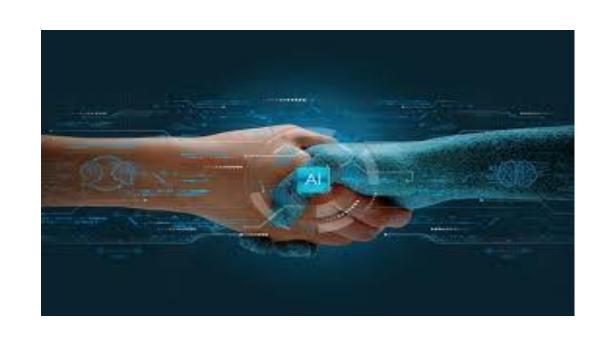
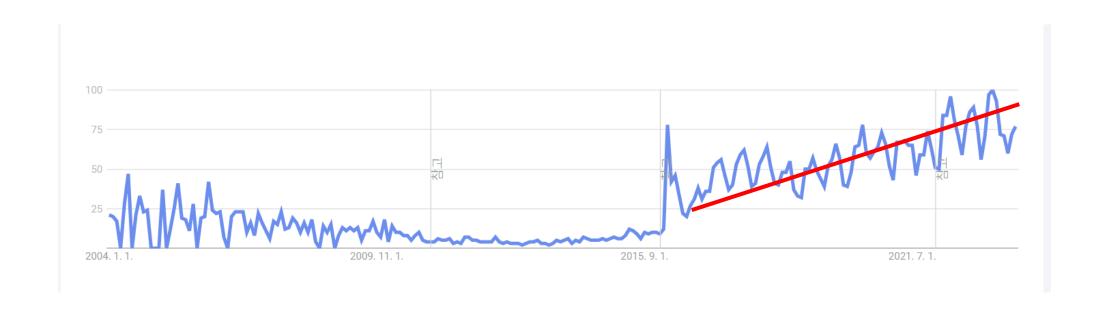
# Human-Al Collaboration in Public Workplaces: Prospects and Challenges



M. Jae Moon
Department of Public
Policy and
Management
Yonsei University

Global HR Forum 2024 October 30 2024

# Increasing Interests in Al



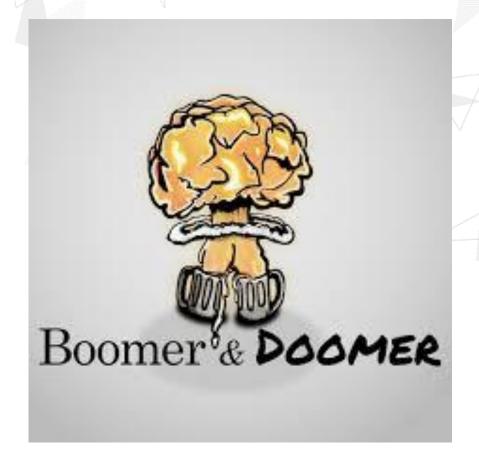
(Moon by Ngram Viewer, 2023)

