MAI4CAREU | Master programmes in Artificial Intelligence 4 Careers in Europe

University of Cyprus

MAI643: Artificial Intelligence in Medicine

Kalia Orphanou

January – May 2023

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

Public Health Applications and Ethics

(Some material drawn from Hersh, WR, 2022. Health Informatics: Practical Guide, 8th Edition slides)

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Public Health Applications and Ethics

1. Public health definitions and history
2. AI in public health
3. Disease surveillance
   a. Social media posts
   b. Crowdsourcing data
   c. Other internet-based data
4. Privacy and security
   a. Terminology
   b. Risks in health applications
   c. Technological approaches to mitigate the issues
5. Discrimination and Bias
6. Future of AI in Public Health

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

INTENDED LEARNING OUTCOMES

Upon completion of this unit on Public health applications and ethics, students will be able to:

1. Explain what public health is
2. Discuss the application of AI methods in public health
3. Provide examples of surveillance data used in public health
4. Point out the current challenges on public health research
5. Understand the importance of ensuring privacy and security in AI health systems
6. Describe the technological approaches used to protect patient data privacy
7. Differentiate between the different categories of bias in health data
8. Point out the limitations of public health research
9. Discuss the future of AI application in public health

What is public health?

- The “science of protecting and improving the health of people and their communities. This work is achieved by promoting healthy lifestyles, researching disease and injury prevention, and detecting, preventing and responding to infectious diseases.” https://www.cdcfoundation.org/what-public-health
- “Public health informatics is the systematic application of information and computer science and technology to public health practice, research, and learning”, Yasnoff et al., J Public Health Manag Pract., 2000
- American Public Health Association (APHA), https://www.apha.org/what-is-public-health

From Hersh et al., 2022

Core functions and activities

- Public health performs its missions through its core functions
  - Assessment
  - Policy development
  - Assurance
- Public health activities include
  - Prevent epidemics and the spread of disease
  - Protect against environmental hazards
  - Prevent injuries
  - Promote and encourage healthy behaviors
  - Respond to disasters and assists communities in recovery
  - Assure the quality and accessibility of health services

From Hersh et al., 2022

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

Public health perspective

- Public health tends to take perspective of health of populations
- One of its basic sciences is epidemiology – study of disease in populations
- However, public health is increasingly involved in other forms of health promotion and prevention, e.g., obesity, nutrition, etc.
- May result in different perspective than individual care
- Population-based view focuses on preventing disease as well as societal impacts on health
- Usually a government (regional or federal) activity

From Hersh et al., 2022

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

A famous early public health story

- More details can be found here.
- John Snow was an early epidemiologist in the London in the mid-19th century.
- In 1854, he investigated a rapid outbreak of cholera in Soho area of London
- He found a common characteristic of those infected: use of water from the Broad Street pump.
- Early use of a geographic information system (GIS)

From Hersh et al., 2022

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Great public health achievements in 20th century (1900-1999)

- Vaccinations
- Control of infectious diseases (clean water, improved sanitation)
- Decline in deaths from coronary heart disease and stroke
- Safer and healthier foods
- Healthier mothers and babies
- Awareness of mental health disorders
- Fluoridation of drinking water
- Recognition of tobacco use as a health hazard

Source: https://www.cdc.gov/mmwr/preview/mmwrhtml/00056796.htm

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Reporting is a long and active tradition of public health

- 1850 – first US federal mortality statistics reported
- 1874 – Massachusetts becomes first state to initiate voluntary reporting
- 1893 – Michigan becomes first state to institute mandatory reporting
- 1961 – CDC becomes responsible for collection and dissemination of disease reports
- 2001 – Anthrax attacks increase interest and funding in disease reporting and surveillance
- 2020 – COVID-19 case reporting for health departments
https://www.cdc.gov/coronavirus/2019-ncov/php/reporting-pui.html

From Hersh et al., 2022

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

What is typically reported

- Foodborne or waterborne diseases – e.g., Cholera, E. coli, Salmonella, etc.
- Sexually transmitted infections – e.g., Chlamydia, Syphilis, HIV/AIDS
- “Traditional” infectious diseases – e.g., tuberculosis (TB), meningitis
- “Exotic” diseases – e.g., SARS, Creutzfeld-Jakob, etc.
- Environmental diseases – e.g., lead poisoning, pesticide exposures, etc.
- Maternal and child health – e.g., infant mortality, birth defects, etc

From Hersh et al., 2022

New opportunities for public health informatics

- Merging of EHR, public health, and other sources of data
- Infectious disease surveillance
- Community health record (CHR) combining EHR, public health, social services, and other data
- Digital Bridge to advance electronic case reporting from EHR and other data – https://digitalbridge.us/
- Adding location-based data to clinical data warehouse
- Use of artificial intelligence in historically resource-poor settings

