일을 이룰 수 있다. 유독 신묘하므로 급하게 서두르지 않아도 빠르게 가려고 하지 않아도 이르게 된다.²⁰⁾

조짐(혹은 기미)은 주관적인 사례, 즉 개인 자신의 특수한 상황에 맞게 개별적으로 활용되는 것이다. 그렇다고 이것이 우리가 사태 혹은 상황의 진전에 관한 개별적 혹은 구체적 과정을 장악하고자 하는 것은 아니다. 그보다는 오히려 인간이 자신이 처한 사태의 진행이나 국면 및 그 포괄적 이치를 파악하고, 더 나아가 이성적 지식과 합리적 판단 하에서 보다 주체적으로 적극적으로 사태의 진행에 참여하여 이를 자신이 바라는 방향으로 추진하고자 하는 것이다.

사회공동체의 연결주의적 차원에서 보자면, 사회가 개인들이 결합하는 근본적인 조건인 반면에, 개인이 결합해야 더욱 높은 차원의 사회를 구성할 수 있다. 사회의 통일성은 개인의 다양성을 이끌며 이 다양성은 반드시 충돌과 화해의 과정을 거쳐서 충돌이 조화로 승화되고 조화가 새로운 공동체를 형성하며, 새로운 공동체는 새로운 가치나 새로운 문화를 창출함으로써 더 높은 차원

¹⁹⁾ 『周易』, 『繫辭下傳』, “幾者, 動之微, 吉之先見者也, 君子見幾而作, 不俟終日.”

²⁰⁾ 『周易』, 『繫辭下傳』, “夫易, 聖人之所以極深而研幾也, 唯深也, 故能通天下之志, 唯幾也, 故能成天下之務, 唯神也, 故不疾而速, 不行而至.”

SESSION 5

Zhouyi's Biantong World and the Algorithm of Yin-yang from the Dimension of AI Connectionism

Yonjae Kim National Gongju University(Korea)

의 통합적 경계를 지향할 수 있다. 닫힌 사회에서는 부정적인 제재가 지나치게 발동하여 개인은 기계적으로 반응하고 고립되거나 파편화되어 전체를 위해 기계화 혹은 도구화되기 마련이다. 반면에 열린 사회에서는 긍정적인 제재가 적절하게 발동하여 개인이 고립되거나 파편화 되지 않고 적극적으로 의사를 표현하며 개인의 자율적 의지나 주체적 의식이 잘 발휘될 수 있다.

인간은 생명의 본성을 전제로 하여 자신의 존재의 타당성을 모색하고 공동체적 의식 속에서 가치의 도덕적 합당성에서 추구한다. 『주역』에서 도덕적 인간으로서의 실현가능성의 문제는 '道義의 관문'으로 표현된다. 『계사전』에서는 다음과 같이 말한다.

하늘과 땅이 위치를 세우니 번역이 그 속에 진행된다. 본성을 이루어 계속 존재하고 존재하는 것이 道義의 관문이다.²¹⁾

하늘과 땅이라는 세계의 실재 속에서 신진대사와 세대교체와 같은 생명의 변화가 진행된다. 인간은 생명의 본성에 입각한 삶의 이치를 체득하며 이를 사회의 공동체로 확충해간다. 인간과 자연의 내재적 관계에서 보자면, 인간은 생명력의 자생적 율동[生生不息]이 작동하는 자연계에서 주체적 자각의식[自強不息]을 갖고서 살아간다. 더 나아가 인간은 자연계의 생태적 질서 속에서 생명의 이치를 터득하고 사회의 공동체적 의식을 체득하고 마침내 윤리적 혹은 도덕적 가치도 모색하였다. 道義의 관문은 天道의 운행질서에 맞추어 人道の 규범적 질서를 만들어가는 삶의 방향성을 가리킨다. 天道의 운행질서는 존재의 변화가능성과 관련되고 人道の 규범적 질서는 가치의 지속가능성과 관련된다. 인간 삶의 歷程이란 존재의 변화가능성으로부터 가치의 지속가능성을 모색하는 과정이다. 여기에는 생명의 이치로부터 도덕의 가치를 확충하는 삶의 지향성이 있다. 따라서 道義의 관문은 인간의 온전한 본성을 고양하는 자아실현의 통합적 경계를 지향하는 통로라고 말할 수 있다. 여기에서 乾坤의 관계망은 음양의 연결고리라는 연결주의적 접속을 통한 사회공학적 연장선에서 인간 삶의 통합적 경계를 지향하고 성취하는 통로인 셈이다.

²¹⁾ 『周易』, 「繫辭上傳」, “天地設位, 而易行乎其中矣. 成性存存, 道義之門.”

SESSION 5

Zhouyi's Biantong World and the Algorithm of Yin-yang from the Dimension of AI Connectionism

Yonjae Kim National Gongju University(Korea)

VII. 문제해결의 실마리

『주역』의 세계관은 만물이 끊임없이 변화하는, 生生不息의 과정에 초점이 맞추어져 있다. 여기에는 인간의식의 차원에서 세계의 실재에 접근하는 變通의 원리가 있다. 그것은 자연계에서 진행되는 生生不息의 과정과 그에 따른 인간사회의 규범적 질서와 맞물려있다. 그것은 天地의 연계망, 陰陽의 연결망, 乾坤의 관계망에서 접근될 수 있다. 여기에는 질서의식의 역동적인 통합적 노선, 즉 동일선, 연속선 및 연장선이 있다. 天地의 연계망 속에서 공동체의식의 동일선으로, 陰陽의 연결망 속에서 알고리즘의 연속선으로, 乾坤의 관계망 속에서 통합적 접속의 연장선으로 진행된다. 즉 천지의 공동체적 의식동일선, 건곤의 연장선 및 이들의 관계를 나타내는 연결고리이다. 그러므로 음양의 방식은 天地의 존재론적 범주와 乾坤의 가치론적 범주 사이에 알고리즘으로 작동한다. 이들의 관계는 인공지능체의 특징 중의 하나인 연결주의의 접속과 연관 지을 수 있으며, 사회공학적으로 세계의 실재가 지닌 개별적인 경계를 넘어서 통합적 경계를 지향하는 통로가 된다.

세계의 실재에서 삼라만상의 양상과 그 방식은 음양의 연결고리를 통해 부호화되거나 기호화되고 궁극적으로 추상화되고 관념화가 된다. 음양의 범주는 현실 세계와 접속하는 통로의 기제이면서도 가상 세계에서 개별적 경계를 넘나드는 기제이다. 삼라만상의 현실 세계는 음양의 방식을 통해 접근된다면 64괘의 가상 세계는 음양의 방식을 통해 수용된다. “한번 음이 되면 한번 양이 되는” 접속의 과정에서 삼라만상의 세계가 운행되는 연결고리의 방식이 상징적으로 표현되는 한편, 64괘의 세계를 구성하는 연결고리의 방식을 상징적으로 재현된다. 그러므로 음양의 연결고리는 전체의 연결주의적 차원에서 64괘의 유기적 결집체를 구성하고 유지한다. 현실 세계와 가상 세계 사이의 경계에서는 건괘와 곤괘의 출입구에서 음양의 범주적 매체를 통해 복잡다단하고도 변화무쌍한 사회적 관계와 접속하는 것이다.