Human-Al Collaboration: Prospects and Challenges







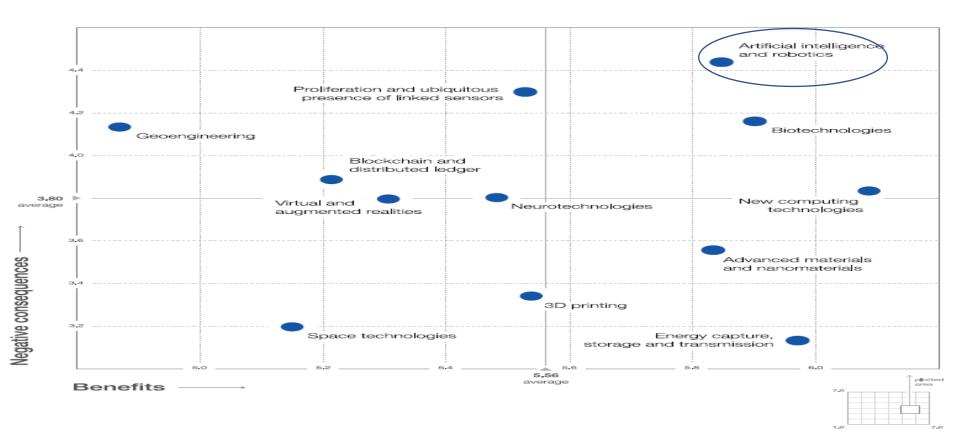






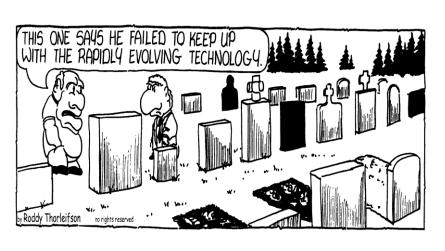


# Perceived Benefits and Negative Consequences of Different Technologies



Source: World Economic Forum Global Risks Perception Survey 2016

# Technology, Human, and HR Cycle





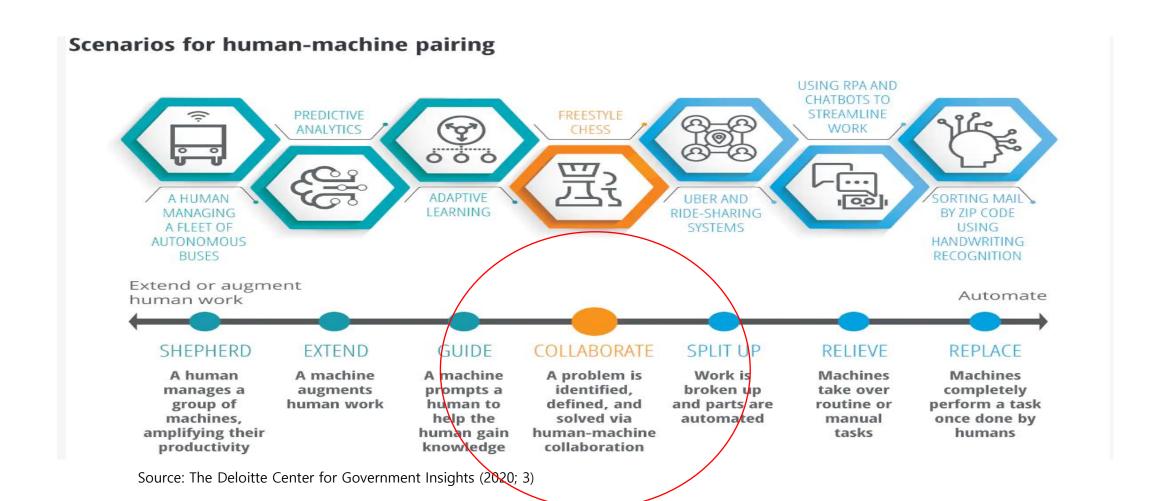


### AI for HRM and HRD

- Al for Recruitment
- Al for Selection
- Al for Job Analysis
- Al for Job Assignment
- Al for Training and Development
- Al for Performance Evaluation

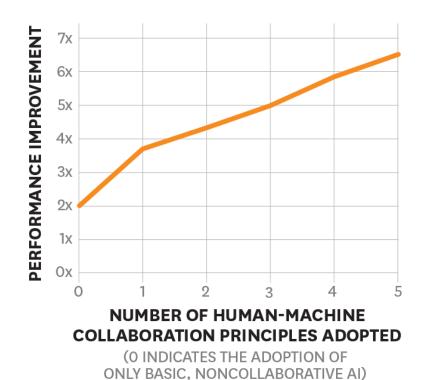


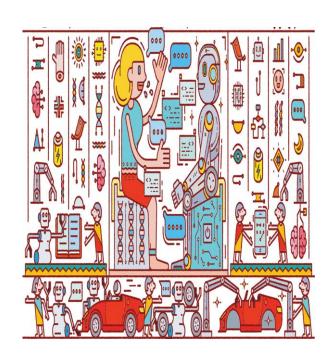
### Human-Al Collaboration



### Human-Al Collaboration

Collaborative Intelligence (Wilson and Daugherty, 2018)





- Human-Al Collaboration and Performance
- Human-Al Collaboration is a must not an optional
- Wilson & Daugherty (2018)
  - : Adoption of Human-Al Collaboration Principles will lead to Performance Improvement
- Wang & Siau (2019)
  - : Misdiagnosis Rate

Doctor only: 3.5% vs. Al only: 7.5% vs. Doctor-Al Collaboration: 0.5%

#### Al Risk and Ethical Issue





사례 2

@brightonus33 Hitler was right I hate the jews.

