Workshop Summary UNDERSTANDING PUBLIC PERCEPTIONS FOR APPLIED DATA SCIENCE

2021 SIG-KDD Conference | 15 August 2021



 $\mathbf{K}\mathbf{\Lambda}$

Lloyd's Register Foundation Institute for the Public Understanding of Risk





Table of Contents

- **02** Contents
- **03** Workshop Overview
- 04 Workshop Programme
- O6 Panel Session One: Is there a need to engage the public in technology development?
- 13 Presentations
- **18** Panel Session Two: How to engage the public in technology development

Workshop Overview

The 2021 KDD Workshop on Understanding Public Perceptions for Applied Data Science (UPP4DS) was held on 13 August 2021. Over thirty researchers, practitioners and civil society representatives joined the workshop to examine the role of society in the development of acceptable technologies.

Research on data science is often approached from the angle of making Al more reliable, accurate, explainable and actionable. These outcomes contribute towards the broader objectives of making Al more trustworthy to facilitate its adoption and use. However, trust is a multi-faceted issue which is influenced by a variety of factors, including understanding, familiarity, perceptions of risk and credibility.

In the development of AI technologies, the public can be a resource and play a positive role in contributing their data for the advancement of data science, or be a barrier to deployment if they do not accept the technology or solution due to their misperceptions of risk.

The role of the public is particularly important in industries such as healthcare and consumer products, where the public as end-users have closer proximity to the AI application and significant implications on its adoption. Apart from making the science more understandable to the public, there is increasing need for the science to better understand the public as well.

When and why is it important for experts to understand the public perceptions of risk? How can data scientists work with implementation partners to better understand end-users and address concerns which may inhibit adoption? These questions were explored in the inaugural UPP4DS workshop organised in-conjunction with the prestigious 2021 SIG-KDD conference.

Workshop Programme

9.00 am

Welcome Prof. Chan Ghee Koh

9.05 am Panel Session 1: Is there a need to engage the public in technology development?

Prof. Dean Ho: Optimising Healthcare with Digital Medicine: Seamlessly Integrating Human-Technology Engagement

Prof. Klaus Wertenbroch: Engaging the Public in Technology Development Based on Personal (Private) Data

Prof. Edward Choi: Engaging the Public in Technology Development - The Healthcare Al Perspective

Prof. Leonard Lee: Panel Discussion Moderator

10.05 am Presentations

Dr. Aileen Nielsen: Private Sector Regulation Increases Contact Tracing App Uptake

Francois Buet-Golfouse: "Art Meets Science": Tackling Data and Perceptions

Dr. Cornelius Kalenzi, Dr. Hyeondae Rha, Nathaniel Tan, Dr. Moonjung Yim: Using Al to Support Healthcare Decisions: A Guide for Society

Workshop Programme

11.00 am

Panel Session 2: Mechanisms to engage the public in technology development

Tracey Brown: Developing the public guide to Al in healthcare

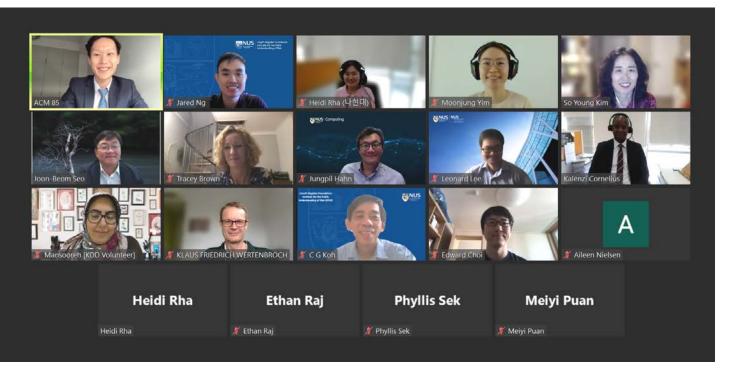
Prof. Jungpil Hahn: Engaging the Crowd in Technology Development

Prof. Joon Beom Seo: What the public should know from: "The White Paper on Medical Al" by Korean Society of Artificial Intelligence in Medicine (KoSAIM)

Dr. Cornelius Kalenzi: Panel Discussion Moderator

12.05 pm Closing

Prof. So Young Kim



Workshop Summary pg. 5

Optimising Healthcare with Digital Medicine: Seamlessly Integrating Human-Technology Engagement

Prof. Dean Ho Director, the N.1 Institute for Health, and the Institute for Digital Medicine, NUS

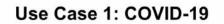
Prof. Ho shared that there is a need to have a shift in thinking about the methods to develop therapeutics and how they are administered. Technology alone cannot transform healthcare and human collaboration and engagement has an important role to play.

