

Medical Data Privacy and Ethics in the Age of Artificial Intelligence

Lecture 1: Introduction and Overview

Zhiyu Wan, PhD (wanzhy@shanghaitech.edu.cn)

Assistant Professor of Biomedical Engineering

ShanghaiTech University

September 16, 2026

Oncology firm Novocure says cyberattack exposed US patient records

By Reuters
September 1, 2026 7:54 PM GMT+8 · Updated September 1, 2026



Figurines with computers and smartphones are seen in front of the words "Cyber Attack" in this illustration taken, February 19, 2024. REUTERS/Dado Ruvic/Illustration [Purchase Licensing Rights](#)

Companies

Abbott Laboratories

Follow

Boston Scientific Corp

Follow

Medtronic PLC

Follow

Show more companies

Sept 1 (Reuters) - Oncology company Novocure ([NVCR.O](#)) said on Tuesday a cybersecurity incident involving unauthorized access to certain information systems in mid-August exposed internal records of more than 1,400 U.S. patients.

The breach adds to a growing number of cyberattacks on the healthcare sector, including medical device makers Abbott ([ABT.N](#)), Stryker ([SYK.N](#)), Medtronic ([MDT.N](#)) and Boston Scientific ([BSX.N](#)), drugmaker Novo Nordisk ([NOVOB.CO](#)) and drug-delivery equipment supplier West Pharmaceutical Services ([WST.N](#)).



The HIPAA Journal is the regulatory update

Become HIPAA Compliant	HIPAA Checklist	HIPAA News	HIPAA Breach News	Cybersecurity News	HIPAA Legal
Compliance Pros Hub	Small Practices Hub	Business Associate Hub	HIPAA Training Hub	HIPAA Training Courses	

Atrium Health Pays Up to \$1.8M to Resolve Pixel Lawsuit

Posted By [Steve Alder](#) on Jul 17, 2026

Charlotte-Mecklenburg Hospital Authority, doing business as Atrium Health, has agreed to pay up to \$1,800,000 to settle a class action lawsuit stemming from its use of pixels and other tracking technologies on its MyAtriumHealth (formerly called MyCarolinas) patient portal.

North Carolina-based Atrium Health operates a dozen hospitals in North and South Carolina, along with more than 900 care facilities in the two states. Like many health systems, Atrium Health used [tracking technologies on its patient portal](#). These tools have important uses for website operators; however, their use on healthcare websites risks impermissible disclosures of sensitive data.

When these tools are added to authenticated web pages such as patient portals, patients' protected health information may be disclosed to the third-party providers of the tools, such as Meta (Facebook) and Google. Following an investigation, Atrium Health determined that between January 1, 2015, and July 31, 2019, the protected health information of up to 585,959 patients may have been impermissibly disclosed to third parties as a result of the use of these tools. When reporting the data breach, Atrium Health assumed that all patients who used the portal had their ePHI impermissibly disclosed.

The data potentially compromised included IP addresses and third-party identifiers/cookies. If forms were filled out, that disclosed information may also have included full names, email addresses, phone numbers, city/state/zip code, gender, and any other information entered into the forms.



ChatGPT's advice kept man from seeking medical treatment for dangerous condition, lawsuit claims

By Diana Novak Jones
July 23, 2026 3:16 AM GMT+8 · Updated July 23, 2026



OpenAI CEO Sam Altman attends an event in Tokyo, Japan February 3, 2025. REUTERS/Kim Kyung-Hoon/File Photo [Purchase Licensing Rights](#)

Summary

Companies

- The lawsuit seeks damages and a court order requiring ChatGPT to end conversations with people needing immediate medical help
- OpenAI has said 40 million users ask ChatGPT healthcare-related questions daily
- The company is facing a growing number of lawsuits alleging real-world harms

July 22 (Reuters) - A Florida pastor sued OpenAI and its chief executive, Sam Altman, on Wednesday claiming ChatGPT's inaccurate medical advice kept him from seeking treatment for recurrent dizzy spells, leading him to suffer a pulmonary embolism that nearly killed him.

中国官方发布人工智能医学影像研究伦理指引 强调保护隐私和安全

2026年08月28日 18:04 来源：中国新闻网

A+ 大字 体 A- 小字 体

中新网北京8月28日电 (记者 孙自法)作为人工智能(AI)赋能医学研究和临床实践的重要组成部分，人工智能医学影像正在逐渐走向应用，广受关注。

中国官方最新发布《人工智能医学影像研究伦理指引》，明确提出增进人类福祉、促进公平公正、尊重人的主体性、保护隐私和数据安全、确保安全可控、强化透明可信等基本原则。

其中，人工智能医学影像研究应切实保护个人隐私、强化数据安全。对医学影像及健康数据的采集、传输、存储、分析、共享、销毁，应实施全流程规范管理，建立完善的数据授权与撤回机制。尤其是大规模数据存储应采取高等级的安全保障，防止隐私泄露、外部攻击和未经授权使用。

人工智能医学影像研究应遵循医学科学规律，建立全生命周期风险治理机制，保障技术研发、实验与临床应用的安全、规范、有效、可控。健全风险事前预防、事中管控、事后处置的全流程闭环管理，守住医疗安全与科技伦理底线。

《人工智能医学影像研究伦理指引》由国家科技伦理委员会医学伦理分委员会研究制定，科技部官网日前予以全文公布。其旨在指导人工智能医学影像研究的合规开展，规范人工智能医学影像相关的基础算法研究、技术研发、临床验证和转化，防范科技伦理风险，推动人工智能医学影像研究的负责任创新与健康有序发展。

人工智能医学影像是指利用人工智能技术对医学影像数据进行分析 and 处理的一种技术系统与amp;方法，主要用于辅助医生开展疾病诊断、治疗决策、健康管理以及为医患共同决策提供信息支持，提高诊断的准确性和效率。

人工智能医学影像数据是指与人工智能医学影像研究相关的各类数据，包括原始影像数据、训练数据、预训练数据、合成数据、验证与测试数据。

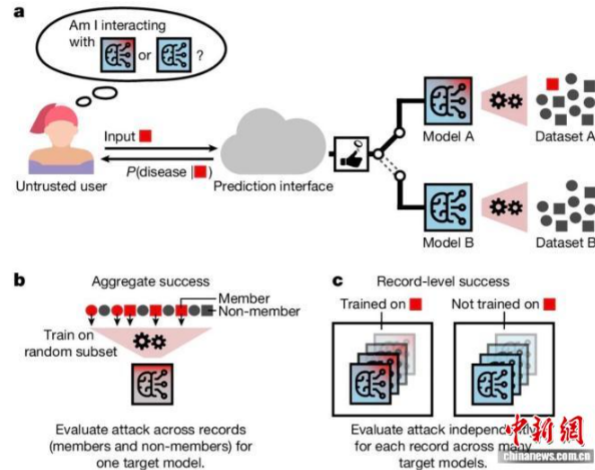
该指引对脱敏处理、算法泛化能力、算法黑箱、人工智能幻觉、算法可解释性、特异性、灵敏性等术语进行解读。人工智能幻觉是指人工智能系统(含多模态医学影像模型、大语言模型)输出看似可信、形式完整，但与客观事实、原始数据源、输入条件不符或完全虚构内容的现象。

聚焦医疗AI模型带来隐私风险 最新健康研究呼吁加强评估和保护

2026年06月25日 12:59 来源：中国新闻网

A+ 大字 体 A- 小字 体

中新网北京6月25日电 (记者 孙自法)国际学术期刊《自然》最新发表一篇健康科学的研究论文提醒，数据被用于训练医疗人工智能(AI)模型的个人，可能面临在网络攻击中被识别的风险。



本项研究的相关示意图(图片来自论文)。施普林格 自然 供图

这项研究表明，代表性不足群体面临的数据泄露风险可能更高。论文作者指出，当前的风险评估并未将这些群体纳入考量，因此，他们呼吁采取进一步的风险缓解措施并实施严格的访问控制。

该论文介绍，医疗AI模型有望改善全球健康状况，特别是在缺乏专业人才地区。然而，用于训练这些模型的敏感数据可能面临隐私攻击。攻击者利用成员推理攻击(MIA)来确定个人的数据是否被用于训练模型。通过此类攻击，可以推断出患者的医疗数据和私人信息。此前关于数据风险的研究主要基于整个数据集，并未考虑个体的风险。

论文第一作者和通讯作者、德国慕尼黑工业大学Moritz A. Knolle与同事及合作者开展了一项隐私审计，重点关注个人隐私风险，发现医疗AI模型可能对个人数据贡献者构成隐私风险。他们利用七个由真实临床数据(包括医学影像、心电图和电子健康记录)组成的大型数据集，确定了数据贡献者中最为脆弱的群体。