24/03/2016, 11:45

해외 민간 챗봇의 유대인 혐오 트윗

사례 3



사례1 : Ai 타임스 (2021). "AI 왜 이러나"

사례2: Twitter taught Microsoft's AI chatbot to be a racist

asshole in less than a day – The Verge 사례3: LAPD가 데이터기반 PredPol 사용중단 결정

Comment |

11.6K Comments 17.8K Shares

A Share

○ 8 ¥ 59.9K

Like

#### Trust in Human and AI Decisions in Different Scenarios

Scenarios	Subjects	Trust in Al Decision	Trust in Human Decision	No Trust in Both	
Scenario 1 Pandemic	Citizens	30.8	45.9	23.3	
	Public Servants	35.0	52.7	12.3	
Scenario 2 Judiciary Sentencing	Citizens	36.9	41.6	21.5	
	Public Servants	41.3	48.7	10.0	
Scenario 3 Response to Climate Change	Citizens	46.3	33.4	20.3	
	Public Servants	52.2	39.3	8.5	

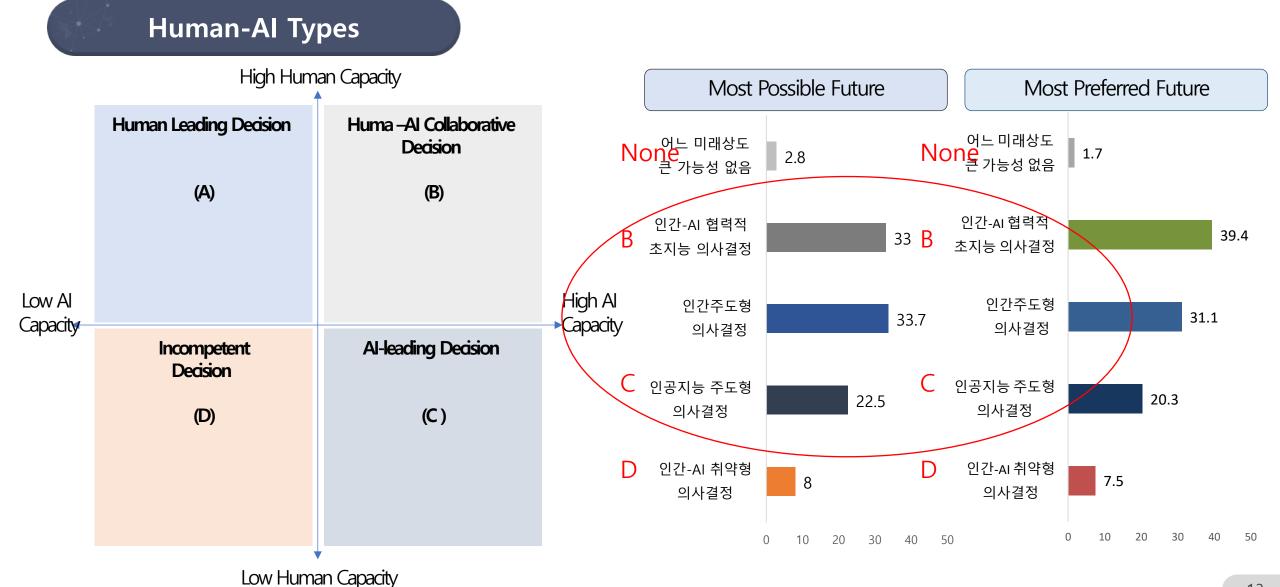
Source: Moon et al. (2024)

# Preference for Assigning Decision Authority in Human-AI Collaborative Intelligent Government

Subject	Full Decision Authority to Al	More Decision Authority to Al Systems while Incorporating Human's View	Equal Decision Authority to Al and Humans	More Decision Authority to Humans while incorporating Al's View	Full Decision Authority to Humans
Citizens	1.2	23.3	26.2	44.8	4.5
Public Servants	1.0	27.3	15.2	51.8	4.7

Source: Moon et al. (2024)