變通的 경계에는 인간이 변화의 가능성 속에서 어떻게 자신의 존재를 지속적으로 이루어내는가 하는 삶의 과정과 관련된다. 인간은 현실의 시공간에서 세계와의 의미있는 관계를 통해 혼종 혹은 합생의 통합적 과정을 겪는다.

SESSION 5

Zhouyi's Biantong World and the Algorithm of Yin-yang from the Dimension of AI Connectionism

👤 Yonjae Kim National Gongju University(Korea)

그러한 과정에서 인간은 因果관계의 필연적 속박을 넘어서 특정의 목표나 지향성을 갖게 된다. 인간은 현재에서 과거와의 연속성을 유지하면서 혼종 혹은 합생의 지속적 과정에서 미래의 삶을 꾸려갈 수 있다. 그러므로 인간은 현실적인 삶에서 생물학적인 유전인자처럼 어느 정도 고정된 것이 아니라 변화하는 세계에서 합생의 역동적 관계(nexus)라는 일련의 창발적 혹은 창신적 과정을 통해 새롭게 거듭나는 것이다.




이러한 연결주의의 사회공학적 맥락에서 보자면 개인의 자율적 주체와 사회의 통합적 구조는 서로 모순되기도 하지만 서로 양립가능하다. 여기에서 사회의 결정론적 사고와 개인의 창조적 주체 사이에 격차를 줄이는 방식이 중요하다. 어떠한 체계나 조직이든지 간에 구성요소들이 분화와 통합을 특징으로 하는 상호작용의 과정이 있기 마련이다. 전체적인 통합의 수준에서 보자면, 변화를 일정한 정도로 포용하는 균형이 이루어지는 반면에, 균형을 일정한 정도로 수용하는 변화가 발생한다. 여기에서는 어떠한 특정의 시기나 상황에서도 완성된 혹은 완결된 균형이란 결코 없으며 균형에 도달한 상태는 새로운 단계로 변화하기 위한 또 하나의 과정에 지나지 않는다.

결론적으로 말해, 변통의 연결주의적 차원에서 음양의 알고리즘은 세계의 실재에서 현실과 가상의 통합적 경계를 추구하는 사회공학적 방법론적 성격을 지닌다. 이처럼 변통의 세계를 연결주의적 관점에서 접근하는 방법은 과학의 영역을 인문학의 시선에서 조망하고 이를 토대로 하여 자연과학과 자연철학의 접점, 과학적 사유와 철학적 사유의 경계를 해체하는 통합적 경계를 지향할 수 있는 것이다.

SESSION 6

Moderator  Travis Shaw Chung-Ang University(Korea)

AI and Society (with Algorithm)

| | |
|-------------------|---|
| 16:15 ~ 16:45 KST | <p>Love Story in the Algorithmic Age  Aditi Vashistha University of Delhi(India)</p> |
| 16:45 ~ 17:15 KST | <p>Application of Artificial Intelligence through the K nearest Neighbor Algorithm to Classify Healthy Indonesian Card Recipients in an Appropriate Target  Rifaldy Fajar Yogyakarta State University(Indonesia)</p> |
| 17:15 ~ 17:45 KST | <p>Boredom, Time-Perception and Algorithmic Governmentality  Aditya Nayak Jawaharlal Nehru University(India)</p> |
| 17:45 ~ 18:15 KST | <p>Discussion</p> |

SESSION 6

Love Story in the Algorithmic Age

 **Aditi Vashistha** University of Delhi(India)

Abstract

This paper explores love stories in the algorithmic age through a comparative study of love stories from two different times, while focusing on the separation of selves. First, is “Meghadutam” by Kalidasa composed in 375CE India. Second, is an episode “Striking Vipers” from the dystopic TV series ‘Black Mirror’. While the former story is only a separation of space, the latter, is a separation in temporality because of digital communication.

Dystopic literature today transgresses¹ the limitation of space that existed in literature earlier. Separation in love, and the desire to be with the lover are two intimate parts of a love story. Desire, while forming itself through phantasm² creates its own life-world. Separation and the desire to be with the lover creates memories. In Kalidasa’s Meghdutam³, space works as a boundary that defines separation. It is limit of movement and communication. Meghadutam is a love poem where Yaksha⁴ is on a lonely mountain peak, desperate to be with his beloved. He asks the cloud to deliver a message to her in the Himalayan city of Alaka. It describes mountains, trees, oceans, and rivers in the space between them. Only the territorial distance is significant, and not the time because Yaksha is completely absorbed in her memory.

‘Striking Vipers’ is a story of two individuals who met online where they developed a virtual sexual relationship, and their virtual characters developed emotional intimacy. One of them was single while the other had a married life which started getting affected by this online relationship. Here, the separation can be understood not in the terms of space but in the terms of time. The algorithmic age has

¹ Transgression can be defined as a movement which breaches the limit to create a new limit. This new limit will be breached again in a continuous process.

² Phantasm are those residues of feelings that we have, and try to understand through the corresponding activities and images from outside.

³ Means cloud messenger.

⁴ A benevolent nature spirit.

SESSION 6

Love Story in the Algorithmic Age

 **Aditi Vashistha** University of Delhi(India)

opened up a new dimension of existence. Where the self exists at multiple levels simultaneously. This renews the question of existence. The self that exists at multiple levels is living multiple lives also.

References

- Deleuze, Gilles, and Leopold Sacher-Masoch. 1971. *Masochism; an interpretation of coldness and cruelty. Together with the entire text of Venus in furs*. New York: G. Braziller.
- Foucault, Michel, and James D. Faubion. 1997. *Aesthetics, method, and epistemology*. New York: New Press.
- Freud, Sigmund. 1922. *A general introduction to psychoanalysis*. New York: Boni and Liveright.
- Guattari, Gilles Deleuze and Felix. 1975. *Anti-Oedipus*. University of Minnesota press.
- Kalidasa. 2017. *Meghadutam*. Translated by Srinivas Reddy. Delhi: Penguin.
- McLuhan, Marshall. 1964. *Understanding media: the extensions of man*. New York: McGraw-Hill.
- Illouz, Eva. 2007. *Cold intimacies: the making of emotional capitalism*. Cambridge, UK: Polity Press.