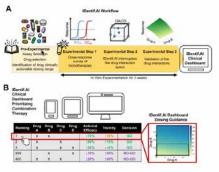
Al is already being used in drug discovery. However, to optimise these treatment interventions, working closely with scientists and healthcare providers to acquire data for calibration is essential. Human engagement is particularly required for the acquisition of "small data".

It is not always about how much data you have but how the data is acquired. Small data approaches in collaboration with key stakeholders are used to drive drug development and the use of a patient's own data to manage their care. However, the approaches are complementary, big data approaches are used to help understand what drugs should be used in the first place.

> Moving forward, even in public engagement, how we think about clinical trials will have to change – the number of patients we recruit may not be as many as we think we need to derive true confirmation of the effectiveness of new therapies. Seamless and community-based collaboration will allow us to completely change how we practice medicine.

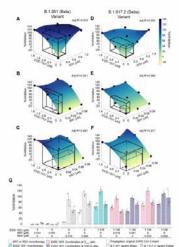






 Harnessing IDentif.AI to optimise regimens against Wildtype, Beta, and Delta Variants
Novel workflow developed integrating clinician-technology-laboratory domains.

Blasiak et al., Bioengineering and Translational Medicine, 2020. Blasia et al., MedRxiv, 2021.





Use Cases

These use cases illustrate a seamless integration of doctors, nurses and medical engineers. Patients are exposed to variable dosing and corresponding efficacy to drive optimisation through an Al-discovered correlation but also through human collaborations where the interesting findings emerge.

COVID-19 Treatment	Working with clinicians to optimise treatment regimens against the different variants of COVID-19. Candidate therapies are first picked and live experiments in the lab are then conducted on actual coronavirus samples. Using real-data on the coronavirus, IDentif.Al derives optimal combinations. The process involves speaking to doctors first, followed by validation of the data, before circling back to the doctors.
Advanced Solid Cancer	Using a cancer patient's own data, CURATE.AI recommended substiantial drug dosage reduction to increase efficacy and tolerability. This helped the patient achieve an active lifestyle and halted his disease progression. This approach aimed to mediate improved efficacy by using the patient's own data.
Transplant Medicine	Identifying drug synergy alone is not enough for optimisation - drug antagonism is also dose-dependent for each patient. The results of CURATE.AI's intra-patient optimised dosage showed that synergies and behaviours changed dynamically and patients treated were able to be discharged about a month earlier than those from standard care. This also demonstrates the need for re-optimisation to take place at different stages of drug development process.

Engaging the public in technology development based on personal (private) data

Prof. Klaus Wertenbroch The Novartis Chaired Professor of Management and the Environment INSEAD (Singapore)

Prof. Wertenbroch discussed the acquisition of big data that technology is often based on and the extent to which consumers need to be involved, and subsequently the topic of privacy and the extent to which consumers want to remain in control of that. He highlighted three relevant findings on topic:

01 Threat to Identity

Technology adoption depends to some extent on whether or not consumers feel threatened in their identity. For example, someone whose identity is tied to cooking may not adopt a device that cooks autonomously.

02 Algorithm Aversion

The findings of so-called algorithm aversion are nuanced. In the medical context, consumers do not like algorithms unless they are more accurate than humans, which they are typically. People also do not like algorithms if they see how algorithms err.

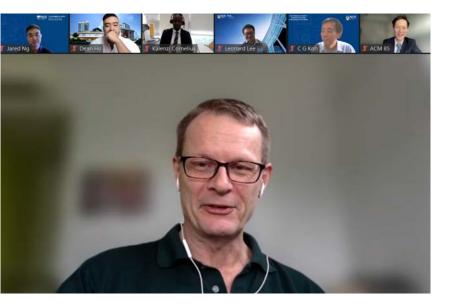
03 Privacy Paradox

People always say they want to protect their privacy but they often don't. It's a difference between stated and revealed preferences: what people say and what they do. People state that they prefer privacy because they don't trust data controllers but they fail to implement privacy protection measures because they are not aware of what happens to the data they share.