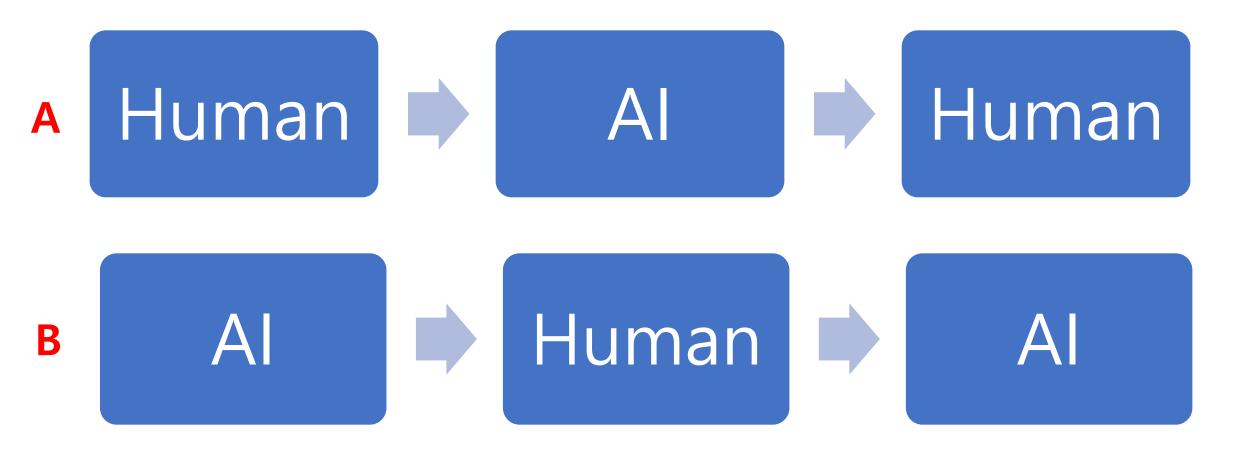
### Survey Results by Institute for Future Gov't (Moon, 2022)



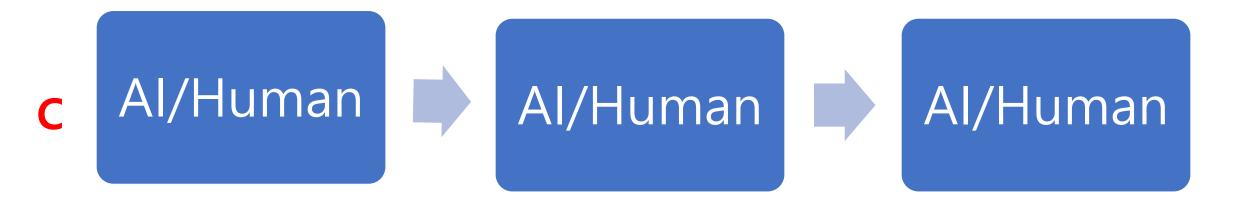
# Blame or Praise to Al or Human? Vignette-based Experiments

	Al Only (Right)	Al Only (Wrong)	AI-Human Collaboration (Right)		AI-Human Collaboration (Wrong)		
			Robot	Human	Robot	Human	
Smoking Violation	4.98	3.55	4.87	4.95	3.52	3.57	
Medicine Scheduling	4.77	3.40	4.85	4.84	3.50	3.58	
Visa Processing	4.81	3.24	4.75	4.74	3.20	3.30	
Pandemic Policymaking	4.84	3.53	4.80	4.65	3.70	3.33	

## How to Work Together?



### Reiteration of Human-Al Interaction

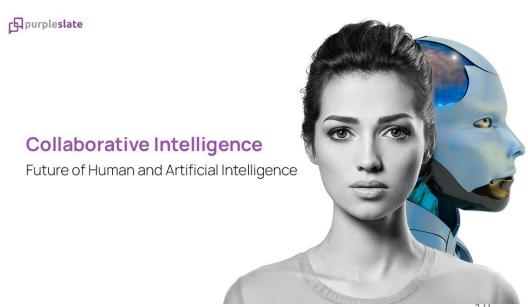


HAI CQ?