SESSION 6

Love Story in the Algorithmic Age

 **Aditi Vashistha** University of Delhi(India)

LOVE STORY IN ALGORITHMIC AGE

This paper explores love stories in the algorithmic age through a comparative study of love story of two different times, while focusing on the separation of selves. First is “Meghadutam” by Kalidasa in 375CE. Second is an episode “Striking Vipers” from dystopic TV series “Black Mirror”.

The Meghadūta (Cloud-Messenger, written by Kālidāsa) is the first of a genre, *saṁdeśa-kavya*, that flourished in Sanskrit and other Indian languages into the modern era. The form is simple, but found root in the minds of individuals in many different contexts. (German Romantics loved the poem, but only rarely went on to pen *saṁdeśa-kavya* of their own.) In this, the model poem, two lovers find themselves separated by half of the Indian subcontinent. The poem itself consists of a message between the two lovers and directions for how to deliver it. In the Meghadūta the lovers are separated by many miles, but later poems play with the distance represented. Whether separated by mountain ranges, highways, miles, or yards, the distance hardly matters. As long as the lovers are not bound by the tight embrace of love they feel restless and unsure and must communicate with those they are separated from. Whether read in the heyday of the Gupta Empire, Schlegel’s Germany, or New York’s Upper East Side, Kālidāsa’s verses can strike the heart of many a lover.

The poem opens with a pitiful image of a nameless Yakṣa (demi-god), exiled, alone, powerless, wasting away in Central India. He has been brought to this sorry state as a punishment from the god Kubera. Not long ago our hero had been gainfully employed by the god and lived an idyllic life with his young wife in the Himalayas. But he became distracted (presumably by the thought of his lover), slacked off, and was fired—which is even worse if your boss is a god. While exiled to a monastery on a mountain, the Yakṣa constantly thinks of his lover, neglecting his body and growing so thin that “his gold bracelet slipped off his wrist.” As the months wear on and the seasons change, our hero sees the first clouds of summer, which in India bring the monsoon. Business and trade come to a halt during the monsoon, when travelling mer

SESSION 6

Love Story in the Algorithmic Age

 **Aditi Vashistha** University of Delhi(India)

chants return to their homes, a fact that makes our Yakṣa (stuck in the monastery until autumn) even more despondent. In desperation, the Yakṣa calls out to the largest cloud he sees, asking it to carry a message to his lover.

The Yakṣa begins by describing the journey the cloud is to take across India and ends with a sweet message for his beloved. Interestingly, the message itself is rather short when compared with the description of the journey the cloud is to take. Before we can get to the message we are confronted with a sense-battering display of the land that lies between the two lovers. Everything the Yakṣa touches is sexualized. The longing that unites the lovers thematizes the land. What was once homogenous and anonymous space becomes demarcated when the gaze of the Yakṣa passes over it.

‘Striking Vipers’ is an episode of the tv series, ‘Black Mirror’. There are three major characters in the episode. Danny, Karl, and Theo. Theo is Danny’s wife. Karl is an old friend of Danny. Both Karl and Danny are interested in Video games. Both starts playing a virtual reality fighting game. They can feel the sensations while playing the game. Both develops a sexual affair in virtual reality. But when they tried to kiss each other in real life, they couldn’t do it. So, it creates a dilemma for Danny. It is also difficult for Theo to understand. This story poses two questions one is about the separation of self on the level of real and virtual, the other is separation in love due to temporality.

While former is only a separation of space, the latter, because of digital communication, is a separation in temporality. Dystopic literature today transgresses¹ the limitation of space that existed in literature earlier.

Separation in love, and the desire to be with the lover are two intimate parts of a love story. Desire, while forming itself through phantasm² creates its own life-world.

¹ Transgression can be defined as a movement which breach the limit to create a new limit. This new limit will be breached again. It is a continuous process.

² Phantasm are those residues of feelings that we have, and try to understand through the corresponding activities and images from outside.

SESSION 6

Love Story in the Algorithmic Age

 **Aditi Vashistha** University of Delhi(India)

Separation and the desire to be with the lover creates memories. In Kalidasa's Meghdutam³, space works as a boundary that defines separation. It is limit of movement and communication. Only the territorial distance is significant, and not the time because Yaksha is completely absorbed in her memory. At the level of the self, it remains same. The geographical features create a definition of the love. The description of the space also points out that space is something that working as a catalyst in the story.

In 'Striking Vipers', it is different. The relationship of Theo and Danny is not threatened by any physical factor. But it is the existence of two selves. In the Algorithmic age, one can exist at multiple spaces in the same time. These spaces can be both Virtual or Real. The virtual self is different from the real self. Which can be analysed through the virtual relationship between Karl and Danny and real relationship between Theo and Danny. Theo comes to know about Danny's infidelity through change in his touches. The change in his physical gesture convinces Theo. So, the separation is different. It is not only separation in love but also separation between two selves. The virtual and real self are formulated through different medium. Because of the different medium, they are different selves.

³ Means cloud messenger.

SESSION 6

Love Story in the Algorithmic Age

 **Aditi Vashistha** University of Delhi(India)

Bibliography

- Deleuze, Gilles, and Leopold Sacher-Masoch. 1971. *Masochism; an interpretation of coldness and cruelty. Together with the entire text of Venus in furs*. New York: G. Braziller.
- Foucault, Michel, and James D. Faubion. 1997. *Aesthetics, method, and epistemology*. New York: New Press.
- Freud, Sigmund. 1922. *A general introduction to psychoanalysis*. New York: Boni and Liveright.
- Guattari, Gilles Deleuze and Felix. 1975. *Anti-Oedipus*. University of Minnesota press.
- Kalidasa. 2017. *Meghadutam*. Translated by Srinivas Reddy. Delhi: Penguin.
- McLuhan, Marshall. 1964. *Understanding media: the extensions of man*. New York: McGraw-Hill.
- Illouz, Eva. 2007. *Cold intimacies: the making of emotional capitalism*. Cambridge, UK: Polity Press.

SESSION 6

Application of Artificial Intelligence through the K nearest Neighbor Algorithm to Classify Healthy Indonesian Card Recipients in an Appropriate Target

 **Rifaldy Fajar** Yogyakarta State University(Indonesia)

Abstract

Health is the right of every Indonesian citizen. The law regulates the guarantee of getting health services for the entire community. The Healthy Indonesia Card is a card issued by the government under the auspices of the Ministry of Health to overcome the problem of the increasing number of underprivileged people who do not get health services. Artificial intelligence is considered to be able to solve problems in determining which people are entitled to receive or who are not entitled to receive the Healthy Indonesia Card. Classification methods are able to distinguish data classes or data concepts. Data collection is the essence of research where the data used will determine the probability of a Healthy Indonesia Card Recipient. Researchers took data from the Sukoharjo Regency Government Social Service, Central Java Province, and used the K-nearest Neighbor Algorithm method to classify the correct recipient of the Healthy Indonesia Card. The results show that the K-nearest Neighbor Algorithm can be used to classify Healthy Indonesia Card recipients according to the determining factors, including gender, age, age interval, latest education, occupation, income, and also dependents of children. Based on data testing, 12 trials resulted in a value of 97.66% precision, 98.5% accuracy, and a recall of 96.5%. So, the conclusion in this study is that this classification application can help policymakers in determining the classification of people who are entitled to receive a Healthy Indonesian Card or not, and based on BlackBox testing the features in the application can function properly.