From Hersh et al., 2022

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

AI in public health informatics

- **Assessment functions i.e. public health surveillance**: Data collection and analysis to monitor population health status.
  - Machine learning (ML) methods can be applied to data collected from wearable devices, social media, electronic health records to predict population health.
  - NLP methods can be used to extract features from the unstructured text data
- **Policy development:** Communication, advocacy and coordination with stakeholders.
  - Knowledge-based systems have been used to support these functions
- **Assurance functions:** Public health training, research and evaluation of public health services.
  - Public health research involves the use of ML, NLP and knowledge-based methods

From Hersh et al., 2022

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Disease Surveillance

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Role of Surveillance Data

- Serve short-term needs (e.g., to respond to an acute infectious disease outbreak or pandemic such as COVID-19) or
- Serve longer-term needs (e.g., to determine leading causes of premature death, injury, or disability)
- Available for querying and visualization through state and federal public health web sites (e.g., data.gov)
- Used by epidemiologists and researchers
  - To impact public understanding of health threads
  - To define priorities of public health priorities actions
  - To guide the development of new policies

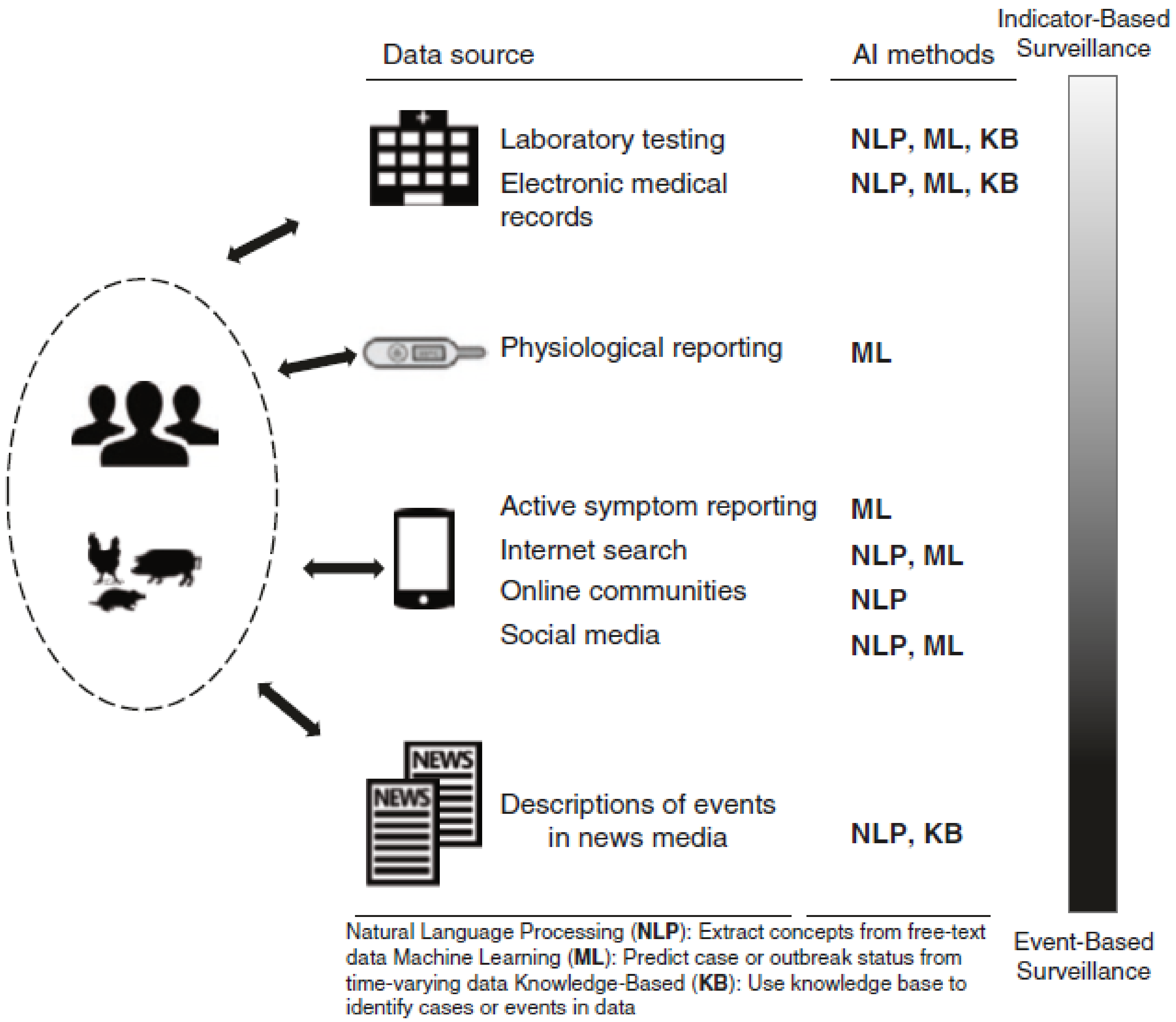
This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Examples of Surveillance Data