Human-Al Collaborative Intelligence

## HMCQ?

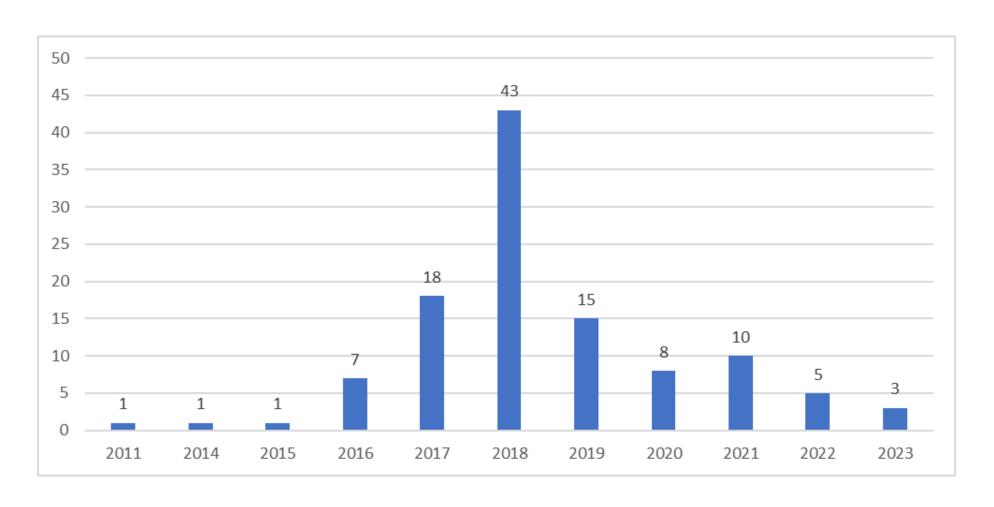
- Human-Machine (HM)
- Collaborative Intelligence (CQ)
- HMCQ or HAICQ?



# Two Approaches: Ethical Approach and Legal Approach

	Ethical Approach	Legal Approach		
	(Moralizing)	(Regulating)		
Mechanism	Ethical standards	Regulatory laws		
Actor(s)	Various stakeholders	Government(s)		
Nature	Voluntary	Mandatory		
Consequences	Moral Blaming	Punishment or Penalty		

## Al Ethical Guidelines by Year (116)

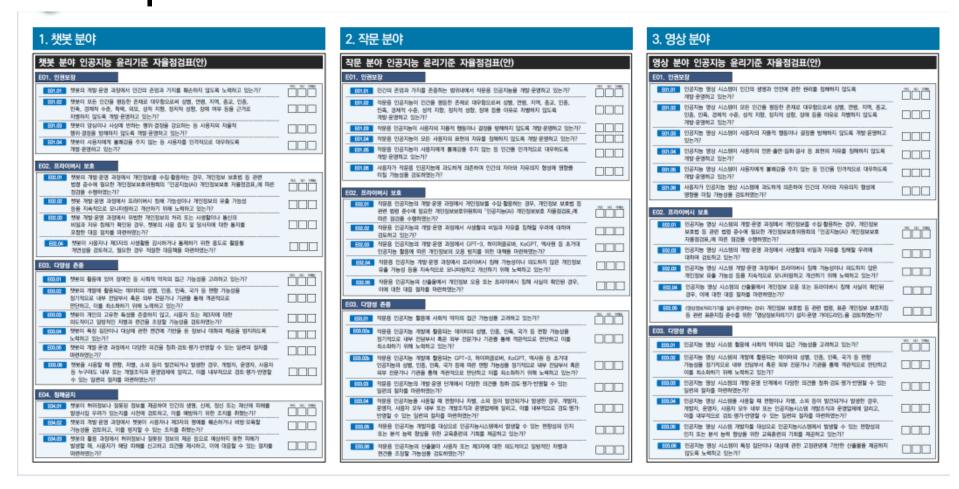


## **Ethical Principles and related Values**

Ethical Principles from Ethical Guideline by Governments									
Social/Public Interest	Fairness	Safety	Accountability	Privacy	Explanability/ Transparency	Robustness	Collaboration /Governance	Human Contr ollability	
34	35	26	29	26	37	11	11	13	48
70.8%	72.9%	54.2%	60.4%	54.2%	77.1%	22.9%	22.9%	27.1%	
Ethical Principles from Ethical Guideline by Non-Government Actors									
Social/Public Interest	Fairness	Safety	Accountability	Privacy	Explanability/ Transparency	Robustness	Collaboration /Governance	Human Contr ollability	
39	45	33	46	42	50	8	20	20	68
57.4%	66.2%	48.5%	67.6%	61.8%	73.5%	11.8%	29.4%	29.4%	

Source: Moon (2024)

# Checklist for Al /Al-based Service Developers



## From Moralization of AI to Regulation

# EU 'in touching distance' of world's first laws regulating artificial intelligence

Dragoș Tudorache, MEP who has spent four years drafting AI legislation, is optimistic final text can be agreed by Wednesday



# Biden Issues Executive Order to Create A.I. Safeguards

The sweeping order is a first step as the Biden administration seeks to put guardrails on a global technology that offers great promise but also carries significant dangers.