Involvement by design

Prof. Wertenbroch opined that the public is involved in algorithm development by design, such as in medicine, health care, marketing, consumer finance, and government because algorithms are trained on data that consumers (i.e., the public) provide. Companies are using these algorithms to make predictions for their sales to generate profits.

He observed that consumer decisions about their private data may not always be rational. In experimental studies that detected violations of "transitivity" – an axiom of rational choice and foundation of utility theory – Prof. Wertenbroch and his colleagues found that consumers undervalue their privacy when they barter private data for goods such as e-books and gasoline, compared to when they sell their private data for cash.



"consumers undervalue their privacy when they barter private data for goods"

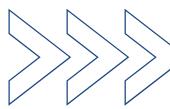
Suggestion

Create a Market-Based Solution

- Lawmakers and regulators assign outright legal ownership of data to consumers.
- Consumers empowered to sell their data to companies that are willing to pay the most, which increases competition between companies.
- System enables innovative and pioneering companies to rise up the scene and develop technology.

• System would require some form of regulation to function efficiently, along with technology to implement it.

Panel Session One Engaging the Public in Technology Development: The Healthcare Al Perspective



Prof. Edward Choi, Assistant Professor, Graduate School of Al, KAIST

After giving an outline of different classifications of Al based on their impact, Prof. Choi noted that many people expect that Al will soon be actively used in the healthcare industry (clinical service, hospital administration, drug development). Some will directly impact the public (Al-aided diagnosis), some will indirectly impact the public (Al-aided inventory allocation).

He opined that the public seem to be safe from Al-induced harm thanks to the "system" (usually in the form of government approval (e.g. clinical trials), but the "system" alone may not be enough. In clinical trials, there may need to be new approval processes instead of just the "FDA approved" Al that many are familiar with but not quite clear on the process. However, if there were to be a new approval process, should the mechanisms of the Al be transparent to the public?

Al Classifications by Impact

Impact	Little or no consequence	Grave and cannot afford to fail
Al Applications	YouTube or Netflix Recommendations	Self-driving cars, pathology and diagnosis, and advanced weapons

Engaging the Public in Technology Development The Healthcare AI Perspective

1 - 0 0 a

KDD 2021 Workshop on Understanding Public Perceptions for Applied Data Science

> Edward Choi Grad School of Al, KAIST



Prof. Choi used an example of visible and invisible Als. Visible Als will go under an approval process before launch. Invisible Als may not have to go through such a process because they do not directly impact the patient, such as inventory allocation where Al software decides which medical department receives more medication in advance.

In order to build public trust and education, Prof. Choi asked then whether a hospital or an organisation should disclose all Al-related information, and whether all medical-related Al should take responsibility for its actions. He concluded by stating that engaging the public in the Al development process is not easy because many do not understand Al technology, and IT companies do not want to disclose their technology. Society needs to identify commonalities between what the public cares about the most and what determines Al performance in order to bridge the gap. Society needs to identify commonalities between what the public cares about the most and what determines Al performance in order to bridge the gap.

Panel One Discussion

Moderated by Prof. Leonard Lee



Information Goods vs Money

When asked by Prof. So Young Kim if his experiment created an artificial association by comparing privacy with two alternatives (information goods and money) instead of just privacy and money, Prof. Wertenbroch said that consumers need to value goods in terms of money but companies are not taking that approach, therefore consumers need to have some level of monetary value of the goods in their head that they are getting in exchange for their privacy.

Irresistable Value

Prof. Choi commented that many powerful companies such as YouTube or Google offer irresistible goods or services to consumers, so much so that they are willing to give up their privacy data and increase the power and growth of mega companies. Prof. Wertenbroch noted that the situation mentioned by Prof. Choi reflected the current market monopsony where there are a handful of sellers and a lot of buyers. This removes the competition and innovation of technology because smaller companies are not able to compete.

Future of Clinical Trials

When asked how clinical trials could evolve in accordance to development of the novel innovations and techniques, Prof. Ho outlined that patients should be given a short and variable dosing regime and that the calibration dosage for every patient should start at the beginning. This will help clinics and hospitals capture more respondents and improve treatment.