- Health conditions (i.e. breast cancer, obesity)
- Threads to health (i.e. smoking prevalence, drug overdose)
- Healthcare capacity (e.g., availability of immunization or medications)
- Emergency or intensive care services
- Other critical needs for delivering required care for a population
- Other events (e.g., births, mortality)

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Applications of AI Methods to Infectious Disease Surveillance



Data such as laboratory results and electronic medical records allow measurement of cases of disease (i.e., indicator-based surveillance), while other sources, such as news media allow detection of events, without measuring individual cases (i.e., event-based surveillance). Regardless of the data source, many approaches to surveillance rely on artificial intelligence methods to extract concepts from free-text, predict cases or outbreaks from time-varying data, or to reason about data using existing knowledge. Source: Shortliffe et al., 2022

Internet-based Surveillance Data for Public Health

- Data from search queries and location logs e.g. Google Trends, Wikipedia page views, Google Flu Trends or even online reservations.
- Social media posts e.g. Facebook posts, Twitter
  - NLP, image analysis used to extract features from users' posts
- Crowdsourcing surveillance efforts (e.g. FluNearYou, Influenzanet)
- Digital health collection platforms based on smartphones and wearable devices.

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

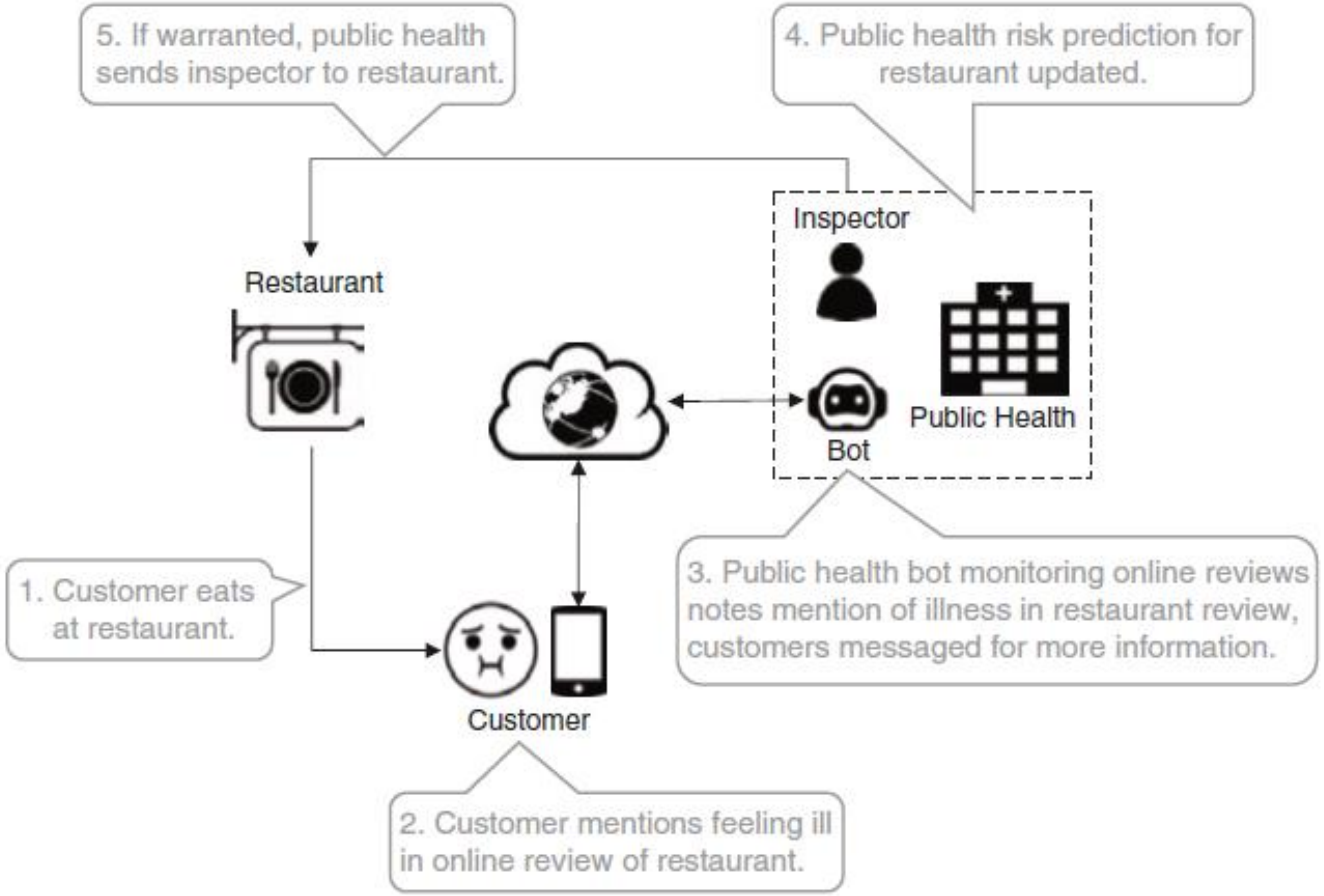
Surveillance Data from Social Posts

- Availability of **geolocated tweets** to monitor spread of a disease
- Twitter is the most popular for collecting surveillance data
- Facebook posts or likes, Instagram timelines such as to identify drug reactions [Correia et al. 2016].
- Application of AI in social medial platforms to prevent foodborne illness due to restaurant dining (see next slide).