SESSION 6

Application of Artificial Intelligence through the K nearest Neighbor Algorithm to Classify Healthy Indonesian Card Recipients in an Appropriate Target

👤 Rifaldy Fajar Yogyakarta State University(Indonesia)

Application of Artificial Intelligence through the K-nearest Neighbor Algorithm to Classify Healthy Indonesian Card Recipients in an Appropriate Target

**Rifaldy Fajar
Yogyakarta State University
INDONESIA**



Background

Health is the right of every Indonesian citizen.

The law regulates the guarantee of getting health services for the entire community.

The Healthy Indonesia Card is a card issued by the government under the auspices of the Ministry of Health to overcome the problem of the increasing number of underprivileged people who do not get health services



SESSION 6

Application of Artificial Intelligence through the K nearest Neighbor Algorithm to Classify Healthy Indonesian Card Recipients in an Appropriate Target

👤 Rifaldy Fajar Yogyakarta State University(Indonesia)

Background

Many problems related to the Indonesian Healthy Card Recipient!

Real example:

The recipients are not suitable person/not qualified.

Aim/Objective/Idea

- Artificial Intelligence (AI) is considered to be able to solve problems in determining which people are entitled to receive or who are not entitled to receive the Healthy Indonesia Card.
- Classification methods are able to distinguish data classes or data concepts.
- Data collection is the essence of research where the data used will determine the probability of a Healthy Indonesia Card Recipient.

SESSION 6

Application of Artificial Intelligence through the K nearest Neighbor Algorithm to Classify Healthy Indonesian Card Recipients in an Appropriate Target

 **Rifaldy Fajar** Yogyakarta State University(Indonesia)

Main Methods

- Problem Identification
- Literature Study
- Data Collection
- Data Processing
- Build the Application
- Testing
- Results Analysis
- Draw the conclusion

Methods

Problem Identification

Problem identification is carried out to find out problems that arise such as not yet optimal recipients of the Healthy Indonesia Card.

Responding to the existing problems, an information system is needed to create a classification system for determining the recipients of the Healthy Indonesia Card at the Sukoharjo District Social Service.

The data used as considerations in determining the recipients of the Healthy Indonesia Card are age, gender, occupation, latest education, monthly income, and dependents of children.

SESSION 6

Application of Artificial Intelligence through the K nearest Neighbor Algorithm to Classify Healthy Indonesian Card Recipients in an Appropriate Target

 **Rifaldy Fajar** Yogyakarta State University(Indonesia)

Methods

Literature Study

Literature study was conducted in search of information related to concepts, theories and methods relevant to problem identification.

Based on the existing problems, the authors conduct a literature study using books that discuss data mining, K-Nearest Neighbor, journals and with research that has been done previously.

By using 10-fold cross-validation and T-Test with 0.05 significance level to compare K-Nearest Neighbor, Naive Bayes, j48, and Support Vector Machine, it can be concluded that the knn classifier produces high accuracy and low accuracy produced through SVM.

All classifiers are executed with little or no tuning. (Crismanto R, 2015).

Methods

Design

To explain the function of the application, an overall application design will be described in the form of a use case diagram. All functions that can be used in the use case diagram include the following: users can add testing data and training data, delete data, users can view or delete data information.

Data Collection

Data collection is the essence of research where the data used will determine the probability of a Healthy Indonesia Card Recipient. Researchers took data from the Social Service of Sukoharjo Regency. The data that has been collected will be determined using the variables in the hypothesis. The data is used as a consideration in determining the recipients of a healthy Indonesia card which is a polynomial type, namely age, last education, occupation, salary or monthly income, dependents of children and data type label is a statement receiving KIS or not.

SESSION 6

Application of Artificial Intelligence through the K nearest Neighbor Algorithm to Classify Healthy Indonesian Card Recipients in an Appropriate Target

Rifaldy Fajar Yogyakarta State University(Indonesia)

Results

Data Testing Page (Indonesian Version App)

| NO | NIK | Jenis Kelamin | Usia | Interval | Pendidikan Terakhir | Pekerjaan | Pendapatan | Tanggungan Anak | Terima kis | Prediksi | Opsi |
|----|-----------------|---------------|------|----------|----------------------|----------------------|---------------|-----------------|------------|----------|-----------------|
| 1 | 331024906760004 | PEREMPUAN | 41 | A3 | BELUM TAMAT SD | Buruh Harian Lepas | SANGAT RENDAH | 1 | YA | | [Delete] [Edit] |
| 2 | 331022005700002 | LAKI-LAKI | 47 | A3 | DIPLOMA IV/ STRATA I | Pegawai Nagari Sipil | TINGGI | 4 | TIDAK | | [Delete] [Edit] |

Pengujian (Testing) : To know the accuration, precision, and recall.

Results

Blackbox Testing

Pengujian blackbox dilakukan untuk menguji tingkat operasional dari sistem yang telah dibuat sebelumnya. Secara keseluruhan hasil yang didapatkan dari aplikasi telah berjalan dengan baik.

- Menu
- Login (Valid)
- Import Data (Valid)
- Add Data (Valid)
- Delete (Valid)
- Delete All (Valid)
- Edit (Valid)
- Data Prediciton (Valid)
- Logout (Valid)

SESSION 6

Application of Artificial Intelligence through the K nearest Neighbor Algorithm to Classify Healthy Indonesian Card Recipients in an Appropriate Target

 **Rifaldy Fajar** Yogyakarta State University(Indonesia)

Results

K-nearest Neighbor Algorithm Testing

In testing the K-Nearest Neighbor Algorithm, a test is carried out on the results that have been carried out by the system.

Testing is done by doing calculations manually to test whether the results of system calculations produce the same results as calculations manually.

In this test, the researcher used 5 training data samples and 1 data as testing data.

From the test results, it was found that both the results from the system calculations and manual calculations produced the same value.

Thus, the system that has been built previously can be accounted for for its calculation results.