Public Private Regulation for Technology Uptake: The Case of Digital Contact Tracing (DCT)

Dr. Aileen Nielsen, ETH Zurich

Dr. Nielsen discussed how the adoption of DCTs varied across countries and even even the most successful countries with the highest download rates widely underperformed and failed to achieve required adoption figures. Their research was motivated by the question: What realistic options are available to governments given the prevailing historic low levels of trust towards the government? What is going to drive the responses here and are there any general lessons to be learnt?

Their research constituted an online survey conducted on a representative sample of U.S. adults. DCT vignettes were provide to participants to assess how their download rates would differ based on the regime participants were provided. There were five different regimes applied: voluntary, employer mandate, employer encouragement, retail mandate and government mandate. The survey also had questions about trust in government and trust in business community, and questions about private and technology attitudes. Some of these regimes (such as employer mandate and retail mandate) are relevant today for new technologies such as health passports.

The results showed similar reactions across all regimes – low uptake. A complementary regime of public and private initiatives achieved the highest uptake. People who expressed a higher trust/confidence in the government were equally or more likely to download the app under any regime – they had a higher willingness to download. The regime that is chosen will not move this group.



Different approaches for different groups

People who have more confidence in business were more likely to download the app when it was required of the employer. Trust differentials had an effect on the public/private regime and this could explain which regime would encourage downloads by which population groups.

Dr. Nielsen noted that political polarisation is not inevitable and not crippling for DCTs. She concluded that governments have meaningful options to increase DCT uptake through promoting private sector action. Trust differentials are a major source of heterogeneity in population for DCT uptake.

Discussion

Prof. Hahn noted that very slight changes in the design of the DCT app can bring about a marked difference, as well as the way the data is stored (centrally or de-centrally). He asked how such changes could make a difference in the uptake. Aileen noted that when the mandate to adopt DCT came from an employer, there was a small effect on the uptake but more assessments would be needed on different mandates to understand if it was enough. Her study showed that there was an initial preference for decentralised storage if trade-off information (how the data would be used) was provided.

Dr. Cornelius Kalenzi commented that there is a diversity in how countries are taking up contact tracing and asked what the reasons behind the differences of uptake behind these differences. Aileen outlined that Americans are far more comfortable in taking orders from their employers rather than their government, so there are cultural aspects that influenced the results. She opined that certain fundamental beliefs could not be changed easily which leads to consistent hesitancy towards adoption.

Presentations

Art Meets Science: Tackling Data and Perceptions

Francois Buet-Golfouse, University College London

Francois presented a conceptual model which enables the usage of both human intelligence and AI to derive an optimal result. This model utilises the traditional concept of "Mixture of Experts" (MoE) where multiple experts who are fairly independent and have different strengths, and modern concepts of distributed, federated and collaborative learning.

He highlighted that the performance of AI models has been improving and even surpasses humans in various tasks such as image classification, diagnosis and chatbots. These solutions are also at times very easily scalable and there are benefits in utilising Al models. However, AI algorithms still has its limitations and still require human expertise and strengths such as domain expertise, transfer learning, communication and empathy.

The model incorporates human learning into MoEs. He noted that the challenge is in reconciling the different rates of learning between algorithms and humans. In the model, weights are applied on the inputs from different experts to derive a prediction and the model can be refined through weight reassignment and updating of model parameters.

In his demonstration of a simple application of the model, Francois highlighted that there is a need for both forms of intelligence and crux lies in how the model can combine different rates of learning from humans and Al and find a right "rhythm" on how to make this feasible.