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Surveillance Data from Social Posts (Example)

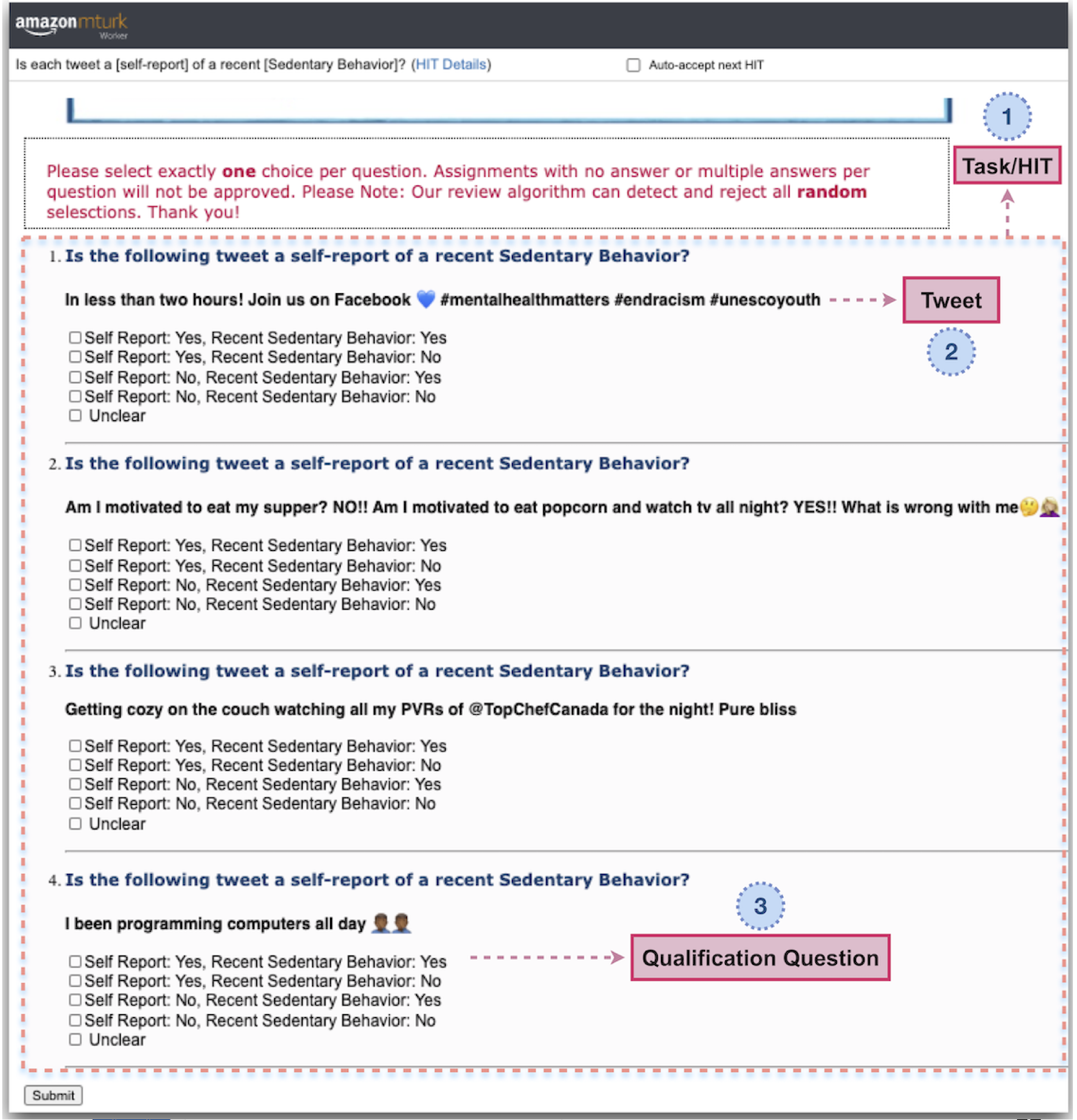


An example of how AI methods can be used to monitor comments on social media and generate information to guide public health interventions. Here, a bot developed by a public health organization detects a comment that may indicate foodborne illness due to a restaurant meal. This information is used to update the risk assessment of the restaurant and possibly trigger an in-person inspection. Source: Shortliffe, 2022 (More details in: Oldroyd, 2018)

Surveillance Data from Crowdsourcing

- Crowd-sourced participatory surveillance systems - Flu Near You [Smolinski et al., 2015] and Influenzanet.
- Crowdsourcing system are online platforms that recruit users to undertake a particular task i.e. an online survey including detailed symptom reports and report on observed disease distribution through online maps and newsletters [Woójcik et al. 2014].