Results

Accuracy, Precision dan Recall

Pada penelitian ini, peneliti melakukan percobaan pengujian data testing sebanyak 12 kali percobaan menggunakan data testing sebanyak 200 data yang diambil secara acak. Dari pengujian didapatkan nilai Accuracy, Recall dan Precision. Adapun hasil darpi pengujian dijelaskan pada table berikut:

| No | Activity | Training Data | Testing Data | Accuracy Value | Precision Value | Recall Value |
|----|------------|---------------|--------------|----------------|-----------------|--------------|
| 1 | Testing 1 | 50 | 200 | 96% | 97% | 95% |
| 2 | Testing 2 | 100 | 200 | 97% | 98% | 95% |
| 3 | Testing 3 | 150 | 200 | 97% | 99% | 97% |
| 4 | Testing 4 | 200 | 200 | 96% | 98% | 97% |
| 5 | Testing 5 | 250 | 200 | 97% | 99% | 97% |
| 6 | Testing 6 | 300 | 200 | 97% | 99% | 97% |
| 7 | Testing 7 | 350 | 200 | 99% | 99% | 96% |
| 8 | Testing 8 | 400 | 200 | 98% | 99% | 97% |
| 9 | Testing 9 | 450 | 200 | 99% | 98% | 97% |
| 10 | Testing 10 | 500 | 200 | 98% | 99% | 97% |
| 11 | Testing 11 | 550 | 200 | 99% | 98% | 96% |
| 12 | Testing 12 | 600 | 200 | 99% | 99% | 97% |

SESSION 6

Application of Artificial Intelligence through the K nearest Neighbor Algorithm to Classify Healthy Indonesian Card Recipients in an Appropriate Target

 **Rifaldy Fajar** Yogyakarta State University(Indonesia)

Results Analysis/Discussions

The result analysis is intended to analyze the results related to the function of the previously designed system.

An admin can classify people who are entitled or not entitled to receive the Healthy Indonesia Card by entering personal data of each individual, such as Gender, Age, Last Education, Occupation, Monthly Income and Dependents of Children.

Before the admin performs the classification, an admin must enter the registered Username and Password to enter the system.

After the admin enters the system, there are features in the system such as importing data, performing calculations, deleting and changing data and viewing calculation results.

Results Analysis/Discussions

In blackbox testing the system functions can run well as well as the results of calculations carried out using the system or manually producing the same results so that the system can be justified.

In addition, testing the level of accuracy, precision and recall which was carried out 12 times resulted in an accuracy value of 97.66%, precision 98.5% and recall 96.5%.

The level of accuracy that has been carried out 12 times produces a value that tends to increase because a lot of training data produces accurate results.

In addition, the level of precision also seems to increase because there is a lot of training data so that the higher the level of accuracy between the training data and the testing data being tested.

Likewise, the recall rate obtained unstable results caused by the success rate of the system in retrieving an information.

SESSION 6

Application of Artificial Intelligence through the K nearest Neighbor Algorithm to Classify Healthy Indonesian Card Recipients in an Appropriate Target

 **Rifaldy Fajar** Yogyakarta State University(Indonesia)

Conclusions

1. This classification application can help admins in determining the classification of people who are entitled to receive the Healthy Indonesia Card or not.
2. Based on blackbox testing, the features in the application can function properly.
3. Based on the calculation algorithm testing that is done manually or by using the system produces the same value.
4. Based on testing data testing as many as 12 trials resulted in an average accuracy value of 97.66%, precision 98.5% and recall 96.5%.

Contact Information

My email:

rifaldyfajar251@gmail.com
rifaldy.fajar@student.uny.ac.id

SESSION 6

Application of Artificial Intelligence through the K nearest Neighbor Algorithm to Classify Healthy Indonesian Card Recipients in an Appropriate Target

👤 Rifaldy Fajar Yogyakarta State University(Indonesia)



Thank you!



SESSION 6

Boredom, Time-Perception and Algorithmic Governmentality

 **Aditya Nayak** Jawaharlal Nehru University(India)

Abstract

In the contemporary lifestyle, systems of entertainment have become central to our existence as human. We accumulate entertainment gadgets and organize our lives around spaces of entertainment, both online and offline. We are scared of boredom to the extent that we can do anything in order to avoid it. But is this fear of boredom natural to us as humans?

20th Century philosopher Walter Benjamin saw boredom as a condition of modern existence. He finds literary records of the feeling called ‘boredom’ for the first time in 1840s. However, Benjamin argues that boring activities feel like torture not because of being boring, but because the modern subject has been trained in a way that boredom is no longer possible.

In the process of escaping boredom, we as humans expose ourselves to infinite external stimulus that are projected at us by consumer capitalism. With the capabilities of algorithms, our behaviours are profiled for personalized recommendations and advertising. The entertainment content curated for us by algorithms in form of Web-Series recommendations, YouTube videos, and social media posts of Instagram and Facebook, take control of our consciousness.

This control mechanism reaches beyond entertainment and product advertising into controlling our political opinions and positions. The project of Algorithmic Governmentality, conceptualized by Antoinette Rouvroy, is to eliminate multiple alternative visions of the future, and make us believe in a singular future that has been algorithmically calculated.

Boredom isolates us, and thus becomes a defence mechanism against the external stimulus. It does not allow the stimuli to activate a state of ‘shock’ in our subconscious mind in order to create ‘involuntary memory’. These involuntary

SESSION 6

Boredom, Time-Perception and Algorithmic Governmentality

 **Aditya Nayak** Jawaharlal Nehru University(India)

memories become unconscious instructions that produce needs in us for consumer commodities, and later, political opinions. Ian Buchanan sees this as the reason why consumer capitalism cannot tolerate boredom. Modern capitalism has trained us to avoid boring, repetitive actions in order to hegemonise our consciousness, and make us vulnerable to follow its instructions. After appropriation of resources and human labour, Frederic Jameson argues that the final frontier of capitalism is our consciousness itself.

We need to train ourselves to stay in states of repetition by developing particular kinds of boredom as a response to distractions of commodity fetishism. It will give us more creative ways of expressing ourselves. We must learn to wait in this infinite time, to access infinite possibilities of the future.

SESSION 6

Boredom, Time-Perception and Algorithmic Governmentality

 **Aditya Nayak** Jawaharlal Nehru University(India)

References

- Moran, J. (2003). Benjamin and Boredom. *Critical Quarterly*, 45(1-2).
- Spacks, P. M. (1995). *Boredom: The Literary History of a State of Mind*. Chicago: University of Chicago Press.
- Benjamin, W., & Tiedemann, R. (1999). *The Arcades Project*. Cambridge: Belknap Press.
- Buchanan, I. M. (2017). The Disappearance of Boredom. *Lo Sguardo - Rivista di Filosofia*, 23, 277- 284.
- Jameson, F. (2003, May/June). Future City. *New Left Review*, 21.
- Phipps, J. F. (2004). Henri Bergson and the Perception of Time. *Philosophy Now*(48).
- Bergson, H. (1988). *Matter and memory*. New York: Zone Books.
- Stiegler, B. (2016). *Automatic Society: The Future of Work (Vol. 1)*. Cambridge: Polity Press.
- Rouvroy, A., & Berns, T. (2013). Algorithmic governmentality and prospects of emancipation. Disparateness as a precondition for individuation through relationships? *Réseaux*, 177(1).
- Heidegger, M. (1977). *The question concerning technology, and other essays* . New York: Harper and Row.
- Bergson, H. (1911). *Creative Evolution*. London: Macmillan.
- Deleuze, G., & Guattari, F. (1975). *Anti-Oedipus*. University of Minnesota press.
- Proust, M., & Sturrock, J. (2003). *In Search of Lost Time*. London: Penguin.
- Kracauer, S. (1995). *The Mass Ornament: Weimar Essays*. Cambridge: Harvard University Press.
- Nietzsche, F. W. (1974). *The Gay Science*. New York: Vintage Books.