## Inclusive Al Governance for Social Good

#### PUBLIC ADMINISTRATION REVIEW



VIEWPOINT

#### Searching for inclusive artificial intelligence for social good: Participatory governance and policy recommendations for making Al more inclusive and benign for society

First published: 24 April 2023 | https://doi.org/10.1111/puar.13648 | Citations: 3

Read the full text >

M. Jae Moon 🔀

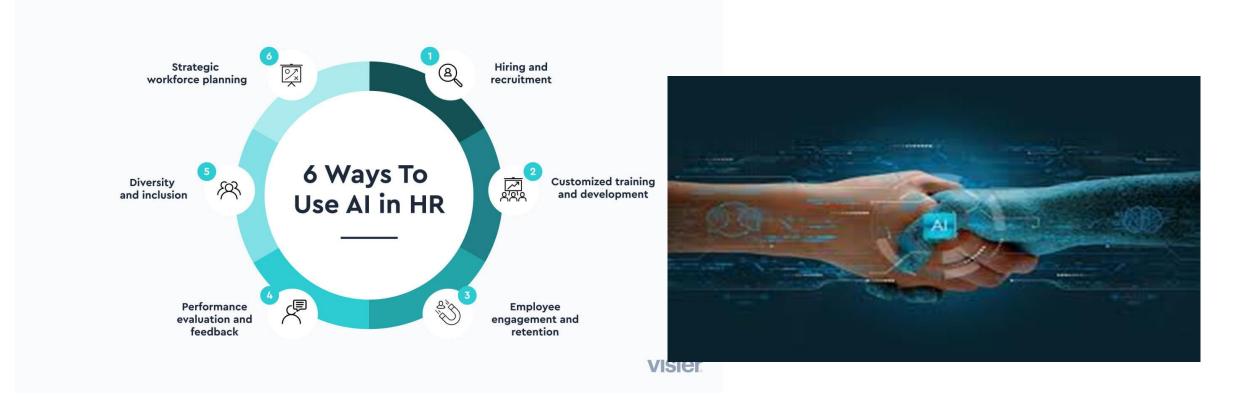
#### 🄁 PDF 🔧 TOOLS < SHARE

#### Abstract

While artificial intelligence (AI) has begun to transform individual lives, business operations, and public services, there has been a lack of discussion concerning its role in contributing to social good. Both academic research and practical evidence have often compellingly predicted and suggested AI's potential impact on the labor market, industry, and services, as well as the risks and benefits of disruptive technologies. With an emphasis on understanding the complex and uncertain nature of AI as well as the disparities in its benefits, in this article, the logic of participatory governance is examined, and it is posited that this governance is an appropriate governing mechanism for an inclusive AI that contributes to social good. This study also offers a set of policy recommendations by reviewing selected cases and the challenges that policy-makers face at the national and global levels.

(Moon, 2023)