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423



Crowdsourcing online survey conducted in Amazon Truck for data annotation.

A sample labeling task (ie, human intelligence task [HIT]) for sedentary behavior. Each HIT contains 4 questions and each asks if the presented tweet is a self-reported physical activity, sedentary behavior, or sleep quality–related behavior. The fourth question is an easy, qualification question that was used to check the quality of the worker. Source: Shakeri Hossein et al. 2022.

Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Limitations of Surveillance Data

- Digital data are not owned by the public
- Digital health surveillance data is often linked to a non-health data source (e.g. pulling health data from a twitter user who posts on a range of topics which may also include mentions of their health-related information).
- Except from crowdsourcing data, the rest are not captured actual occurrence of the disease but public awareness.
- Inaccurate predictions of Google Trends
- Frequent changes to search query algorithms

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Privacy and Security

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Societal and ethical risks on health informatics applications

- Patients' **data privacy**
- Patients' **confidentiality**
- Legal risks impact the **accountability** of the data
- **Transparency of the system** – the individuals should be aware of any data collection or use of their personal health data.
- **Discrimination** in medical data (i.e. train a model with less data of a representative population for development of a new treatment/vaccine/drug or diagnosis)

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Terminology

- **Privacy:** Concerns who has access to the patient information and how do we declare who has unauthorized access.
- **Confidentiality:** Data or information is not made available or disclosed to unauthorized persons or processes.
- **Security:** Protecting information and information systems from unauthorized access, use, disclosure, disruption, modification, or destruction in order to provide confidentiality, integrity, and availability.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Terminology (ctd.)

- **Accountability:** The requirement for actions of an entity to be traced uniquely to that entity.
- **Availability:** Data or information is accessible and useable on demand by an authorized person.
- **Integrity:** Data or information have not been altered or destroyed in an unauthorized manner.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Terminology (ctd.)

- **Individually identifiable health information (IIHI)** – Any data that can be correlated with an individual.
- **Protected health information (PHI)** – IIHI as defined by HIPAA Privacy Rule
- **Consent** – (in context of privacy) written or verbal permission to allow use of your IIHI.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Data Privacy Risks

- Large amounts of data are being collected on multiple devices i.e. mobile devices, wearable devices.
- Exchange of data/Information sharing
- Health-care data may contain **highly sensitive information** (e.g., sexually transmitted diseases, behavioral health information, or financial information).
  - Sensitive information should be protected from access and misuse by unauthorized persons.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Medical Identity Theft

- A growing concern, emanating from general identity theft, defined as use of IIHI for obtaining access to property or services
- https://oig.hhs.gov/fraud/medical-id-theft/
- Victims are not only individuals but also healthcare providers and plans as well as society at large
- Electronic health records (EHRs) deemed extremely valuable, e.g., 100 times more than credit card numbers

From Herch et al. 2022

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Privacy Risks and Confidentiality in Public Health

- Individuals and health care providers will not use the systems if they are not sure that their data would be kept **private** and **confidential**.
- Data collected for public health research could also benefit the patients but they should be handled with confidentiality.
- The ability of health care researchers to **anonymously pool** data from patient cases that limits the natural history of the disease and the effects of various treatments.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

Privacy Risks and Confidentiality in Public Health

- **Challenge:** to find a balance between free access to information (useful for personal and public health informatics systems) and the protection of patients' privacy and confidentiality.
  - ➢ the ethical obligations to privacy and confidentiality against the social goals of public health

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

New Challenges with the Technology Advancements

- Growing use of electronic data in clinical workflows
- Health information exchange (HIE) moves data across networks
- Cloud computing changes perimeter of data protection
- New models of care (e.g., accountable care organizations) require more members of team to access information
- Clinicians want to bring their own devices (e.g., personal laptops, smartphones, etc.)

From Herch et al. 2022

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

De-identified Data

- Growing numbers of repositories of de-identified data used for research purposes.
- Structured data easier to de-identify, though methods develop for text (Li, 2017; Heider, 2020)
- Various methods can be used to ensure population databases do not allow identified of individuals, e.g., when small cell numbers (O'Keefe, 2018)
- Machine learning methods achieve best results, although still not perfect (Hartman, 2020)
- **Challenge:** De-identification of data may also eliminate the access to patient's data that are important for a medical task.