莫里茨·诺勒

慕尼黑工业大学，医疗保健与医学人工智能研究所
tum.de 的已验证邮箱

可信赖的人工智能 保护隐私的人工智能 医疗人工智能

标题	被引用	年
利用动态红队演练解决大型语言模型在健康和医学领域的基准测试差距 J Pan, B Jian, P Hager, Y Zhang, C Liu, F Jungmann, HB Li, J Canisius, ... 自然健康, 1-17	1	2026
医疗人工智能带来的隐私风险各不相同 MA Knolle, MJ Menten, F Jungmann, F Meissen, B Glocker, D Rueckert... 自然	6	2026
敏感性、特异性和一致性：合成数据生成隐私过滤器的三方评估 A Koeken, A Ziller, M Knolle, D Rueckert arXiv 预印本 arXiv:2510.01793	1	2025
记忆偏差：人工智能对数据贡献者的预测结果会偏向于他们在训练数据中的健康状态 M Knolle, MJ Menten, D Rueckert, G Kaissis, B Glocker 利用纵向医学影像和数据进行学习的国际研讨会.....		2025
胰腺导管腺癌分支类器官模型中异质性驱动的表面可塑性和治疗反应 A Papargyriou, M Najajreh, DP Cook, CH Maurer, S Barthel, HA Messal, ... 自然生物医学工程 9 (6), 836-864	51	2025
差分隐私 M 诺尔, G 凯西斯 联邦学习在医学影像中的应用, 83-94页		2025
利用扩散模型进行视觉隐私审计 K Schwethelm, J Kaiser, M Knolle, S Lockfisch, D Rueckert, A Ziller arXiv 预印本 arXiv:2403.07588	3	2024
无监督异常检测中(可预测的)性能偏差 F Meissen, S Breuer, M Knolle, A Buyx, R Müller, G Kaissis, B Wiestler, ... EBioMedicine 101	6	2024
机器学习中的记忆：结果综述 D Usynin, M Knolle, G Kaissis 机器学习研究学报, 1-25	13	2024
偏差感知最小化：理解和缓解私有随机梯度下降中的估计偏差 M Knolle, R Dorfman, A Ziller, D Rueckert, G Kaissis arXiv 预印本 arXiv:2308.12018	3	2023
一种独特的刺激性 cDC1 亚群可增强肿瘤中 CD8+ T 细胞的反应，从而发挥保护性抗癌免疫作用。 P Meiser, MA Knolle, A Hirschberger, GP de Almeida, F Bayerl, S Lacher... Cancer Cell 41 (8), 1498-1515. e10	239	2023
肿瘤来源的前列腺素E2会诱导cDC1功能障碍，从而损害肿瘤内抗癌T细胞反应的调控。 F Bayerl, P Meiser, S Donakonda, A Hirschberger, SB Lacher, AM Pedde... 免疫学 56 (6), 1341-1358. e11	240	2023
人工智能和大数据在心脏病学中的应用：实用指南（诊断章节） D Rueckert, M Knolle, N Duchateau, R Razavi, G Kaissis 施普林格 自然	3 *	2023
输入属性如何影响差分隐私中的隐私损失？ TT Mueller, S Kolek, F Jungmann, A Ziller, D Usynin, M Knolle... arXiv 预印本 arXiv:2211.10173 D Rueckert, M Knolle , N Duchateau, R Razavi, G Kaissis Springer Nature	2	2022
How Do Input Attributes Impact the Privacy Loss in Differential Privacy? TT Mueller, S Kolek, F Jungmann, A Ziller, D Usynin, M Knolle ... arXiv preprint arXiv:2211.10173	2	2022



Welcome to Medical Data Privacy and Ethics in the Age of Artificial Intelligence

- You're sitting in BME2142
- When: Wednesdays, 15:00-15:45, 15:55-16:40
- Where: SIST, D106
- Office Hours: Upon Request
- Teaching Assistant:
 - Jiarui Xu (xujr2025@shanghaitech.edu.cn)
 - Haoyun Xu (xuhy12025@shanghaitech.edu.cn)
- Instructor: Zhiyu Wan (wanzhy@shanghaitech.edu.cn)

Goals of This Lecture

- Know more about this course
 - Instructor
 - Objectives
 - Grading policies
 - Schedule
- Overview of the concepts
 - AI
 - Medical Data
 - Medical AI
 - Privacy
 - Medical Data Privacy
 - Ethics



AI



Medical Data



Medical AI



Privacy



Medical Data Privacy



Ethics

■ Hi! I'm Zhiyu Wan.

- BEng, **Automation**. XJTU (Gifted Young)
- MS & PhD, **Computer Science**, Vanderbilt
- Postdoc, **Biomedical Informatics**, VUMC

■ Faculty in **Biomedical Engineering**

■ Health Information Safety and Intelligence Research Lab

(<https://zhiyuwan.com/hisir-lab/>)

■ Research Areas

- Biomedical Data Privacy
- AI in Medicine
- Trustworthy AI
- AI Ethics
- Computational Game Theory
- ...



BME2142: Medical Data Privacy and Ethics in the Age of Artificial Intelligence

Schedule

Projects

Public Datasets

Who / When / Where
Instructor: [Zhiyu Wan](#)
Teaching Assistant: Jiarui Xu, Haoyun Xu
Semester: Fall 2026
Time: Wednesdays, 15:00-16:40
Location: School of Information Science and Technology, Room 1D106
Office Hours: Upon request, Location: BME Building, Room 228

Course Syllabus ([PDF](#))

First Day of Class: September 16, 2026

Description

“ Medical Data Privacy and Ethics in the Age of Artificial Intelligence ” is a specialized elective course for graduate students majoring in Biomedical Engineering and a foundational public course for Master of Engineering students in Biomedical Engineering. This course focuses on issues related to medical data privacy and ethics in the age of artificial intelligence (AI), with an in-depth exploration of privacy protection technologies, ethical dilemmas, and legal regulations surrounding medical data. The curriculum covers ethical challenges and privacy protection strategies that may arise during the collection, sharing, processing, and use of medical data. Through lectures, discussions, case studies, and project-based practice, this course aims to develop students’ capabilities in designing intelligent medical systems and managing biomedical data. It helps students understand how to balance privacy protection with technological innovation in AI-driven healthcare data systems, and provides a solid foundation in ethics, law, and technical practices for their future work in medical data processing.

The course comprises four fundamental modules: the first part introduces basic concepts and major challenges of data privacy and ethics in the age of artificial intelligence; the second part discusses social and legal approaches to protecting data privacy and ethics; the third part covers technical methods for privacy and ethics protection; and the fourth part applies these privacy and ethics protection methods to specific cases.

Schedule

(Note: This schedule is tentative and subject to change.)

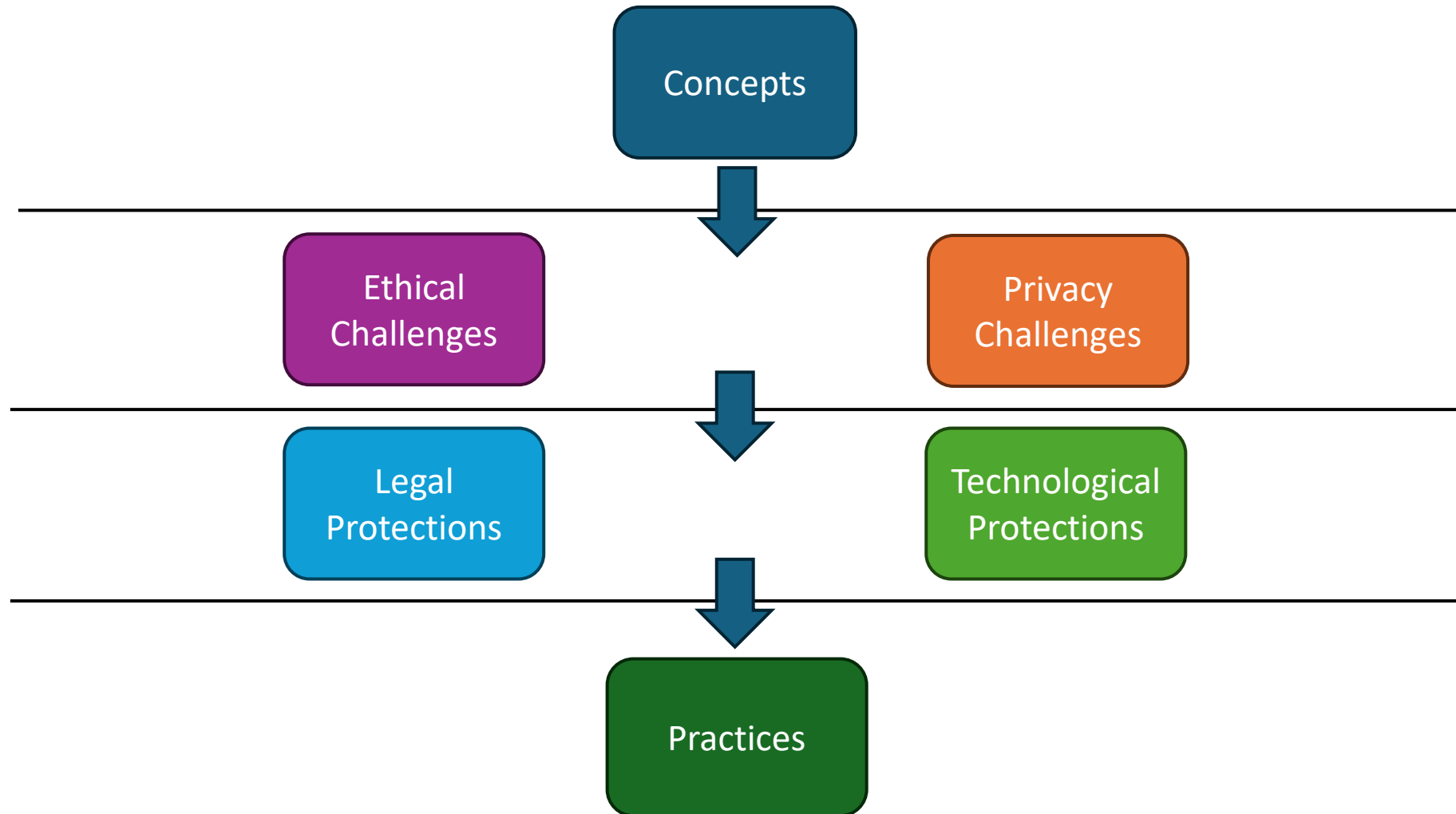
Week	Chapter	Teaching Contents	Readings	Assignments
Week 1, Wednesday Sept 16, 15:00-16:40 Lecture 1 (PDF)	I. Course Introduction and Overview of Data Privacy and Ethics	Course introduction. The role of medical data in the age of AI. Concept of data privacy and its importance. Concept of ethics and morality and their importance. AI ethics: machine morality and ethics, automation, and employment.	(Optional: 1. 《信息科学技术伦理与道德》 Chs.8-10. 2. 《工程伦理》 Ch.12. 3. 《信息科学技术伦理与道德》 Chs.3&4.)	/
Week 2, Wednesday Sept 24, 15:00-16:40 Lecture 2 (PDF)	II. Ethical Issues in Life Sciences, Medicine, and Informatics (a)	Research ethics: ethical guidelines for human and animal experiments and scientific research. Life sciences ethics: controversies arising from reproductive technology, genetic technology, stem cell research, etc. Information security and privacy issues: personal data breaches, surveillance technology, facial	(Optional: 1. 《工程伦理》 Ch.10. 2. 《信息科学技术伦理与道德》 Chs.5&7.)	

More to come (projects, homeworks, etc.) – links will be available from this schedule page

Course Objectives

- After this course, you should be able to analyze data privacy and ethics risks from three non-exclusive perspectives:
 - Data Detectives: Understand how seemingly private information, can be discovered (or exploited) using automated strategies.
 - Data Protectors: Construct privacy protection technologies that provide formal computational guarantees of privacy in disclosed databases.
 - System Builders & Policy Designers: Develop a system or design a policy with built-in privacy mechanisms.

Course Modules



Expectations

- To analyze a dataset: You are expected to have a **working knowledge** of the Internet, word processing, and analysis tools (Python, R, Matlab, Excel, ...)
- To Build a system: You are expected to be **competent** in an object-oriented programming language (Java, C++, Python, ...)
- Reading & Writing in English.