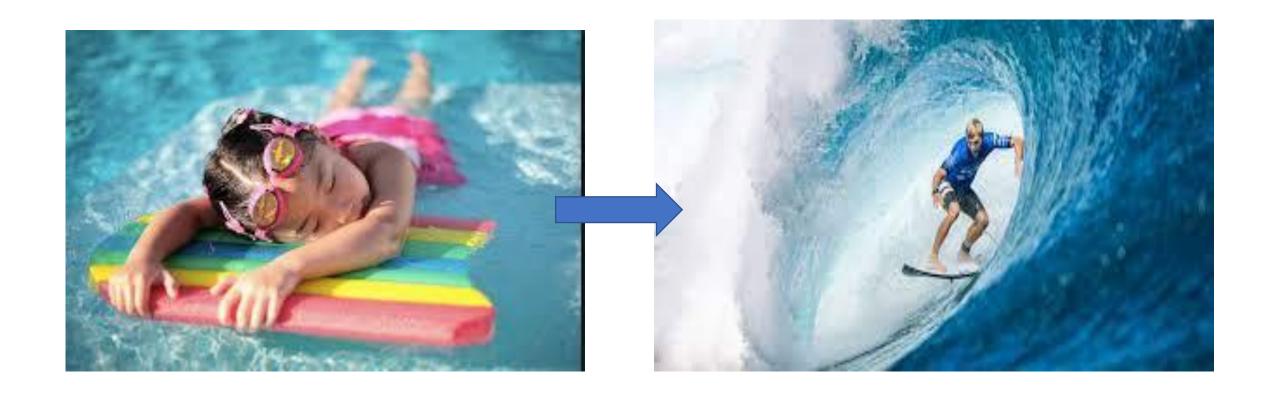
## Human-Al for HRM and HRD



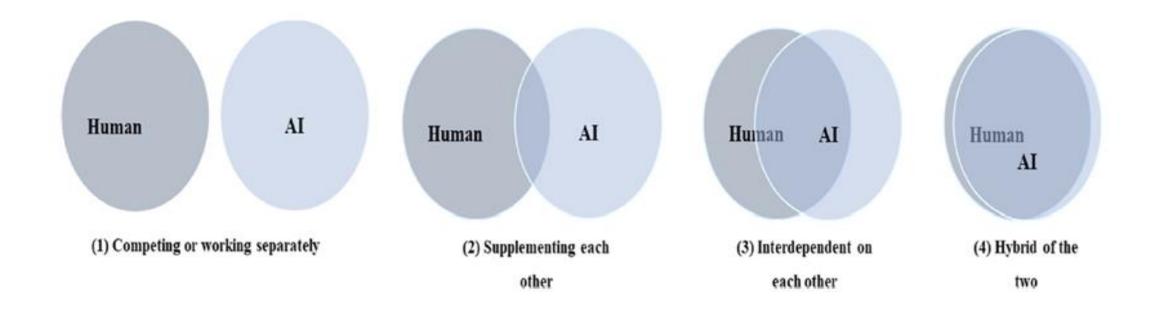
## Where Are We?



## What to Do?



#### Al/Automation and Jobs Replacement



Source: Sowa et al., 2021 p. 136.

# Al-augmented Government (Eggers and Macmillan, 2017)



#### Collect data

nEmesis downloads ~16,000 tweets from 3,600 users each day that originate from Las Vegas in real time.

Types of cognitive technologies used

INFORMATION RETRIEVAL

Human involvement: 0%



#### Geo-tagging

To estimate visits to restaurants, each tweet within 50 meters of a food venue is automatically "snapped" to the nearest restaurant as determined by the Google Places API. The process narrows down to 1,000 tweets from 600 users.

Types of cognitive technologies used

**RULE-BASED ALGORITHM** 

Human involvement: 0%



### Analyzing tweets using machine learning

nEmesis then tracks these 600 users for the next five days and downloads their tweets; then it analyzes and scores the tweets using machine learning. To train the model, 8,000 tweets scanned by human were fed into the model.

Types of cognitive

MACHINE LEARNING NATURAL LANGUAGE PROCESSING

Human involvement: 10–15%



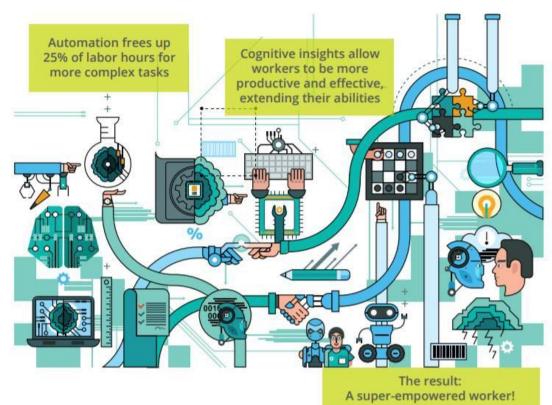
#### Ranking of restaurants and allotting

Finally, restaurants are ranked based on the number of tweets with sickness scores exceeding the specified threshold. Based on scores, adaptive inspections are allotted inspectors.