From Herch et al. 2022

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Technological Approaches

- Deterrents (Awareness)
  - Alerts
  - Audit trails
- System management precautions
  - Software management
  - Analysis of vulnerability
- Obstacles
  - Authentication
  - Authorization
  - Integrity management
  - Digital signatures
  - Encryption
  - Firewalls
  - Rights management

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Encryption

- Necessary but not sufficient to ensure security
- Should, however, be used for all communications over public networks, e.g., the Internet
- Information is collected using a key
- **Types**: symmetric vs. asymmetric
  - Asymmetric, aka public key encryption, can be used for digital certificates, electronic signatures, etc.

From Herch et al. 2022

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Authentication and passwords

- Authentication is a process of gaining access to secure computer
  - Most commonly done with user passwords
- Typical user interacts with many professional and personal sites for which he/she must use password
  - Many desire “single sign-in,” especially in healthcare, where users authenticate just once

From Herch et al. 2022

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

Authentication and passwords

- Growing use of two-factor authentication, i.e., in addition to passwords ("what you know"), may add physical entities ("what you have"), e.g.,
  - Biometric devices – physical characteristic, such as thumbprint or face (e.g., iPhone), retina (e.g., NCIS)
  - Physical devices – smart card, smartphone, or some other physical "key"

From Herch et al. 2022

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

HIPAA privacy and security

- Health Insurance Portability and Accountability Act (HIPAA)
- Privacy Rule: https://www.hhs.gov/hipaa/for-professionals/privacy/laws-regulations/
- Security Rule: https://www.hhs.gov/hipaa/for-professionals/security/laws-regulations/
- Enforced by HHS Office for Civil Rights (OCR): https://www.hhs.gov/ocr/
- Both rules extended with HITECH legislation in 2009
- Patient rights
- Get data, check data, see who has seen it: https://www.hhs.gov/hipaa/for-individuals/guidance-materials-for-consumers/

From Herch et al. 2022

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

HIPAA not only US law or regulation governing privacy

- Common Rule
- Governs rights of subjects participating in human research, updated in 2017- https://www.hhs.gov/ohrp/regulations-and-policy/regulations/common-rule/
- Followed by many federal agencies – HHS rule governing human subjects research is 45 CFR 46 - https://www.hhs.gov/ohrp/regulations-and-policy/regulations/45-cfr-46/
- Family Educational Rights and Privacy Act (FERPA)
- Prevents disclosure of personally identifiable information without approval of parents or (if over 18) student

From Herch et al. 2022

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Bias and Discrimination

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Trust in Clinical Decision Support Systems (CDSS)

- Patients are often frightened and vulnerable
- Build trust between physicians and patients
- Treat data with confidentiality
- Enhance the communication between physicians and patients
- Transparency of the system and the process of collecting/processing data

- “Many clinical decisions are not exclusively medical—they have social, personal, ethical, psychological, financial, familial, legal, and other components; even art might play a role. (Miller and Goodman 1998)”

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

Trust in Clinical Decision Support Systems (CDSS)

- Computer software developers should strive to warn caregivers whenever it appears that a mistake is about to happen.
- Practitioners who use informatics tools should be clinically qualified and adequately trained in using the software products.
- The tools themselves should be carefully evaluated and validated, in vitro and in vivo.
- Health informatics tools and applications should be evaluated not only in terms of performance but also in terms of their influences on individuals or particular communities.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

Bias and Discrimination in Biomedical Informatics

- Bias in medical field divided into three areas:
  - Data bias
  - Model bias
  - Human bias
- Statistical bias vs Social bias
- Statistical bias can cause an algorithm to produce an output that differs from the true estimate [Parikh et al., 2019].

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Bias and Discrimination in Biomedical Informatics