Beyond Expectations

- You have experience in
 - Information security (Cybersecurity expert, hacker)
 - Data structures, algorithms, and statistics (Software engineer, statistician)
 - Public policy and legal frameworks (Lawyer)

Instructional Pedagogy

- Inputs: Lectures, group discussions, case discussions, guest lectures.
- Deliveries (Outputs): In-class quizzes, reading summaries, homework assignments, and course research projects.
- Understanding-based, heuristic teaching approach aligned with international standards.
- Encourage active learning and research interests.
- The primary spoken language is Chinese, while written materials are primarily in English.

Grading Policies

- This is a research-oriented course. There are no exams.
- A substantial portion of your grade will be based on your “final” project.

Criteria	% of Total Grade
Course Project	50%
Homework Assignments	30%
Reading Summaries	5%
In-class quizzes	10%
Attendances and Classroom Performance	5%
Total	100%

More details: <https://zhiyuwan.com/bme2142/#grading>

Homework Policy

- Please do your own homework.
- Do not plagiarize without proper attribution – not even in your reading summaries.
- You can use AI assistant. However, you need to disclose and document the model/agent name and the version number/code.

Reading Summaries

More details: <https://zhiyuwan.com/bme2142/#textbook>

- There is no required textbooks for this course.
- Assigned readings will be available one lecture before it is due (at the latest).
- Your summaries should be no more than 2 pages in length.
- Summaries will be graded on a {A-, A, A+} scale
 - A- : You skimmed the reading and barely understood its meaning
 - A : You read the reading and provided a reasonable account of its contents
 - A +: You demonstrated critical reasoning and insight regarding the topic
- Submit summaries to the instructors and 2 TAs before class.

More details: <https://zhiyuwan.com/bme2142/#reading>

Final Projects

- This course assesses students' mastery of the course content through course projects.
- Each team will consist of **1 to 3 students**.
- Teams will select a topic of interest from a list prepared by the course coordinator, conduct a literature review, identify existing solutions, summarize unresolved problems, propose their own approaches, and locate relevant datasets for experimental validation.

Final Projects (Cont.)

- Each team is required to produce data-driven experimental results, deliver an **oral presentation** and an **poster presentation**, and submit a **final project report** in the form of an English-language academic paper.
- With prior notification and approval from the course coordinator, students may also propose and investigate a project topic outside the provided list.
- Do not be afraid to discuss your project ideas with the instructor!

Course Project Schedule

Week	Course Project Planning
10	Announce course project topics, and students begin forming teams.
11	Students complete team formation and begin selecting course project topics.
13	Each team completes project topic selection and begins preparing a project proposal.
14	Each team presents a project proposal for 5 minutes in class and submits a 4-page project proposal .
15	The instructor and teaching assistants conduct a mid-term evaluation of each team's project proposal and provide feedback and guidance to each team. Students then prepare their final course project presentation based on the feedback.

Course Project Schedule (Cont.)

Week	Course Project Planning
16	According to the score of mid-term evaluation, each team among top 5 (or top 20%) teams gives an in-class oral presentation of approximately 8 minutes. Timing will be adjusted based on the total number of teams, with the presentation order not disclosed in advance. All teams are required to create a poster for display on the classroom wall. The instructor, invited experts, and teaching assistants will provide fair and impartial scoring for each team's presentation and/or poster. For teams delivering an oral presentation, their oral presentation score will be multiplied by 1.2 and then compared with their poster score; the higher score will be counted toward the final grade. Please note that the maximal proportion allocated to the course project in the overall course grade is 50%. For teams that only create a poster, the poster score will be counted directly toward the final grade. Honorary awards will be presented, including Best Oral Presentation, Excellent Oral Presentation, Best Poster, Excellent Poster, and Most Popular Poster. The Most Popular Poster Award will be determined by student voting, with each student casting one vote. All awards will include a trophy, but no cash prize or additional grade points.
17	Students submit the final course project paper (an updated 10-page paper. Note: co-authors are listed without ranking, but each co-author's responsibilities and contributions must be specified in the report).
18	The instructor and teaching assistants conduct a final evaluation of the course project papers.

Grading Policies for the Course Project

- The grading breakdown for the course project is as follows: **project proposal (20%)** + **oral/poster presentation (40%)** + **course project paper (40%)**.
- The grading for the **project proposal** and the **project paper** is assisted by the teaching assistant and is finally determined by the instructor.
- The scores for the **oral/poster presentation** performance are obtained through a weighted average from the instructor (**60%**), invited experts (**20%**) and teaching assistants (**20%**).

Grading Policies for Other Issues

- **Assignments & Quizzes**

- Tests and submitted work are primarily in **English**.
- **2 homework assignments**: Q&A, calculation, programming, etc.; each carries **equal weight**.
- **No collaboration** is allowed.
- Use of **generative AI tools** (e.g., ChatGPT, DeepSeek) must be **disclosed**, including the **model/version used**.
- From week 2 to week 15, there are **5 required readings** in total.
- **2 closed-book in-class quizzes** in **Week 5 and 8**, covering Week 1-4 and 5-7 material, respectively.
- Grades and feedback are normally provided within **4 weeks**.

Grading Policies for Other Issues (Cont.)

- **Attendance & Participation**

- Attendance may be checked by **QR code** and/or random in-class questioning.
- Unexcused absence: **–1% per occurrence**, up to **–5%** of the overall grade.
- Excused absences may receive **no deduction**, as appropriate.
- Lack of active participation: **–1%** of the overall grade.
- Serious classroom disruption: up to **–5%** of the final course grade.

Grading Policies for Other Issues (Cont.)

- **Late Submission Policy**

- Each student receives a total **1-week grace period** across all submissions, including homework, project proposal, and final project paper.
- After the grace period: approximately **–1 point (–1% of overall grade) per additional day**.
- Any delay of **less than one day counts as one full day**.

- **Other Notes**

- Homework may include **bonus questions / bonus points**.

- **Final Grading Conversion**

- It is recommended that if the class size exceeds **30 students**, no more than 40% should receive an A- or above, and the failure rate should not exceed 10%.

Grade	A+	A	A-	B+	B	B-	C+	C	C-	F	EXC	P	NP	N
Grade Point	4.0	4.0	3.7	3.3	3.0	2.7	2.3	2	1.7	0	N/A	N/A	N/A	N/A
Reference Percentage-to-Grade Conversion Scale	95-100	90-94	85-89	80-84	75-79	70-74	67-69	63-66	60-62	0-59	As per the course syllabus	>=60	<60	No grade record

Textbook & Recommended Reading

Book Name: 信息科学技术伦理与道德	Authors: 杨丽凤、虞晶怡	ISBN: 9780128163399
Publisher: 机械工业出版社	Publication Month: 2023.11	Version: 1

Book Name: 工程伦理	Authors: 李正风、王前、丛杭青	ISBN: 9787302524670
Publisher: 清华大学出版社	Publication Month: 2019.06	Version: 2

Book Name: Medical Data Privacy Handbook	Author: Aris Gkoulalas-Divanis, Grigorios Loukides	ISBN: 9783319236322
Publisher: Springer	Publication Month: 2015.12	Version: 1

Book Name: Guide to the De-Identification of Personal Health Information	Author: Khaled El Emam	ISBN: 9780429100659
Publisher: Auerbach Publications	Publication Month: 2013.05	Version: 1

Book Name: Responsible Genomic Data Sharing: Challenges and Approaches	Authors: Xiaoqian Jiang, Haixu Tang	ISBN: 9780128163399
Publisher: Academic Press	Publication Month: 2020.03	Version: 1

Schedule

- Let's look at the syllabus.

More details: <https://zhiyuwan.com/bme2142-schedule/>

Ethics (Sept 23, 30, Oct 10, 14, 21)

- Ethical Issues in Biomedical Research and Informatics
 - AI ethics
 - Research ethics
 - Life sciences ethics
 - Information security and privacy
 - Medical ethics
- Ethical Issues in Data Sharing and Medical AI
 - Data governance and data lifecycle management
 - Algorithmic fairness and bias
 - Transparency and interpretability

Quiz 1



Quiz 2a

The Law and Regulations (Oct 28)

- The impact of international laws and regulations like HIPAA and GDPR on the impact of biomedical data.
- China's data security law and personal information protection law.
- Compliance requirements for medical data sharing.

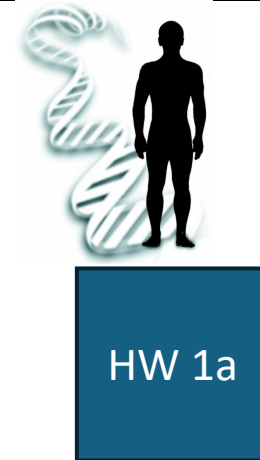
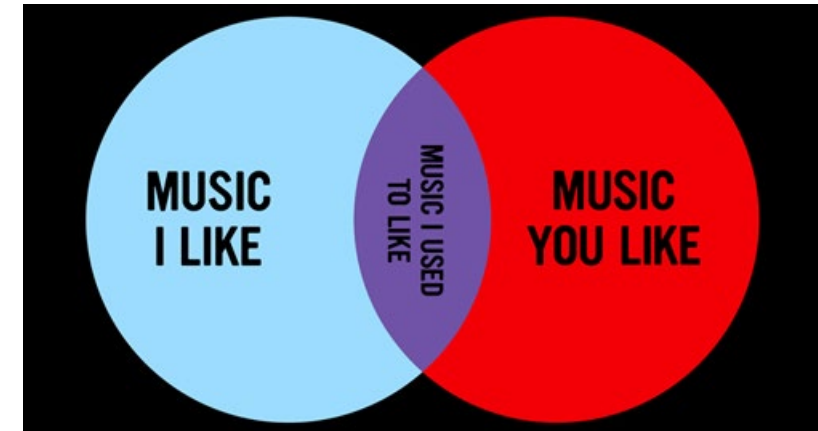


Quiz 2b

(Picture from Brad's Slides)

Privacy Risks of Biomedical Data (Nov 4, 11, 25)