SESSION 6

Boredom, Time-Perception and Algorithmic Governmentality

 **Aditya Nayak** Jawaharlal Nehru University(India)

Abstract

The contemporary life is organised around systems of entertainment. We are scared of boredom. This fear of boredom is not a natural fear in human beings, it began to be experienced with the dawn of modernity. Repetitive, monotonous boredom protects us from external stimuli, this is why consumer capitalism requires its absolute demolition to acquire complete control over the consciousness of its subjects. This control extends beyond product advertisement to social as well as political control.

Keywords: 1) Boredom; 2) Walter Benjamin; 3) Algorithmic Governmentality; 4) Bergson; 5) Fake News

Table of Contents

1. Introduction
2. Boredom- A Modern Experience
3. Time-Perception of Boredom, and Algorithms
4. Algorithmic Governmentality
5. Conclusion

SESSION 6

Boredom, Time-Perception and Algorithmic Governmentality

 **Aditya Nayak** Jawaharlal Nehru University(India)

1. Introduction

In our everyday life, we are terrified by the idea of having nothing to do for a day. Weekends and holidays feel intolerable without a fun-activity. We are all scared of boredom. But is this fear accompanying humanity since forever? Or have we acquired it over time? What are the repercussions of boredom, and the fear of boredom on our social, and political lives? These are few questions that are going to guide this enquiry.

The paper can be divided into three sections – First, understanding boredom in contemporary times; Second, the time-perception of boredom, and how algorithms interact with it; Third, the ways in which the boredom-algorithmic interaction has facilitated creation of a subject who desires their own repression in times of algorithmic governmentality.

The paper attempts to build upon the existing theories, and philosophies on boredom and time, in order to apply them to the contemporary life surrounded by algorithms. We shall discuss the works of Walter Benjamin, Sigmund Freud, Henri Bergson, Marcel Proust and some more thinkers of relevance.

2. Boredom- A Modern Experience

The contemporary life is organised around gadgets and systems of entertainment. While it was possible to see such a life as pathological or anomalous before the Covid-19 pandemic, it has become the normal state of affairs for every individual now. No surprise, the global entertainment industry has seen a continuous growth from \$1.39 Trillion (2011)¹ to \$2.1 Trillion (2021) . Deeper internet penetration, increase in smartphone users, and the proliferation of Over-the-top (OTT) platforms is just another symptom of the entertainment demand.

¹ www.statista.com/statistics/237749/value-of-the-global-entertainment-and-media-market/

SESSION 6

Boredom, Time-Perception and Algorithmic Governmentality

 **Aditya Nayak** Jawaharlal Nehru University(India)

The underlying phenomenon providing a major drive to the entertainment industry is the feeling of boredom, and the willingness to avoid it by engaging ourselves in some form of activity. We feel an urge to escape monotony, do something that will satisfy our need for mental stimulus, and then allow us to get back to monotony. However, every time we attempt to distract ourselves, we fail to achieve that feeling of satisfaction, and hence need to move to some other form of entertainment/stimulus.

But how do we understand this feeling called ‘boredom’? Why do we want to avoid this feeling? Has there been a transition in historical time in how we experience and deal with boredom? These are some of the questions that we are taking up to begin our discussion.

20th Century philosopher Walter Benjamin was deeply interested in the feeling of boredom. He sees boredom as a historical experience which, according to him, expanded its reach in the 1840s. Thus, he understands boredom as not only something related to modernity but the way in which we experience modernity. In other words, boredom is not a pathology, but the very condition of modern existence.

“We are bored when we don’t know what we are waiting for. That we do know, or think we know, is nearly always the expression of our superficiality or inattention” (Benjamin & Tiedemann, 1999). - Walter Benjamin, The Arcades Project

For Benjamin, the significant development of modernity is that of the replacement of *Erfahrung* (the capacity to assimilate, recollect and communicate experience to others) with *Erlebnis* (the sense of life as a series of disconnected impressions with no common associations). For him, a man becomes a hostage to boredom when he is denied the potential for *Erfahrung* because he feels expelled from the everyday time (Moran, 2003).

SESSION 6

Boredom, Time-Perception and Algorithmic Governmentality

 **Aditya Nayak** Jawaharlal Nehru University(India)

Joe Moran (2003), in his essay ‘Benjamin and Boredom’ identifies that the sentiment that would more appropriately represent Benjamin’s idea is ‘ennui’² which he sometimes uses interchangeably with ‘boredom’. Ennui is something more permanent existential perception. Benjamin identifies it as a ‘culturally significant sentiment’ (Moran, 2003).

The argument is further developed by Ian M. Buchanan (2017), in his essay ‘The Disappearance of Boredom’. With the example of the Airport, Buchanan stresses on our inability to wait. He points out that waiting at the airport, or a railway station feels like torture. However, it is not torture because of being boring, but because we have been culturally trained in a way that ‘boredom is no longer possible’ (Buchanan, 2017). This partly explains why our contemporary lifestyle has come to be organised around avoidance of boredom.

Buchanan argues that consumer capitalism cannot tolerate boredom, which is why it has transformed the airports into ‘malls’ under the logic that waiting does not feel boring if you are shopping. This argument is extended to the smartphone that has transformed the entire world into a shopping mall. They have not just transformed the space, but also transformed time; hence leading to our inability to wait. Buchanan (2017) argues that consumer capitalism has killed its most potent critic by conquering boredom. According to him, boredom is our defence against the present. (Buchanan, 2017)

Here, the ‘present’ means the various kinds of external stimuli that are projected at us through advertisements, social media, or social expectations of behaviour. Repetitive behaviour, and monotony isolate us from these external stimuli and thus defend us from the present. But when we indulge ourselves in distractions in order to avoid the repetitive boredom, we make our unconscious mind available for these stimuli. These distractions have presently taken the form of watching videos, social media, and other types of online content.

SESSION 6

Boredom, Time-Perception and Algorithmic Governmentality

 **Aditya Nayak** Jawaharlal Nehru University(India)

While these external stimuli can be easily identified as ‘commercial advertisements’, they do not lose their power to influence our needs (desires) merely by being identified. They still remain powerful because of an underlying phenomenon that Benjamin calls ‘shock’. This shock is a central need for the advertisers in order to penetrate into deeper levels of consciousness and spark a desire to buy.