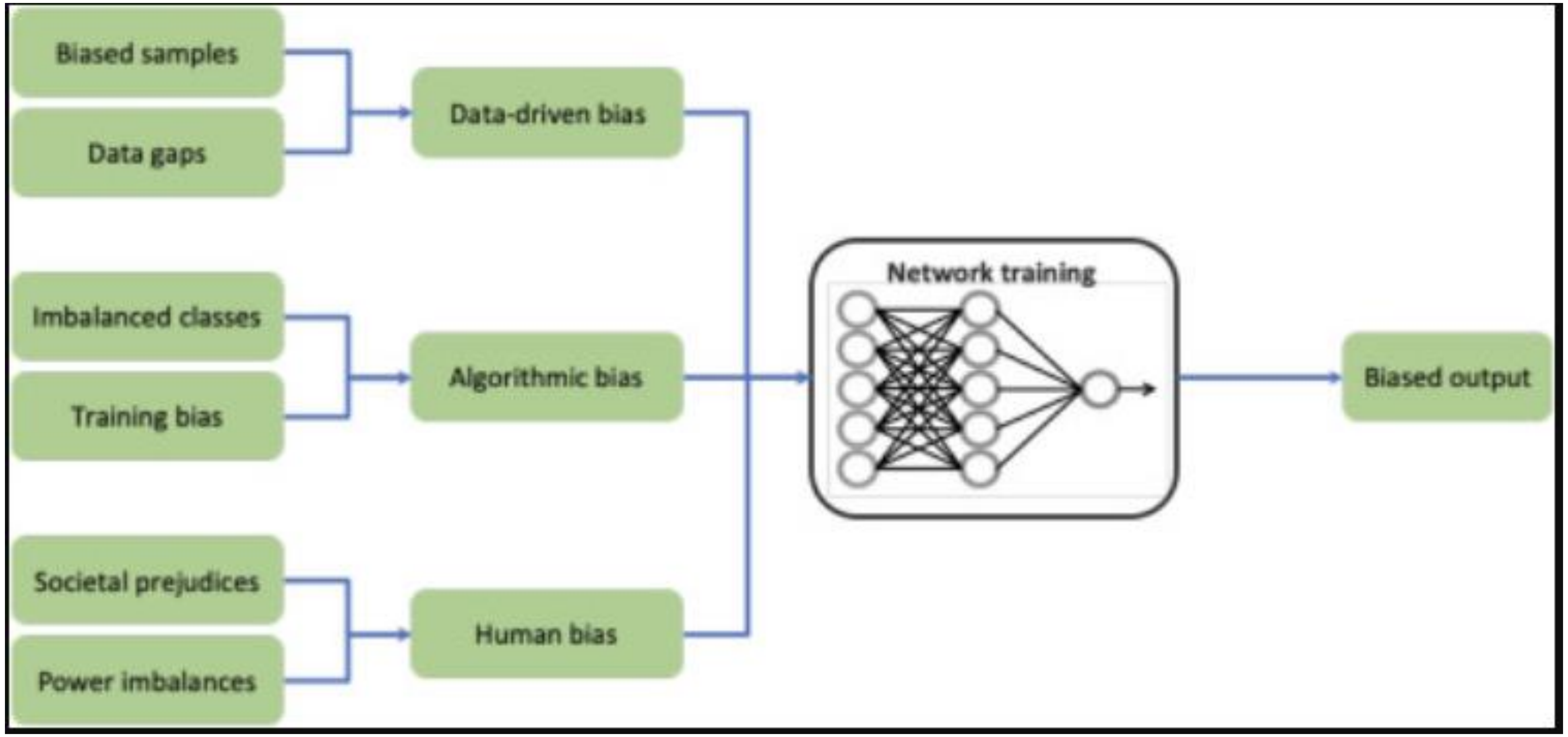
- Social bias refers to inequalities that may result in suboptimal outcomes for given groups of the human population i.e. demographic bias[Davenport et al. 2019].
- Misdiagnosis of certain groups due to underrepresented cases of gender or ethnic minorities

Auditing approaches can help to detect bias and promote fairness

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Fairness and Discrimination in Biomedical Informatics



Source: Norori et al., Patterns, 2021 - Illustration of different sources of bias in training machine learning algorithms

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Racial Bias in ML in Medical Applications

- Example of data (social) bias
- Convolutional neural networks (CNN) has been applied in **skin lesion classification** with high accuracy but the proportion of the Black patients in the training sample is approximately 5% to 10%.
  - ➢ Black patients are less likely to be diagnosed with melanoma even if their mortality rate for melanoma is high.
  - ➢ Misdiagnosis may cause skin cancer at a more advanced stage in Black patients.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Racial Bias in ML in Medical Applications (ctd)

- **Model bias**
- Racial bias has been identified in a widely used algorithm for predicting patients with complex health-care needs.
  - Algorithm uses health costs (feature) to make decisions for health needs.
  - Since less money is spent on Black patients who have the same level of need, the algorithm falsely concludes that Black patients are healthier than equally sick White patients [Obermeyer et al., 2019].
  - Data bias reflected also on the implementation of the algorithm.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

Gender Bias in ML in Medical Applications

- In cardiology, a heart attack is overwhelmingly misdiagnosed in women
  - ➢ Prediction models for cardiovascular disease are mostly trained in men datasets
- Gender bias in data is more often in clinical trials and drug development
  - ➢ In clinical trials, most participants are male of a limited age group
  - ➢ Preclinical studies for drug development usually include either a vast majority, or exclusively male animals.
- Bias due to socioeconomic status or sexual orientation are often impossible to infer in a biomedical dataset unless this information has been explicitly collected and included as metadata.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Human Bias

- Inequalities of patient diagnosis/treatment due to bias in health care practitioners
- Patients suffering from depression, with ethnic minorities experiencing more severe symptoms and receiving medication less often than white patients.
- Medication dosages were adjusted for patient size, without considering gender or racial differences.
- In health care, gender differences can include differences in gene expression
- Age, gender differences can include differences in the prevalence of particular diseases such as coronary heart disease, stroke, and different types of cancer.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Transparency and Fairness