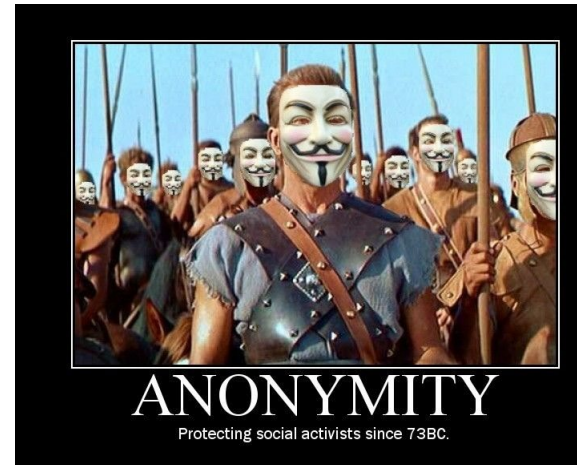
- Biomedical data:
 - Electronic health records
 - Genomic data
- Attack models:
 - Re-identification attacks
 - Membership inference attacks
 - Reconstruction attacks
- Risk Assessment Methods



(Pictures from Brad's Slides)

Common Privacy Protection Techniques for Medical Data (Nov 18, Dec 2)

- Data de-identification techniques
 - K-anonymization
- Game-theoretic models



HW 1b

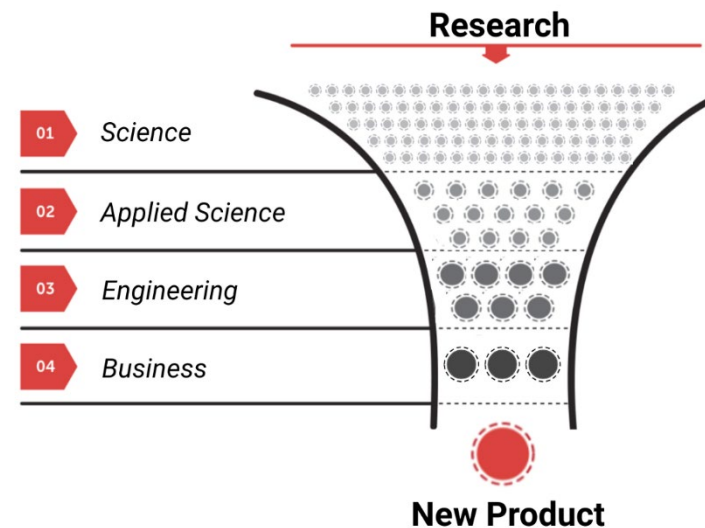
(Pictures from Brad's Slides)

Advanced Privacy Protection Techniques for Medical Data (Dec 9, 16)

- Cryptographical methods
 - Homomorphic encryption
 - Secure multi-party computation
 - Encrypted hardware
- Differential privacy
- Federated learning
- Synthetic data generation



HW 2

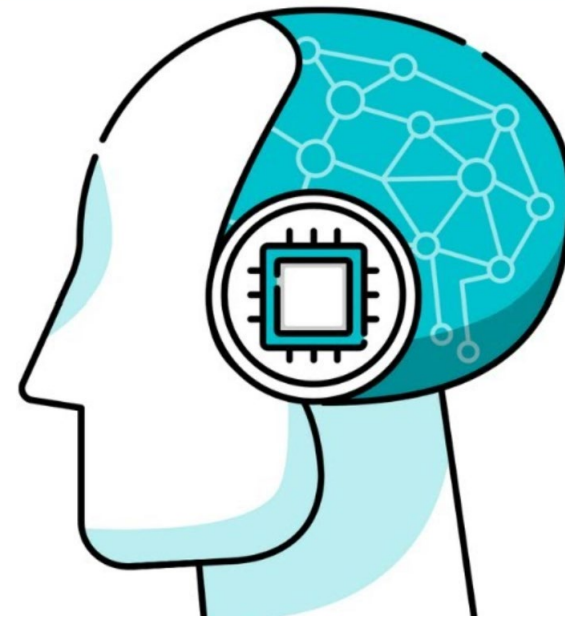


(Picture from Brad's Slides)

Privacy and Ethical Issues in Cutting-Edge AI Technologies for Healthcare (Dec 23)

■ Large Language Models & Generative AI

- Introduction
- Applications
- Privacy issues
- Fairness issues
- Other ethical issues
- Solutions and future directions



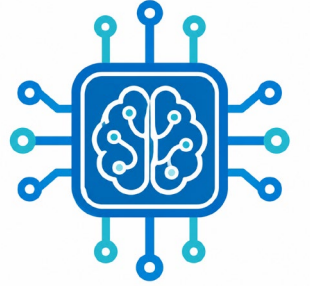
Project

Course Project Presentation (Dec 30)

- The students are in control
- You'll be graded by a committee of special reviewers



(Picture from Brad's Slides)



Introduction to AI

- **Definition of AI:** AI refers to the simulation of human intelligence in machines that are programmed to think, learn, and make decisions.

- **Types of AI:**
 - Narrow AI: Specialized in performing specific tasks.
 - General AI: Hypothetical, with abilities similar to human cognitive functions.

Development of AI

- **Early Development:** The origins of AI date back to the 1950s, with pioneers like Alan Turing and John McCarthy.



- **Advancements in Machine Learning:**
 - 1990s: Introduction of supervised and unsupervised learning.
 - 2010s: Emergence of deep learning, enabling AI to handle vast datasets like images and speech.

Age of AI

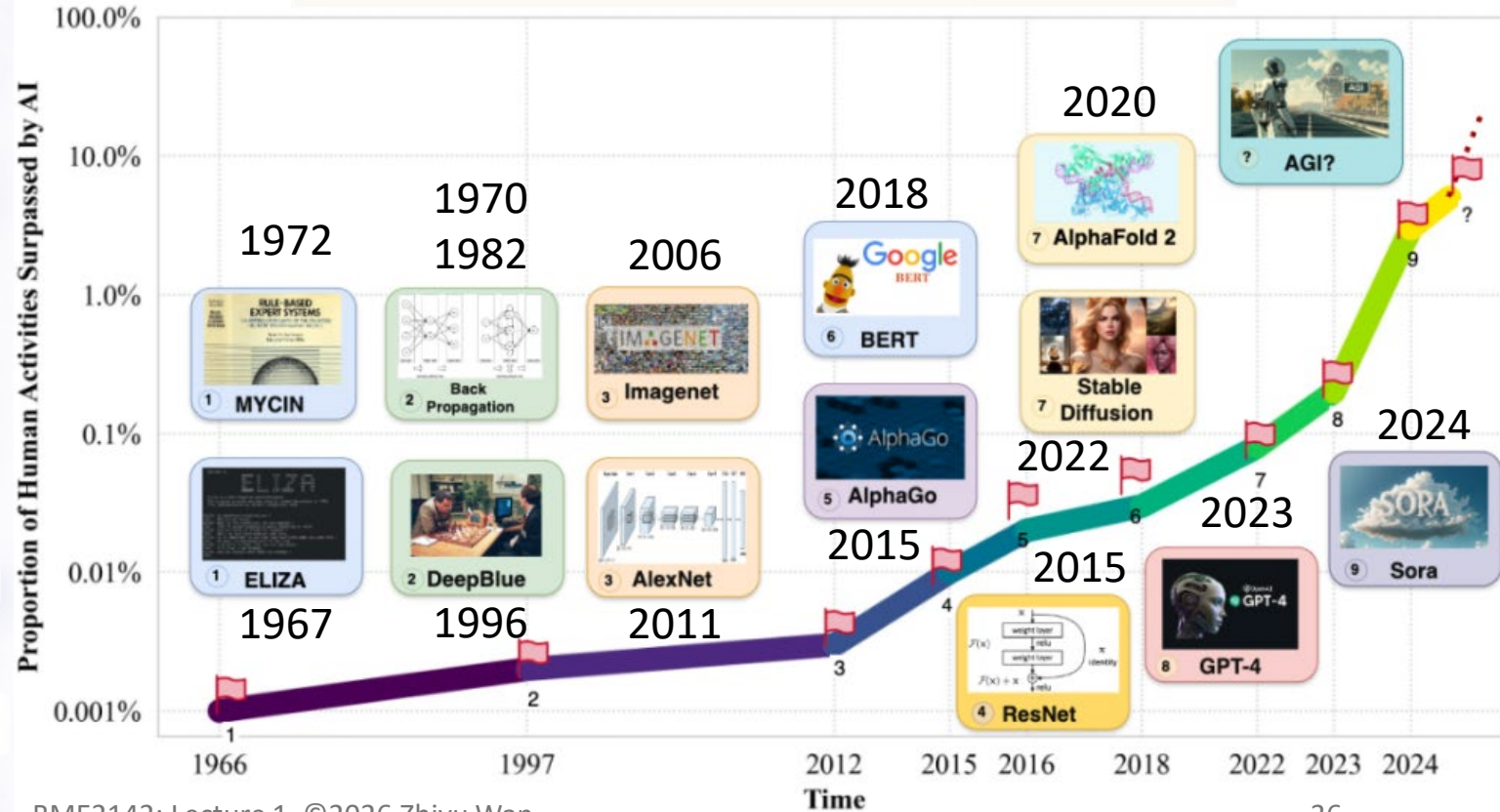
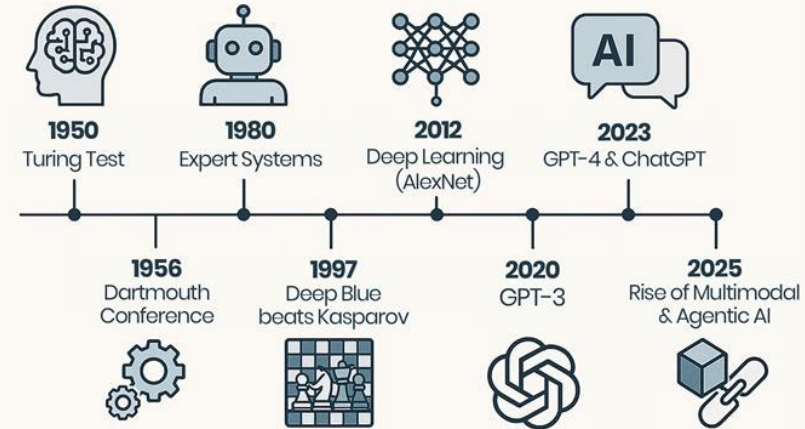
The Evolution of AI (1956 - 2026)

From Idea → to Smart Machines



Next: AI Everywhere - Work, Education, Healthcare, Robots

HISTORY OF ARTIFICIAL INTELLIGENCE



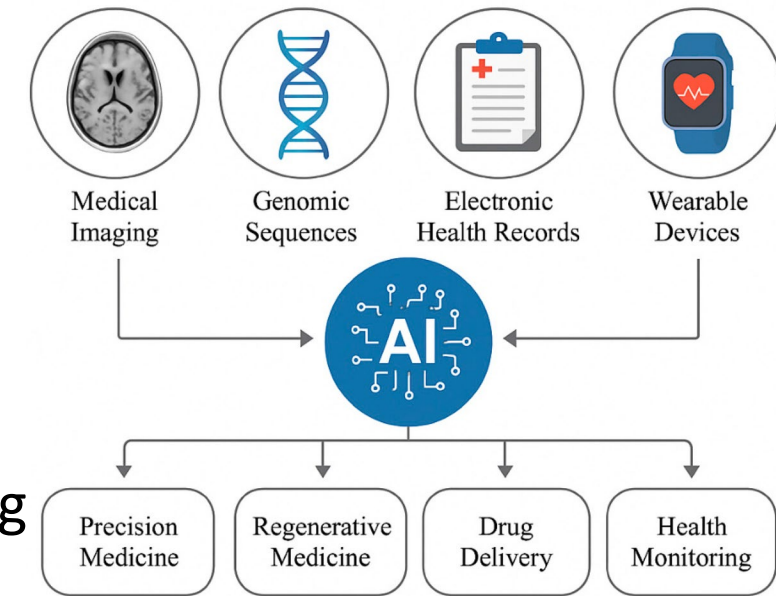
Medical Data in the Age of AI

■ Types of Medical Data:

- **Electronic Health Records (EHRs):** Patient data including medical history, diagnoses, and treatments.
- **Medical Imaging:** X-rays, MRIs, CT scans, etc.
- **Genomic Data:** DNA sequences and genetic information.
- **Wearable Device Data:** Vital signs, activity data, etc.