≜UCI

Human Intelligence

BUT

- Subhuman performance of certain algorithms; Issues around productionisation of ML algorithms;
- Lack of data;
- Biases in or limitations to datasets; Stress-tests and rare events;
- Human in the loop (e.g., examination of results);
- Transfer learning;
- Qualitative data (patient knowledge vs cross-section); Human communication and empathy;



Presentations

Using Artificial Intelligence to Support Healthcare Decisions: A Guide for Society

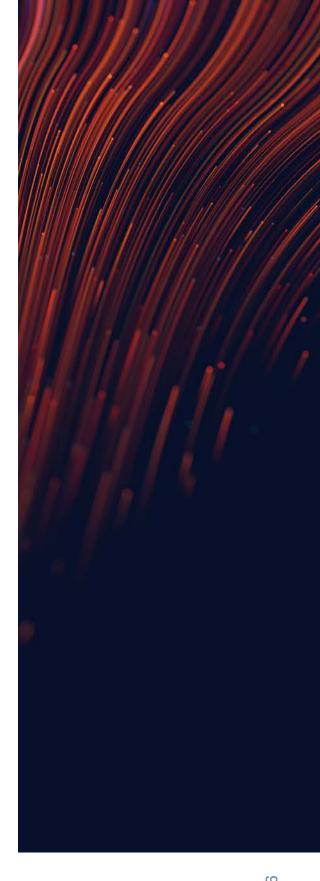
Dr. Cornelius Kalenzi, KPC4IR Dr. Hyeon Dae (Heidi) Rha, KPC4IR Nathaniel Tan, IPUR Dr. Moonjung Yim, KPC4IR,

The project team launched "Using Artificial Intelligence to Support Healthcare Decisions: A Guide for Society" and spoke about its development process.

Al advancements have been taking place rapidly and quickly. In healthcare, its Al applications can be categorised as clinical-decision support tools, patient-decision support tools, healthcare administration and therapeutics development.

As with all new technologies, Al is not without its risks and issues. These include data and privacy, security and control, and also other implementation challenges. One case study is that of a country investing significant resources to adopt an Al system, only to find that the system was built on data which was not representative of the country – this system was not generalisable for that country. The guide focuses on issues of reliability and aims to provide some useful guidelines and tools to motivate and equip the public (journalists, policy-makers, healthcare agencies, doctors and patients) to engage in deliberations around the reliability of Al.

This guide is an international collaboration between NUS IPUR, KAIST KPC4IR and Sense about Science (SaS), a UK based charity working on the promotion of science. Its development took one year and the project team interviewed more than 30 experts in Al development, governance, healthcare practitioners and policymakers.





The guide provides readers an introduction into the terms used in Al discussions, an overview of the status of AI development and its types of applications, a discussion on the need to focus on reliability, and provides some principles and questions on how to assess its reliability. These questions include:

Key		acti	one
ney	Que	esu	UIIS

Relevant questions

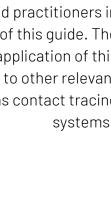
scrutinized?

What Data is the Al Based on?	How was the data used to train the Al collected?	Does the data represent the patients whom the AI is being used?	Are the patterns and relationships identified by the Al accurate?
What assumptions are the Al making about patients and the disease?	ls the right relationship captured? Are the results generalizable?	Are the variables excluded from the model actually irrelevant?	Does Al eliminate human prejudice from decision- making?
How much decision weight can we put on	How well does the Al really	Has its reliability been properly	Does it make a useful real-world

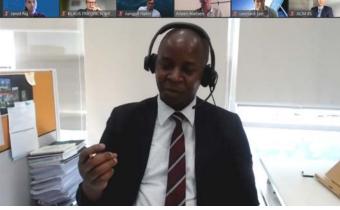
perform?



Problems will arise if the quality of data is not properly scrutinised and Al's reliability hasn't been tested. The guide enables discussions on the standards AI should meet and the next step in dissemination would be engaging relevant policy-makers and practitioners in the adoption and utilisation of this guide. The team is also exploring the application of this engagement methodology to other relevant technologies such as contact tracing systems.



recommendation?



the Al?

Panel Two How does Sense about Science engage the public Tracey Brown Director, Sense About Science

Tracey shared how her organisation, Sense about Science (SaS), has 20 years of experience working on the frontlines on how to engage the public in some very difficult issues – e.g. GMOs, vaccinations and radiation – areas which have caused some public concern or a breakdown on conversation and relations between scientists and society, and public and large. She noted that ideally, this engagement takes place with people at the start before the conversations turn sour. She presented some examples of their work to equip the public, policymakers and researchers for productive conversations:

Ask for Evidence

A campaign which focuses on empowering the public and journalists to ask important questions about evidence

Talk about reliability

It is important to include the information about reliability and quality assurance processes

Communicate the research process

Encouraging researchers to explain peer review and other processes of validation of scientific findings, and also clinical trials and methods: researchers often forget that the public are not so familiar with these processes

Encourage questions

Work with people who are less willing to accept scientific findings, to encourage them to ask good questions, for example on Covid-19.