- For an AI system to be fair – avoid discrimination and bias -, transparency is important.
- Transparent systems can easily be assessed for fairness.
- Documentation of every step of the development process.
- Intellectual property of how someone can reuse the code.
- Humans should be aware that they interact with an AI i.e. suicide prevention system.
- A factor that may lead to biased results may also be associated to a differential impact on certain outcomes.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Future of AI in Public Health

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Limitations of Public Health

- Limitations in training, resources, and data access in some public health settings.
- Minimal reporting guidelines have been developed for clinical applications (and public health) of AI.
- Awareness and knowledge of AI methods is limited in public health.
- Training in public health informatics and the application of AI in public health is critical.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Future Applications of AI in Public Health

- ML and other AI methods could be used to support decisions about the effective use of interventions in specific communities.
  - Public health organizations should systematically track the delivery of interventions.
- Improved integration of individual-level and public health data.
- Coordinate clinical and public health interventions allowing a “patient-centered” approach
- Digital transformation of public health systems using AI methods.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

SUMMARY

- Public health is the science concerns the monitoring of the population health and disease prevention along with laws and policies development regarding the health of the population i.e. use of vaccinations.
- AI methods such as machine learning, knowledge-based systems and NLP have many applications in public health.
- Surveillance data can be collected from social media posts, web searches, location logs, wearable devices along with other online data such as online reservations but even from crowdsourcing.
- Public health researchers associate the environment and other socio-economic factors with the population health.
- Privacy, security, transparency and confidentiality are important when processing health data.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

SUMMARY

- Anonymization and encryption are just few examples of technological approaches for maintaining privacy in health data.
- Bias and discrimination can be detected in any type of health data such as gender or racial bias which might affect the developed AI clinical system and its output for a particular group of individuals.
- Bias i.e. social, statistical or human bias can be detected by applying auditing in the data by third parties.
- For future applications of AI in public health, training in public health informatics and in AI methods are critical.
- One of the future achievements of AI in medicine is the improved integration of individual-level and public health data.

Discussion

- Who are stakeholders and why are they important for ethical AI development?

- What is the most important barrier to the application of AI in public health?

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

References

- Oldroyd RA, Morris MA, Birkin M. Identifying Methods for Monitoring Foodborne Illness: Review of Existing Public Health Surveillance Techniques. JMIR Public Health Surveill. 2018;4(2):e57.
- Norori N., Qiyang Hu, Florence Marcelle Aellen, Francesca Dalia Faraci, Athina Tzovara, Addressing bias in big data and AI for health care: A call for open science,Patterns,Volume 2, Issue 10, 2021,100347,ISSN 2666-3899, https://doi.org/10.1016/j.patter.2021.100347.
- Centers for Disease Prevention and Control. (1999). Ten great public health achievements – United States, 1900–1999. MMWR; 48(12);241–243.
- Smolinski MS, Crawley AW, Baltrusaitis K, Chunara R, Olsen JM, et al. 2015. Flu Near You: Crowdsourced Symptom Reporting Spanning 2 Influenza Seasons. Am. J. Public Health 105:2124–30
- Ginsberg J, Mohebbi MH, Patel RS, Brammer L, Smolinski MS, Brilliant L. 2009. Detecting influenza epidemics using search engine query data. Nature 457:1012–4
- Koppeschaar CE, Colizza V, Guerrisi C, Turbelin C, Duggan J, et al. 2017. Influenzanet: Citizens Among 10 Countries Collaborating to Monitor Influenza in Europe. JMIR Public Health Surveill 3:e66.

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

References

- Woójcik OP, Brownstein JS, Chunara R, Johansson MA. 2014. Public health for the people: participatory infectious disease surveillance in the digital age. Emerg. Themes Epidemiol 11:7.
- Francisco Nauber Bernardo Gois, Alex Lima, Khennedy Santos, Ramses Oliveira, Valdir Santiago, Saulo Melo, Rafael Costa et al. Predictive models to the COVID-19,Editor(s): Utku Kose, Deepak Gupta, Victor Hugo C. de Albuquerque, Ashish Khanna, Data Science for COVID-19, Academic Press,2021,Pages 1-24, ISBN 9780128245361,https://doi.org/10.1016/B978-0-12-824536-1.00023-X
- Shortliffe, Edward H., and James J. Cimino, eds. Biomedical informatics: Computer applications in health care and biomedicine. Springer Science & Business Media, 2021.
- Shakeri Hossein Abad Z, Butler GP, Thompson W, Lee J, Crowdsourcing for Machine Learning in Public Health Surveillance: Lessons Learned From Amazon Mechanical Turk J Med Internet Res 2022;24(1):e28749
- Elaziz MA, Hosny KM, Salah A, Darwish MM, Lu S, Sahlol AT (2020) New machine learning method for image-based diagnosis of COVID-19. PLoS ONE 15(6): e0235187. https://doi.org/10.1371/journal.pone.0235187

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423