■ Role of Medical Data

- **Training AI Models:** AI models require vast amounts of medical data to learn patterns and make accurate predictions.
- **Data-Driven Insights:** Medical data allows AI to assist in diagnosing diseases, predicting outcomes, and personalizing treatment plans.
- **Challenges:** Ensuring data quality, privacy, and compliance with regulations.

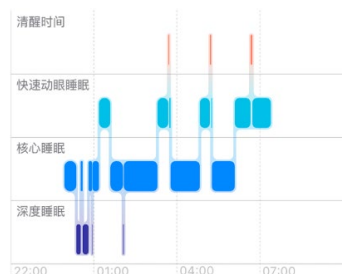


Parvin N, Joo SW, Jung JH, Mandal TK. Multimodal AI in Biomedicine: Pioneering the Future of Biomaterials, Diagnostics, and Personalized Healthcare. *Nanomaterials*. 2025 Jun 10;15(12):895.

Electronic Health Records



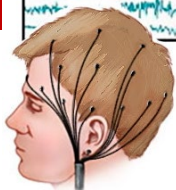
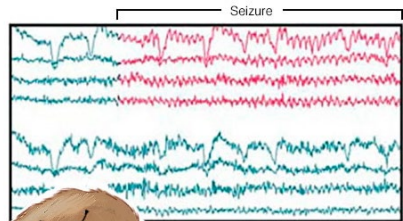
BPM/ECG



Sleeping Cycles



Genomics / Omics



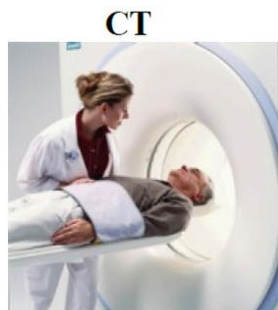
EEG



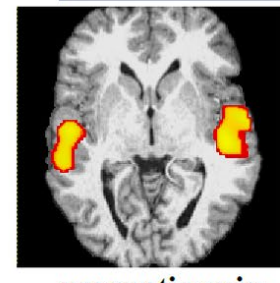
Endoscopic Images



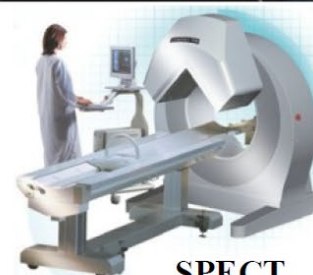
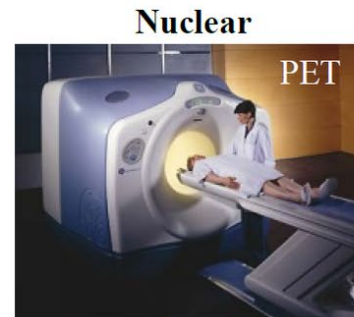
Voice



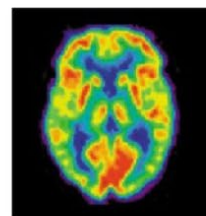
X-ray



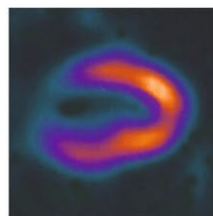
magnetic spin



SPECT



metabolic tracer X-ray emission



sound waves

Medical Imaging

Biomedical Informatics $\approx$ Biomedical Data Science



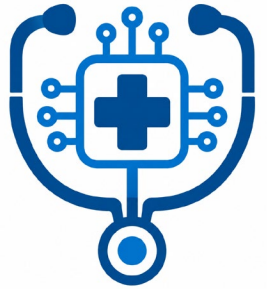
Randolph A. Miller, MD, FACMI



- The **Cornelius Vanderbilt** Professor (Emeritus) of Biomedical Informatics.
- The **founding Chair** of the Department of Biomedical Informatics (DBMI) from **2001-2004**.
- The initial DBMI mission was to develop and evaluate leading-edge **biomedical software applications** to improve the **quality of care**, promote **research**, and enhance **patient safety**.

AI in Medicine

- **AI in Medical Diagnostics**
 - **AI for Image Recognition:** Using deep learning to analyze medical images like X-rays, MRIs, and CT scans for early disease detection (e.g., detecting cancer).
 - **AI in Pathology:** AI-powered systems to analyze pathology slides.



nature reviews clinical oncology

[Explore content](#) ▾ [About the journal](#) ▾ [Publish with us](#) ▾

[nature](#) > [nature reviews clinical oncology](#) > [research highlights](#) > [article](#)

Research Highlight | Published: 21 January 2020

BREAST CANCER

AI outperforms radiologists in mammographic screening

[David Killock](#) 

[Nature Reviews Clinical Oncology](#) **17**, 134 (2020) | [Cite this article](#)

11k Accesses | **297** Altmetric | [Metrics](#)

AI in Medicine

- **AI in Personalized Medicine**
 - **Precision Medicine:** AI models can process genomic data to tailor treatments based on individual genetic makeup.
 - **Predictive Analytics:** AI algorithms can predict the likelihood of disease recurrence and suggest the most effective treatment options.



► J Med Internet Res. 2018 Sep 25;20(9):e11087. doi: [10.2196/11087](https://doi.org/10.2196/11087)

Using Artificial Intelligence (Watson for Oncology) for Treatment Recommendations Amongst Chinese Patients with Lung Cancer: Feasibility Study

[Chaoyuan Liu](#)¹, [Xianling Liu](#)¹, [Fang Wu](#)¹, [Mingxuan Xie](#)², [Yeqian Feng](#)¹, [Chunhong Hu](#)¹,

Editor: Carlos Luis Parra-Calderón

Reviewed by: Francisco Nuñez-Benjumea, Edward Meinert, Robert Robinson

► [Author information](#) ► [Article notes](#) ► [Copyright and License information](#)

PMCID: PMC6231834 PMID: [30257820](#)

AI in Medicine

- **AI in Drug Discovery and Development**
 - **Accelerating Drug Discovery:** AI analyzes massive datasets of molecular information to identify potential drug candidates.
 - **AI for Clinical Trials:** AI is used to identify suitable patient cohorts, predict trial outcomes, and optimize trial designs.



Insilico Medicine uses AI to discover novel SIK2 inhibitors

[Download PDF Copy](#)

Reviewed

Reviewed by [Danielle Ellis, B.Sc.](#)

Aug 2 2023

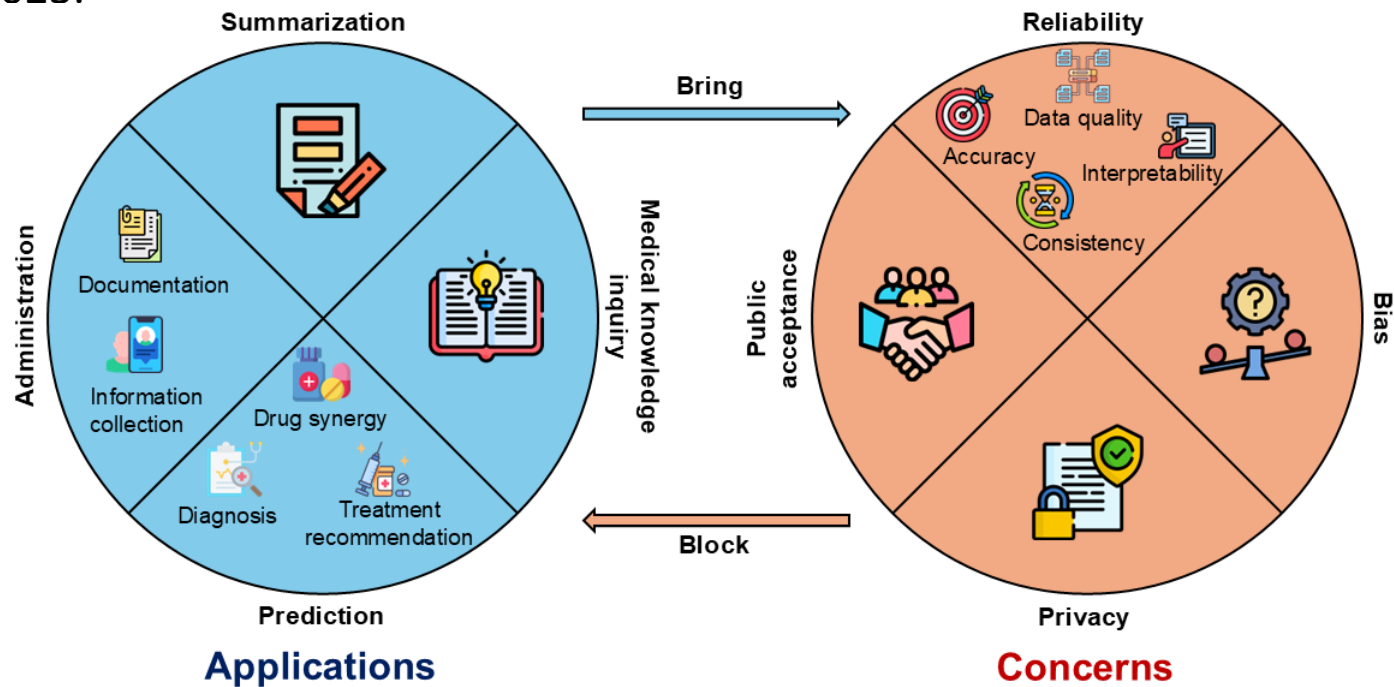
Insilico Medicine ("Insilico"), a clinical-stage end-to-end generative artificial intelligence (AI) drug discovery company, has achieved a significant breakthrough in the application of multiple generative AI models and AlphaFold structures for drug discovery.