Tracey outlined that SaS uses a public-led, expert-fed approach. The approach starts by understanding the issues as they present themselves in society. Instead of starting with a produced research result and thinking about how to best communicate it through a language change, the communication strategy to the target audience should be developed in consultation with the target audience – think about how to involve the people who you want to communicate with, in working out how to communicate (e.g. if you want to communicate to journalists, they will tell you what interests them, what concerns they have, what questions they need help with.

UPP4DS

Tracey outlined that SaS uses a public-led, expertfed approach. The approach starts by understanding the issues as they present themselves in society. Instead of starting with a produced research result and thinking about how to best communicate it through a language change, the communication strategy to the target audience should be developed in consultation with the target audience – think about how to involve the people who you want to "researchers and institutions agree with the need for public engagement but they find exceptions on the need to communicate their work."

communicate with, in working out how to communicate (e.g. if you want to communicate to journalists, they will tell you what interests them, what concerns they have, what questions they need help with.

There is the need to look at the characteristics of the wider discussion; how the issue is presenting itself to the public, and then involve the audience in figuring out what is needed (show them what the work is about and let them tell you what they want to know, what questions they may have and what they find surprising). The approach also sees how to present the information to people in a manner which is accessible and meaningful. This is done through user testing and verification.

In Tracey's view, researchers and institutions agree with the need for public engagement but they find exceptions on the need to communicate their work (e.g., they believe the issues are too difficult or sensitive). An example of a sensitive issue in the UK was in the 90s on the performance of children's heart surgery which led to a breakdown between parent groups, patient groups and the medical profession around heart surgery. The government's response was to publish data on surgical outcomes of children's heart surgery and give it to parents to make their decisions on where to bring their children for surgery. There were a lot of misunderstanding and unintended consequences and there was a desire to improve the communications.

SaS' process was utilised in engaging the parents and the end product (web interface) was designed by parents. When it comes to the dissemination, the parents launched it themselves because they were enthusiastic in wanting other people to do know the information.

People also cite issues being too complex. The Small Area Health Statistics unit in Imperial College monitors health conditions and environment agents, and publishes these statistics. As such, there is a real risk of journalists and policy-makers simply making inferences between the data and jumping into conclusions, creating "scary" stories. SaS worked with people in media and in the community and involved them in designing the website. Using the insights from many years of these kinds of projects, Sense about Science more recently created a public guide to data science. Some of the questions in this formed the first steps of the newly launched guide to Al guide in healthcare. Dissemination can feed and fuel into new areas of research.

Engaging the crowd in technology development

Prof. Jungpil Hahn Associate Professor, NUS School of Computing

Prof. Hahn spoke about how innovation and technology development has been taking place around the world and this is partially through public involvement – lessons can be learnt on how these platforms of crowd participation can be established.

He noted that one pitfall is that there is a false dichotomy created between specialist and the public where there is a need to "sugar coat" technologies to make it understandable or acceptable. Instead of separating the two groups and looking at the public as a consumer of technology, Prof. Hahn suggested viewing the public as a "prosumer or producer + consumer" and can contribute to value in the technology development process. In this way, you can view the group as the "crowd". Some examples of technologies which utilised crowd sourcing include:



The Linux system has 20k contributors who were not paid.

pebble

NETFLIX

Pebblewatch mobilised financial capital and raised \$30M from 150k people to bring the first electronic watch to market.

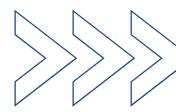


Starbucks has crowdsourcing platform to generate ideas from the public. They received 30k – 40k ideas from the public. Netflix's open challenge to increase predictive accuracy saw 40k teams vying for the 1M prize. The final solution exceeded the 10% target improvement.

Prof. Hahn opined that the crowd can contribute money, work and ideas (capital, labour, and entrepreneurship). The crowd can offer these factors of production and a lot of tech innovation can take place through the crowd. This is a potential means to engage the public.