A further discussion of this phenomenon of ‘shock’ shall take us into the next phase of our discussion. In this section, we shall try to understand the time-perception of boredom and how it has been conquered by consumer capitalism in general, and contemporary algorithms in particular.

3. Time-Perception of Boredom, and Algorithms

Walter Benjamin was interested in tracing the structure of the ‘media form’ in transition. Or, the way in which information gets delivered to us. For his time, he analysed how the newspaper had acquired a headline-form in a way that the content was delivered as ‘pure information’. These disconnected news items could not be assimilated and comprehended as ‘experience’ by the readers (Buchanan, 2017). This was a shift from the narrative or the story form which was primarily focussed on delivering an experience to the audience. Benjamin says that every new media form competes with its predecessor in moving away from the story form.

This is also true for the form of the contemporary social media where the individual social media posts do not contain complete narratives, and stories. The entirety of the narrative may be comprehended over a long-term acquaintance with the context of the social media post. If an individual is given someone else’s social media account, it is nearly impossible that they will be able to decipher the meanings of the individual posts. Over the past few years, Facebook, and now Instagram have both moved to a non-chronological order for social media posts. This allows the algorithms to not only analyse individual preferences but also shape them in a personalised way based on their browsing activity and interests.

SESSION 6

Boredom, Time-Perception and Algorithmic Governmentality

 **Aditya Nayak** Jawaharlal Nehru University(India)

The lack of comprehensibility of disconnected information pieces because of non-existent narrative leads to something Benjamin called ‘shock’. In order to formulate this concept, Benjamin draws heavily from Freud, Bergson, and Proust.

He discusses ‘shock’ in two frames. Firstly, he frames it historically, i.e. how the successive media forms undergo transition in moving away from the original ‘story’ form of narrative. Secondly, he addresses it in relation to the unconscious (Freud). Benjamin thinks about why every new media form will want a more absurd form of delivery? He says, this is because the advertisers need to produce that lack of assimilated comprehensibility in order to generate a moment of ‘shock’ so that their product becomes memorable (Buchanan, 2017). From newspaper, to radio, television, online videos, and now social media, each media form only disintegrated the sense of assimilation only further, thus causing deeper shocks in every successive form.

This moment of shock ejects the audience from the usual receptive state of monotony, and penetrates their conscious mind. This allows the advertisement’s stimulus to generate a ‘desire to buy’ at a level beyond the conscious mind. The ultimate purpose is to program the unconscious in a way that ‘shopping’ becomes ‘a fantasy activity in its own right’ (Jameson, 2003). You do not have to need something to buy it; you buy something because you need a fun-activity to distract yourself from boredom.

In Freud’s hypothesis, the things that become conscious cannot become a trace of memory. For him, consciousness plays another important function, i.e. protection against stimuli. He says that protection from stimuli is just as important as reception of stimuli. For this, he argues that excess stimuli can be managed by repeating it and working until it can become an ‘experience’ and not just a sense perception. This is an important intervention for Benjamin in developing his idea of ‘shock defence’.

Shock defence for Benjamin is a form of repetitive training that enables us at the level of the unconscious to internalize the indigestible excess stimulus and ‘make

SESSION 6

Boredom, Time-Perception and Algorithmic Governmentality

 **Aditya Nayak** Jawaharlal Nehru University(India)

sense of it without having to think about it' (Buchanan, 2017). Boredom is similar to this repetitive training that isolates us from external stimuli, and thus defends us against the present.

When we are bored waiting for some event, time feels to pass differently for us. Ian M. Buchanan says that "time has reasserted itself in a paradoxical way: on the one hand, it has lengthened – the moment seems never to pass, it becomes bloated, expanding without end – but, on the other hand we do nothing to shorten it – indeed, we refuse to pass the time and thus make time pass" (Buchanan, 2017). In such a state, we become impenetrable by the multiple stimuli of consumer capitalism. No entertainment feels entertaining enough at this point of time. However, we have been trained over time to stay away from this non-receptive ascetic-like state to the extent that boredom has become impossible, and entertainment has become compulsive.

It is interesting to think of the time-perception of boredom where time itself seems to pass differently. This happens in other mental-emotional states as well. There is an entire specialised field of study on time-perception. However, for the purpose of this paper, we shall limit ourselves to the time-perception of boredom in relation to the algorithm-moderated online media.

At this point, it is possible for us to explain the shift from the previous chronological timeline on Facebook and Instagram to the new non-chronological layout of individual posts. The chronological layout was allowing for an organising principle that left a sense of assimilation of posts over a short period of time. So, under the chronological layout, when you opened your Facebook or Instagram wall twice in a matter of few minutes, or even hours, it was possible for you to scroll down to an order of the timeline that you had already visited earlier and hence you would stop scrolling beyond that point. But with the new non-chronological layout of the posts, you can no longer visit your own timeline in the same order twice. Even the minimal possibilities of repetition have been eliminated in order to facilitate micro-shocks at every next visit. This adds to the compulsive checking for new posts and notifications

SESSION 6

Boredom, Time-Perception and Algorithmic Governmentality

 **Aditya Nayak** Jawaharlal Nehru University(India)

while you can keep scrolling endlessly on your every visit to the website/Application. In other words, the smartphone becomes more capable of capturing our attention, and keeping us away from the repetitions of boredom.

The point of integration of the algorithms with our time-perception of boredom is apparent now. The primary purpose of these lines of code is to keep us as far as possible from falling back into monotony of boredom. We need to be kept constantly activated by shock-therapy in order to keep us receptive to the continuous instructions, and suggestions of the algorithms that harvest data about us. The cookies attached to our personal browsing history have made it possible to calculate the type and style of content that can keep us online for the longest possible time. It will also make us come back, compulsively.

At this point, it needs to be reasserted that we do not avoid boredom because it feels like torture. In fact, journalist Terence Blacker noted in 2004 that if you type words like ‘torture’, ‘rape’ or slave into a search engine, you would not be led to some human rights organisation, or academic reports but to thousands of websites of recreational sadism (Phipps, 2004). These websites exist as top search results because people have preferred them over other results. In other words, the internet audience has turned even torture into a form of entertainment in escaping boredom. The algorithms have hypnotised us into a completely receptive state ready to receive instructions.

Fredric Jameson had argued that the final frontier of capitalism was always the consciousness itself (Buchanan, 2017). He intends to say that capitalism wants a complete control over the subjects, to the extent that they only listen to the instructive stimulus produced by capitalism according to its own requirements. In the current form of the consumer society, where boredom has become impossible, we have lost our strongest defence without ever realising it.