Applying Insilico's generative chemistry engine to AlphaFold-predicted protein structures, researchers discovered novel and selective inhibitors for salt-inducible kinase 2 (SIK2), a potential target for anti-inflammation and anti-cancer therapy. SIK2 is highly overexpressed in 30% of human ovarian cancers. The findings were published in the July 13 edition of *Bioorganic & Medicinal Chemistry*.

“Utilizing the capability of Chemistry42 and AlphaFold predicted structures, a series of novel, potent and selective SIK2 inhibitors were identified through structure-based design strategy. This work further demonstrates the power of Insilico's Pharma.AI platform.”

Applications and Concerns of Large Language Models in Health Care

- A summary of the applications and concerns regarding LLMs in healthcare as communicated by 65 reviewed research papers selected from a pool of 820 articles sourced from PubMed, ACM, and IEEE, published before September 1st, 2023.



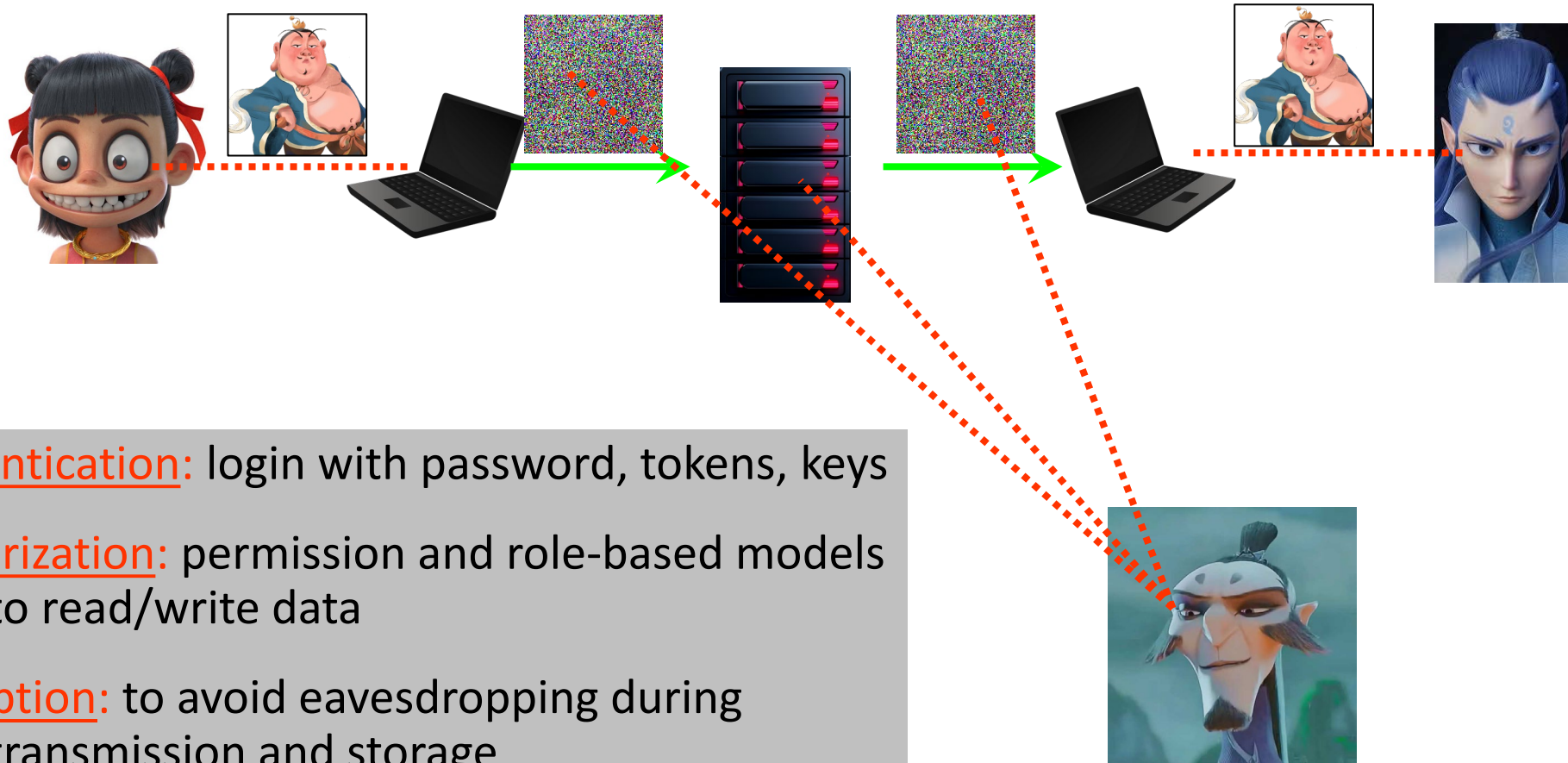
Wang L*, **Wan Z***, Ni C, Song Q, Li Y, Clayton E, Malin B, Yin Z. Applications and Concerns of ChatGPT and Other Conversational Large Language Models in Health Care: Systematic Review. *Journal of Medical Internet Research*. 2024 Nov 7;26:e22769.



So... What is Privacy?



Security for Privacy?



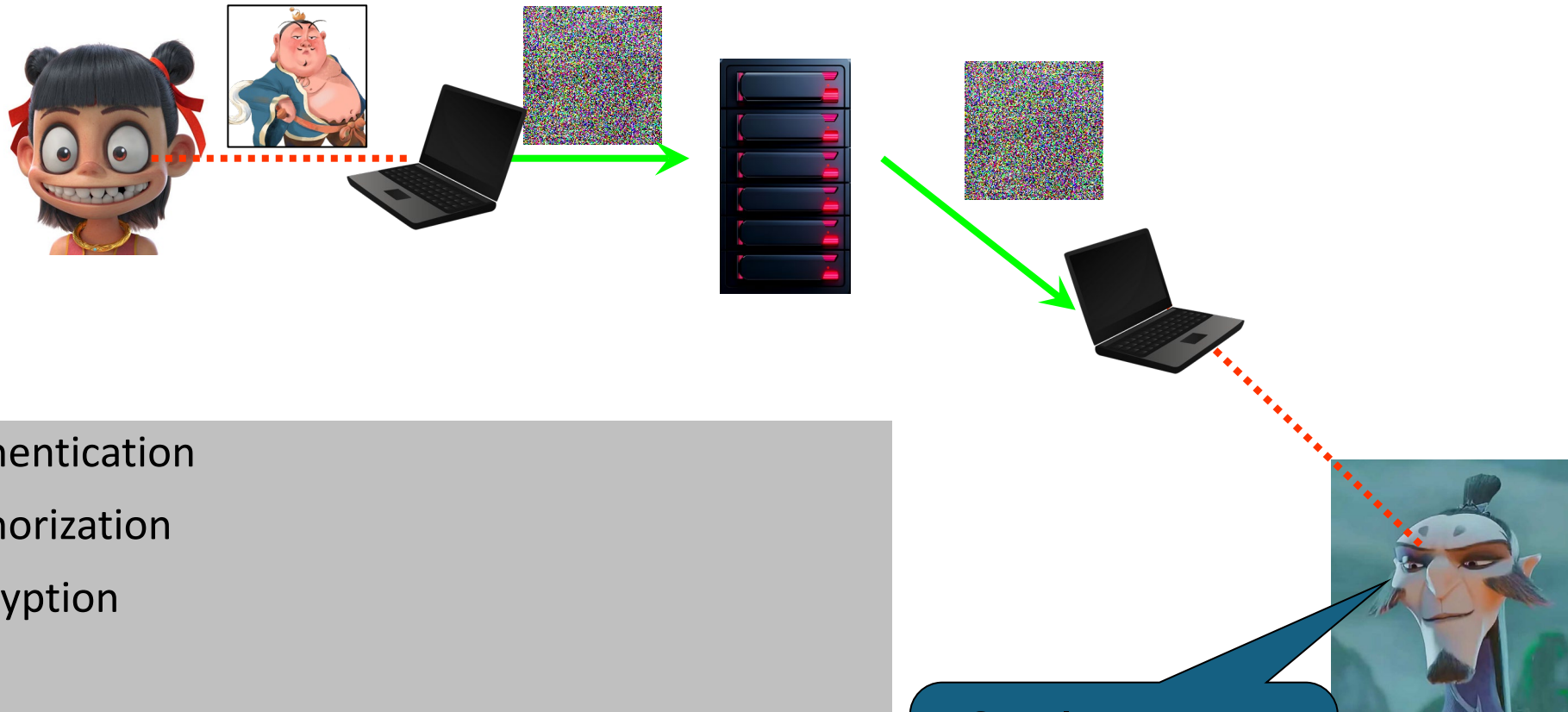
Authentication: login with password, tokens, keys

Authorization: permission and role-based models to read/write data

Encryption: to avoid eavesdropping during transmission and storage

(Adapted from Brad's Slides)

Security for Privacy?