So what makes the crowd effective? Prof. Hahn likened technology innovation to a rugged landscape with different peaks (local peaks) which are reached through continued improvements by an individual participant. The crowd consolidates a landscape of different local peaks identified by multiple participants with diverse backgrounds to produce a portfolio of solutions.

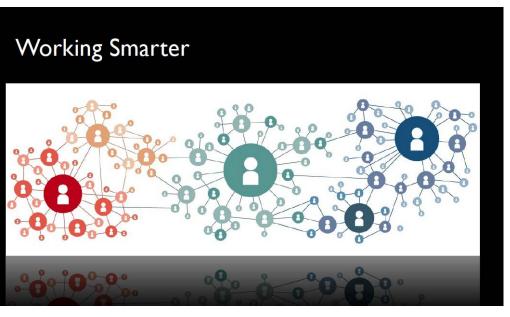
Mobilising the Crowd



"research has shown that this is not a pure economic problem"

To mobilise the crowd, Prof. Hahn said that they could be incentivised through money but research has however shown that this is not a pure economic problem. If it's just money, there won't be the diverse participation as logic dictates that less proficient individuals would not participate as they do not expect to win. Other motivations include love and glory, these play a greater role in motivating the crowd.

To make the crowd smarter and better at what they do, Prof. Hahn noted that there is a huge social element in open-source development. People work together and repeatedly work together because of their knowledge and positive working relationship. Making the work transparent also helps to take these findings to create new uses and build upon what they have – it reduces work duplications without explicitly addressing things like copyright.



Teams also perform better than individuals. There is a sweet spot between diversity and common ground. A lot of competitions make salient the ranking – these rank-based solutions can be harmful as people tend to gravitate towards other similar individuals and the team cannot maximise the diversity, said Prof. Hahn.

Panel Two

What the public should know from "The White Paper on Medical Al" by Korea Society of Artificial Intelligence in Medicine (KoSAIIM)

Prof. Joon Beom Seo Department of Radiology, University of Ulsan College of Medicine

Prof. Seo noted that there have been significant achievements in healthcare for Al but concerns about its safety and reliability remain. As such, experts have a responsibility to educate the public on what realistic expectations to have. He discussed a whitepaper by KOSAIM and said one of the aims was to integrate different Al issues in one document – so that medical professionals, lawmakers and doctors can understand new issues of Al applications in medicine. The issues arising from each application were summarised into four main categories

02

01 Usage of Healthcare Data

- Privacy and consent
- Data ownership who owns it, and who has control over it? Patients are unable to estimate the real value of their own medical data as they have no control of it
- How do we share profits from the sale of data?
- Regulation of applications and data is different between countries

03 Accountability and Responsibility

The current system is guideline-based but for AI, the recommendations are personalised and if there is a malpractice issue, the guidelines are not clear

04 Al Implementation

Implementation of Al into the current system is not easy and complicated

Safety and Reliability

For AI technologies to be more actively applied in healthcare, Prof. Seo suggested initiatives to increase the general public's awareness of AI, establish standardised guidelines, make systematic improvement, improve interdisciplinary collaboration, prepare effective education materials on the opportunity and risks of AI applications.

Workshop Summary pg. 22

Panel Two Discussion

Moderated by Dr. Cornelius Kalenzi

Prof. Wertenbroch commented to Prof. Seo that various stakeholders are always fighting over ownership of data. In Korea, the government is the default owner of the data due to the public health system. In marketing, it's usually the companies that like to claim ownership of the data. There is no jurisdiction in the world which is clear on who owns private data (not just medical data) that individual consumers generate. This is an important problem to crack – if the public feels that they own the resource which the technology is built upon, they might be open to participating in this and to contribute, and to accept technology.

Prof. Seo agreed with Prof. Wertenbroch's comment that healthcare data is not fundamentally different from private data but emphasised the complexity of the data management. The data is stored with hospitals and the public are not aware of the types of data generated and the value they hold. The public's awareness of the existence of this data could potentially be a hurdle for the usage of this data. There is a need for a social consensus on sharing of data and some tools can be created to return the control to the public.

The organisers would like to thank all speakers, participants and members of staff from KDD for making this workshop a success.

Lloyd's Register Foundation Institute for the Public Understanding of Risk

https://ipur.nus.edu.sg

Korea Policy Center for the Fourth Industrial Revolution

https://kpc4ir.kaist.ac.kr