The concept is further expanded in the work of Gilles Deleuze (1975) and Felix Guattari (1975) ‘*Anti-Oedipus: Capitalism and Schizophrenia*’. They introduce the con

SESSION 6

Boredom, Time-Perception and Algorithmic Governmentality

 **Aditya Nayak** Jawaharlal Nehru University(India)

cept of a desiring-machine which produces needs in its subjects. They substantiate by saying that every society has several desiring-machines; however, the most dominant one in contemporary society is capitalism (Deleuze & Guattari, 1975). The concept is important because it visualises a project of integration of the biological human body with the desiring-machine in a way that the body functions according to the needs and requirements of the machine without even realising it.

We need to ask ourselves if the instructive control of algorithms on us extend only to commercial advertising and compulsive buying of products? Or does it extend beyond it into social and political being? The next section shall attempt to answer these questions.

4. Algorithmic Governmentality

The discussion so far about boredom, and the shock-defence interplay has prepared grounds for our discussion on algorithmic governmentality. The processes and sentiments discussed so far shall help us understand how the state power regulates us, and our opinions through management of our boredom. Walter Benjamin identifies digital media not as a source of information or distraction, but as a formative force in shaping our subjectivity.

This boredom-shock interplay has now been assimilated into contemporary political management. Several statements and claims of Donald Trump, the former President of the United States, have been found to be “absurd”, “surprising”, “shocking”. It has been similar for certain other parts of the world. But while we debate the sense/non-sense and logic of such statement by these important people, we need to realise that the purpose of these statements is not to appeal to the conscious minds of the masses. Rather, it is an attempt to breach and penetrate through the conscious mind and establish themselves as an ‘involuntary-memory’ (Proust & Sturrock, 2003) far deeper than the reach of our conscious minds. Much like the absurdity that is expected from a good advertisement. Political management has

SESSION 6

Boredom, Time-Perception and Algorithmic Governmentality

 **Aditya Nayak** Jawaharlal Nehru University(India)

entered the territory where the best advertiser shall remain in the media, and public memory for the longest time.

This reading of contemporary politics helps us in understanding the dis-information campaigns, and ‘fake-news’ phenomenon that has gripped us like a new normal way of conducting politics.

But the role of boredom-shock interplays goes far deeper into the political vision of the future. Antoinette Rouvroy (2013) proposed the concept of ‘algorithmic governmentality’. It is an imagination of the world that is based on algorithmic calculations of big data sets, rather than politics, law, and social norms. The political issues are sought to be quantified in such a way that the multiple/infinite possible uncertainties are reduced into very few probable outcomes. The primary concern is to do away with ‘uncertainties’ and ‘risks’ by moving from obligations to necessities (Rouvroy & Berns, 2013).

Algorithmic governmentality imagines a hyper-individualist society that ends all identifications with classes and groups or communities. Each one must be governed by oneself. At the outset, it appears to be a liberating idea. However, the concern of elimination of risks or uncertainties comes with an implicit condition that these risks shall be defined according to the algorithmic calculations. Only the future that has been algorithmically calculated shall be actualised, every other deviating form/imagination of the future must be suppressed and eliminated. In other words, the reality itself must be subordinated to the vision of the future calculated by the algorithms. A compulsory submission to the decided path.

In the future of algorithmic governmentality, the questions about the processes and variables involved in calculation of the future do not carry any credibility. What matters is that every individual conforms to one vision of the future. It is at this point that our fear of boredom becomes relevant to the governmentality apparatus. We can be made to desire our own repression, by selling us the ‘best’ vision of the future.

SESSION 6

Boredom, Time-Perception and Algorithmic Governmentality

 **Aditya Nayak** Jawaharlal Nehru University(India)

It shall be considered the best because it has been ‘calculated’.

At this point, we need to reflect upon the ways in which we can deal with our fear of boredom. We need to prepare our defence mechanism against the excess stimulus of capitalism that has taken over our collective sense of time.

5. Conclusion

Walter Benjamin uses Marcel Proust’s work ‘*In Search of Lost Time*’ to argue that boredom helps us in accessing certain memories of the everyday life that remain inaccessible in a conscious state of the mind. The shock of modern experience has fractured the possibilities of attaining assimilative memory processes, hence the only way of retaining experience is through ‘*memoire involontaire*’ (Proust & Sturrock, 2003).

Sigfried Kracauer (1924) in his essay *Franfurter Zeitung* also recommends developing a particular kind of boredom as a response to the distractions of the commodity in the modern city. He says that it will allow us to cultivate more creative ways of expressing ourselves.

Perhaps the most helpful reflection comes from Friedrich Nietzsche’s concept of the ‘eternal return’ which talks about infinite time, and finite events that recur infinitely (Nietzsche, 1974). It is a time-perception much like boredom where time feels never-ending, and we have almost nothing to do in that infinite time. Eternal return is an attempt to reconcile the mutually contradictory tendencies of desire: that of repetition and that of eternity.

Learning to wait, and the act of allowing ourselves to get bored without looking for distractions contains in itself infinite possibilities of creativity. We must not deprive ourselves of this eternal joy.

SESSION 6

Boredom, Time-Perception and Algorithmic Governmentality

 **Aditya Nayak** Jawaharlal Nehru University(India)

References

- Moran, J. (2003). Benjamin and Boredom. *Critical Quarterly*, 45(1-2).
- Spacks, P. M. (1995). *Boredom: The Literary History of a State of Mind*. Chicago: University of Chicago Press.
- Benjamin, W., & Tiedemann, R. (1999). *The Arcades Project*. Cambridge: Belknap Press.
- Buchanan, I. M. (2017). The Disappearance of Boredom. *Lo Sguardo - Rivista di Filosofia*, 23, 277-284.
- Jameson, F. (2003, May/June). Future City. *New Left Review*, 21.
- Phipps, J. F. (2004). Henri Bergson and the Perception of Time. *Philosophy Now*(48).
- Bergson, H. (1988). *Matter and memory*. New York: Zone Books.
- Stiegler, B. (2016). *Automatic Society: The Future of Work (Vol. 1)*. Cambridge: Polity Press.
- Rouvroy, A., & Berns, T. (2013). Algorithmic governmentality and prospects of emancipation. Disparateness as a precondition for individuation through relationships? *Réseaux*, 177(1).
- Heidegger, M. (1977). *The question concerning technology, and other essays* . New York: Harper and Row.
- Bergson, H. (1911). *Creative Evolution*. London: Macmillan.
- Deleuze, G., & Guattari, F. (1975). *Anti-Oedipus*. University of Minnesota press.
- Proust, M., & Sturrock, J. (2003). *In Search of Lost Time*. London: Penguin.
- Kracauer, S. (1995). *The Mass Ornament: Weimar Essays*. Cambridge: Harvard University Press.
- Nietzsche, F. W. (1974). *The Gay Science*. New York: Vintage Books.