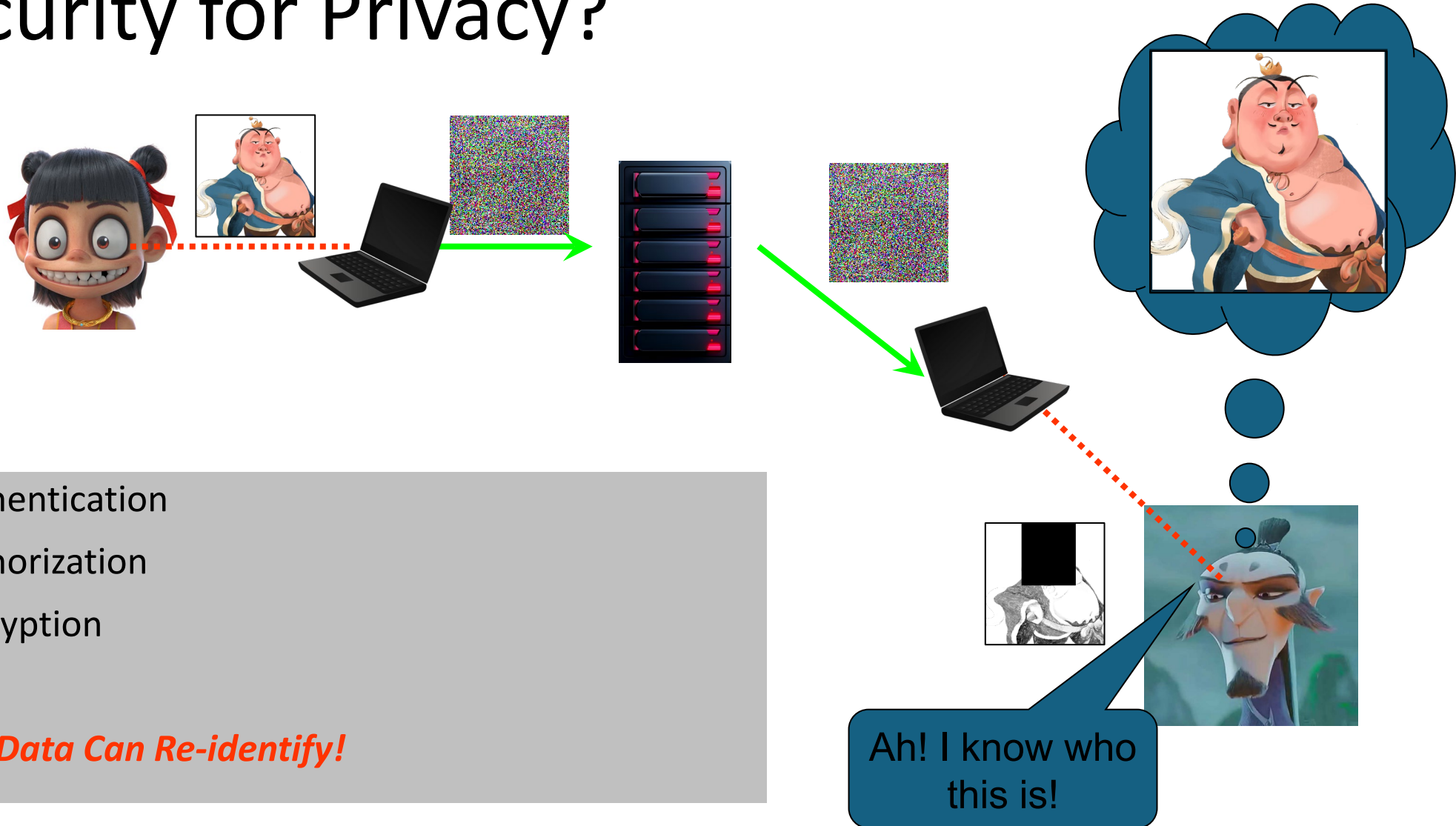


- Authentication
- Authorization
- Encryption
- ***But Data Can Re-identify!***

Can I see some anonymous data?

(Adapted from Brad's Slides)

Security for Privacy?



- Authentication
- Authorization
- Encryption
- ***But Data Can Re-identify!***

(Adapted from Brad's Slides)

Data Privacy Definitions

- **Data privacy** refers to the principles and practices that protect individuals' rights and autonomy regarding the collection, use, storage, sharing, and disclosure of their personal data. (GPT-5.6 Sol)
- **Data privacy** is the condition in which human autonomy and dignity are safeguarded through the controlled collection, use, and disclosure of personal information, achieved via mechanisms such as confidentiality, predictability, and disassociability. (DeepSeek-V4.1-Flash)
- **Data privacy** is a condition that protects human autonomy and dignity through properties including confidentiality, predictability, manageability, and disassociability. (NIST, National Institute of Standards and Technology)

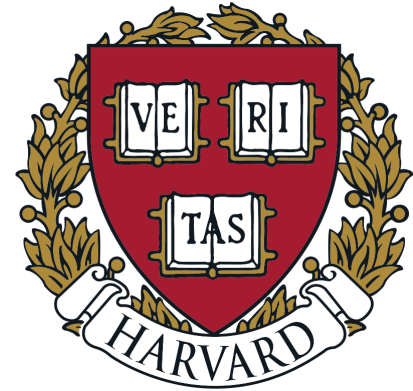
Data Privacy Definitions (Cont.)

- **Data privacy** refers to the principles and regulations governing how personal information is collected, stored, used, and shared, emphasizing an individual's right to control their own details.
(Dictionary.com, AI generated)
- **Data privacy** is the study of computational solutions for releasing data such that (paraphrase Sweeney)
 - a) the data is practically useful (utility) while
 - b) the aspects of the subjects of the data are not revealed (privacy).

Pioneer of Data Privacy

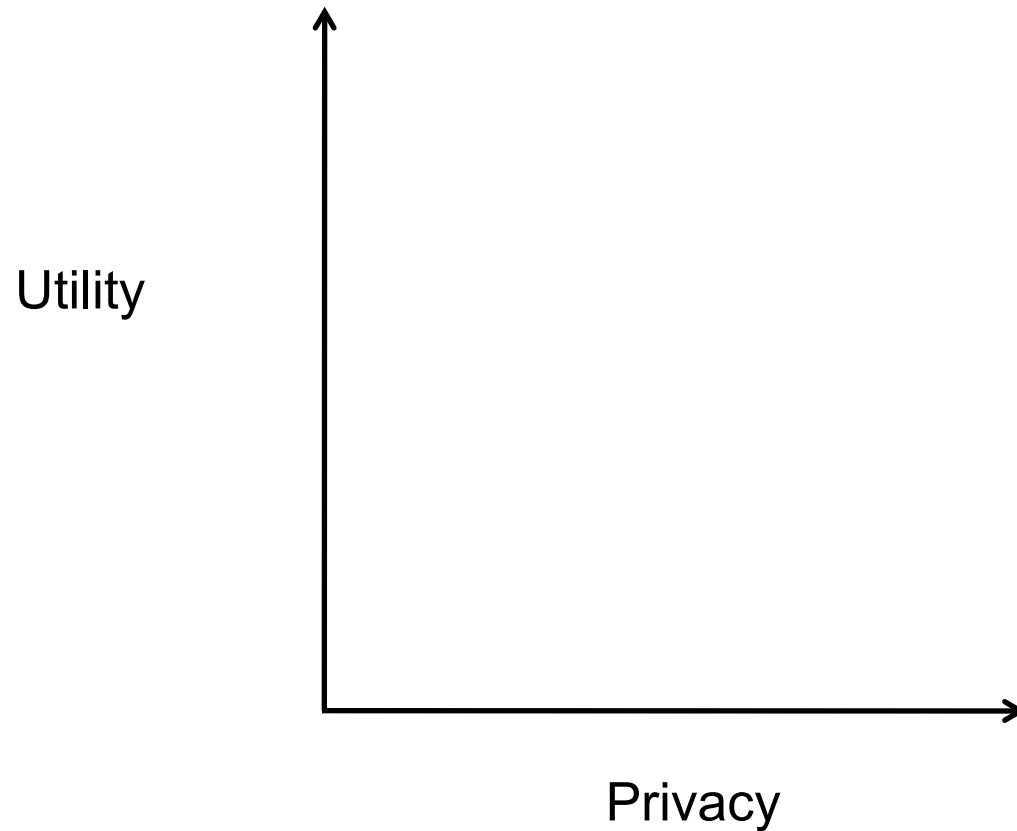


Latanya Sweeney, PhD, FACMI

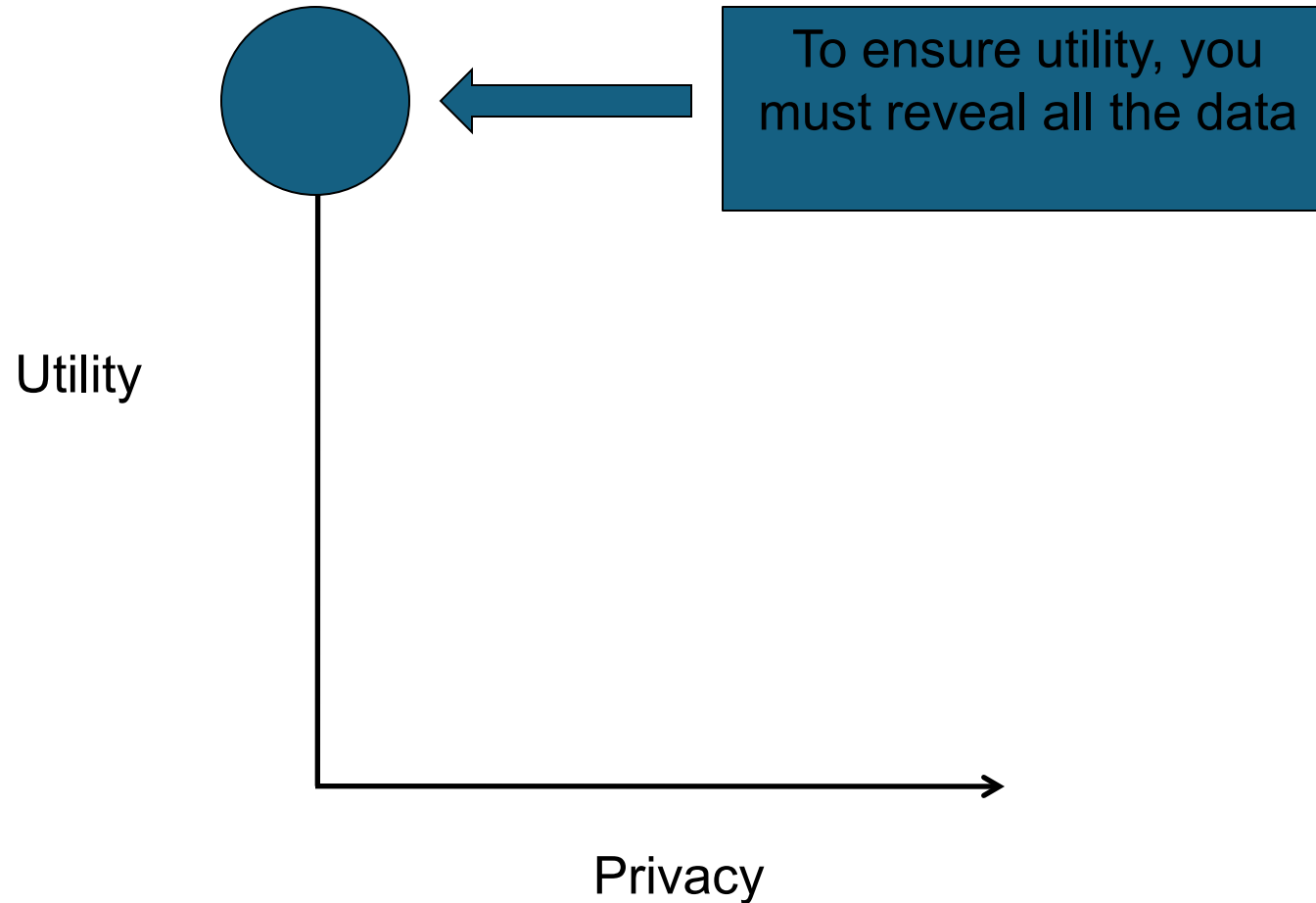


- The Daniel Paul Professor of the Practice of Government and Technology at the Harvard University.
- In 2001, she founded the **Data Privacy Lab** at Carnegie Mellon University.
- She pioneered the field known as **data privacy**, launched the emerging area known as **algorithmic fairness**.
- Her best-known academic work is on the theory of **k-anonymity**.