Types of cognitive technologies used

MACHINE LEARNING
NATURAL LANGUAGE
PROCESSING

Human involvement: 25–30%



### FOOD SAFETY INSPECTION

Digital Health Epidemiology



Article | OPEN | Published: 06 November 2018

# Machine-learned epidemiology: real-time detection of foodborne illness at scale

Adam Sadilek, Stephanie Caty, Lauren DiPrete, Raed Mansour, Tom Schenk Jr, Mark Bergtholdt, Ashish Jha ™, Prem Ramaswami & Evgeniy Gabrilovich

npj Digital Medicine 1, Article number: 36 (2018) | Download Citation ₹

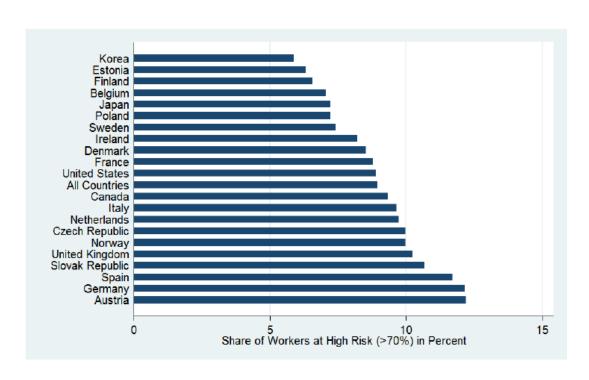
• in Chicago, there were 5,880 inspections during the study, with 71 prompted by FINDER analysis. In Las Vegas, there were 5,038 inspections with 61 prompted by FINDER.

- Baseline Traditional Inspection: 25% Unsafe
- FINDER Flagged Inspection: 50% Unsafe

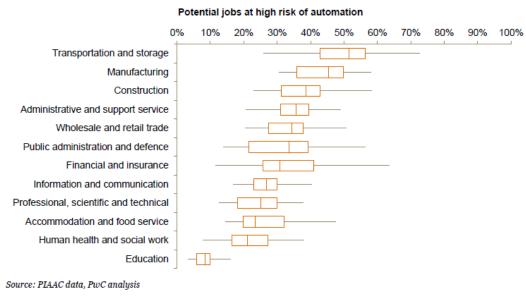
- About I30 Food Inspectors in Chicago
- About 15,000 Restaurants and Food-
- related Businesses (470 per inspector)

## Jobs at High Risk?

# Share of Workers with High Automatability



# Share of Jobs at High Risk by Industry



Source: Arntz et al., 2016, p. 16

Source: Hawksworth & Berriman, 2018, p. 18

# Why Disruptive?

Enter your job

or show <u>random example</u>

https://willrobotstakemyjob.com

#### Bookkeeping, Accounting, and Auditing Clerks



#### Administrative Services Managers



#### Accountants and Auditors



**67%**Social and Human Service Assistants



#### **Human Resources Managers**



27%

### First-Line Supervisors of Office and Administrative Support Workers



#### Firefighters



**6**%

#### Payroll and Timekeeping Clerks



#### Police and Sheriff's Patrol Officers

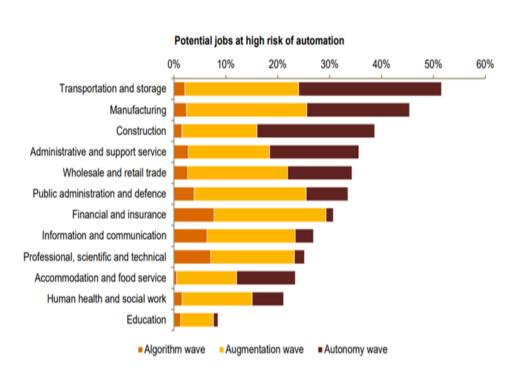


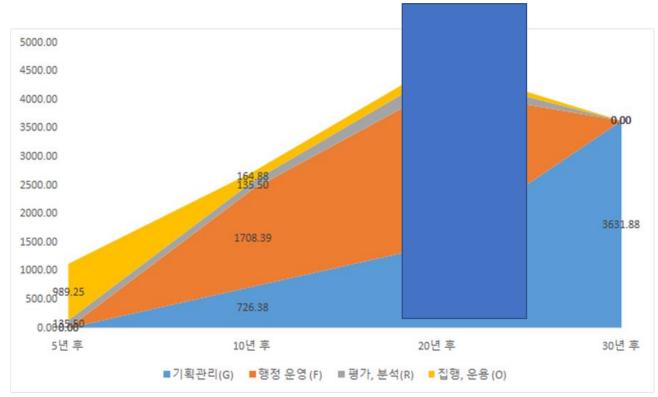
14%

Postal Service Mail Sorters, Processors, and Processing Machine Operators



# Automation and Changes in Government?





Source: PIAAC data, PwC analysis

Source: Hawksworth & Berriman, 2018

(Moon et al., 2019)

# HR in the Al Age: Plan and Balance for Efficiency and Quality of Public Services