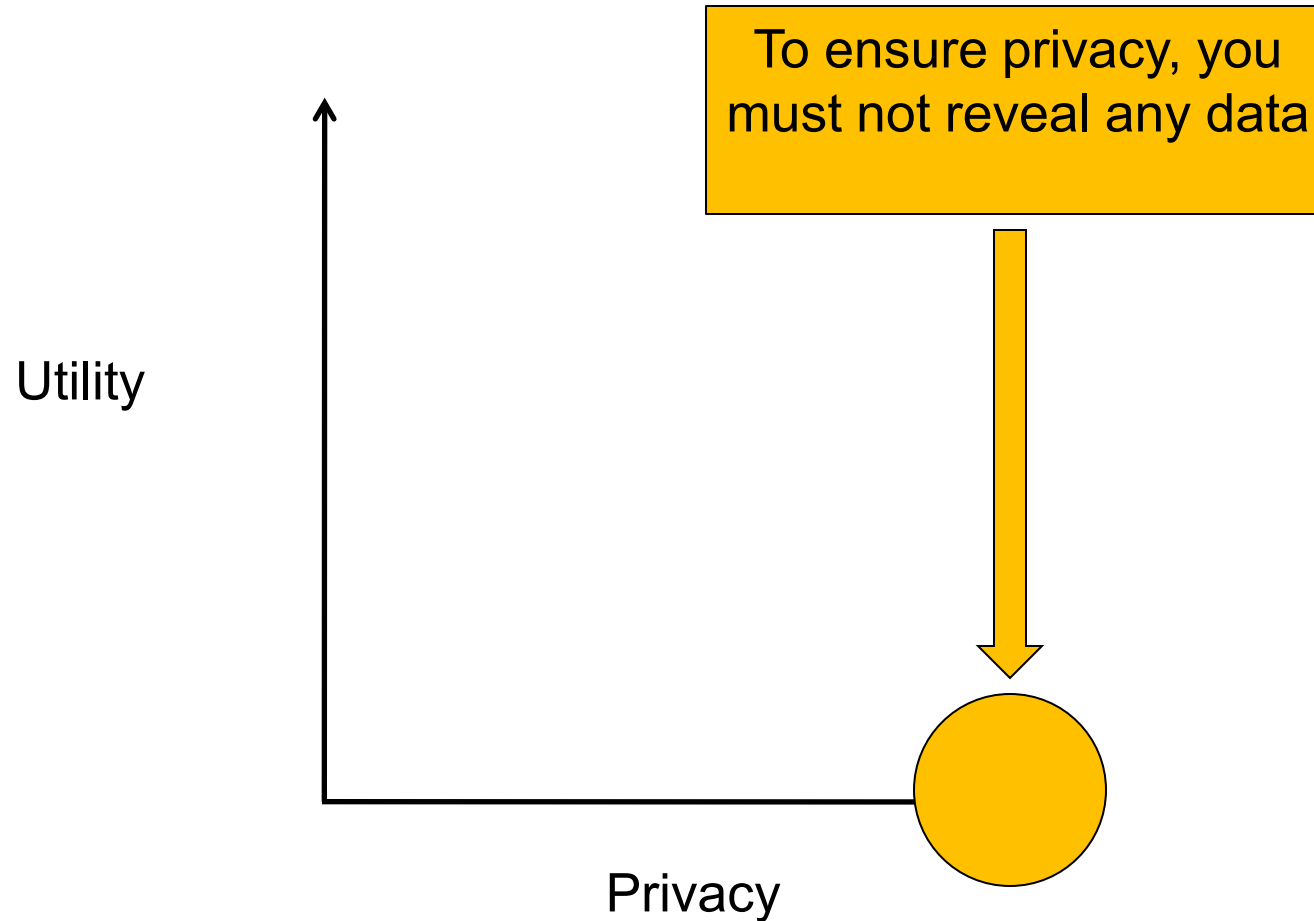
A Visual Perspective



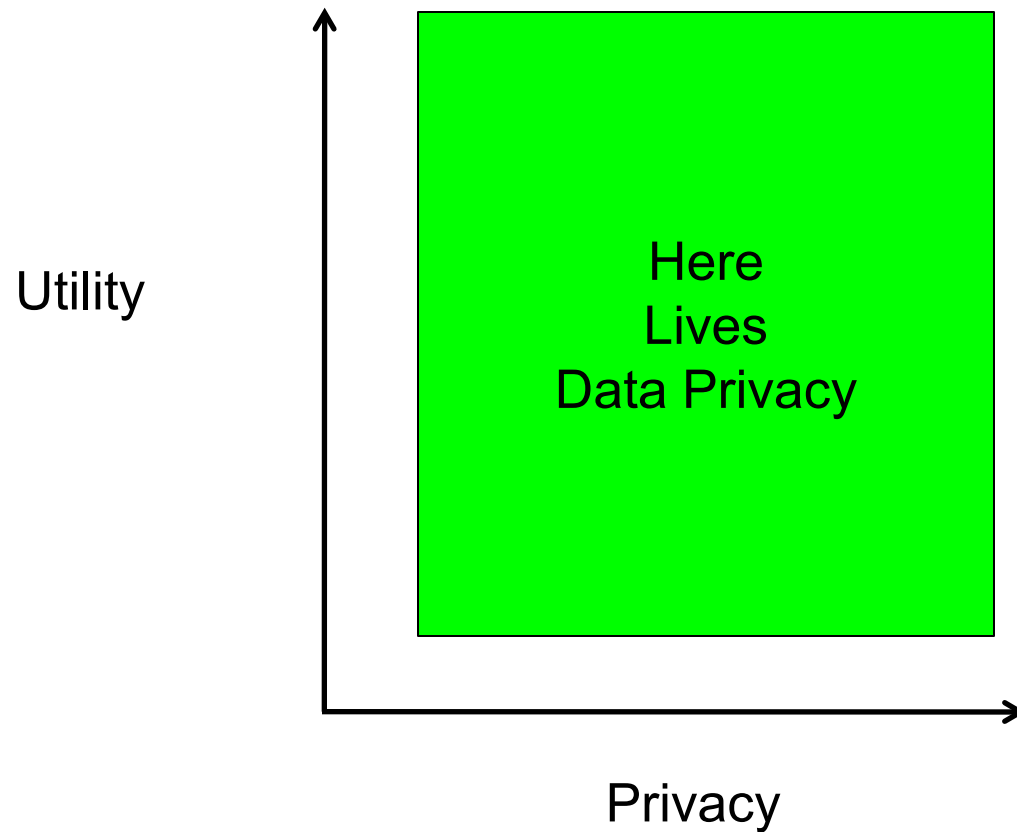
A Visual Perspective



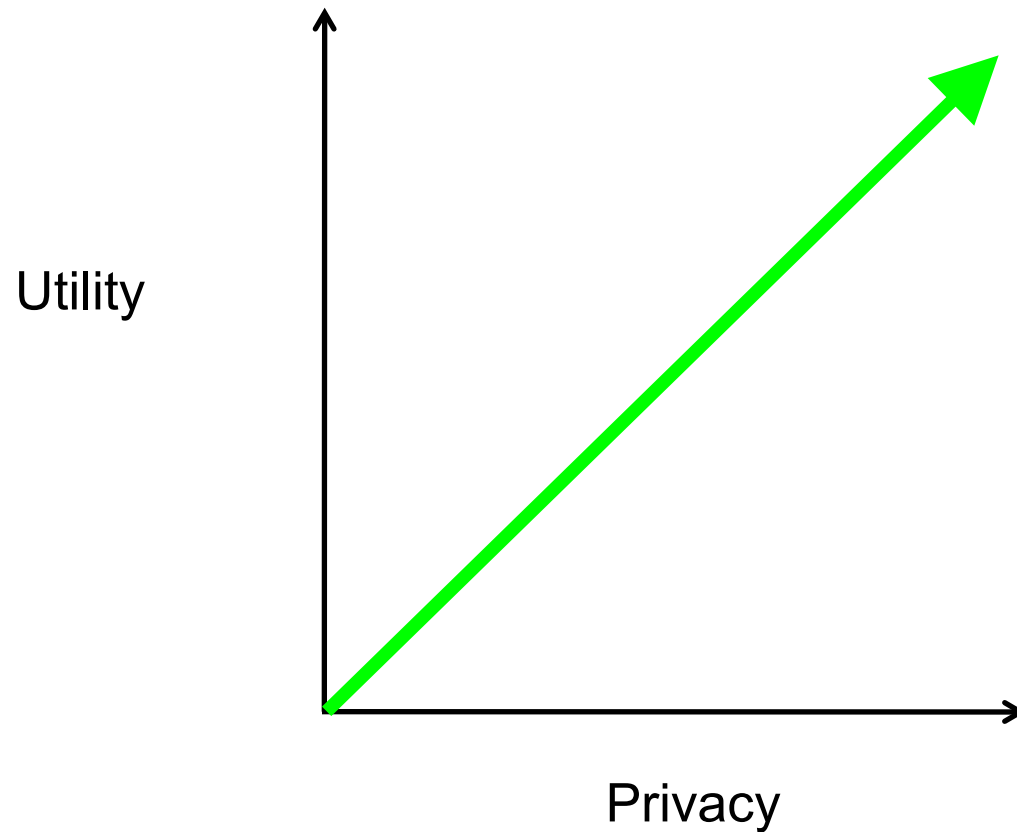
A Visual Perspective



A Visual Perspective

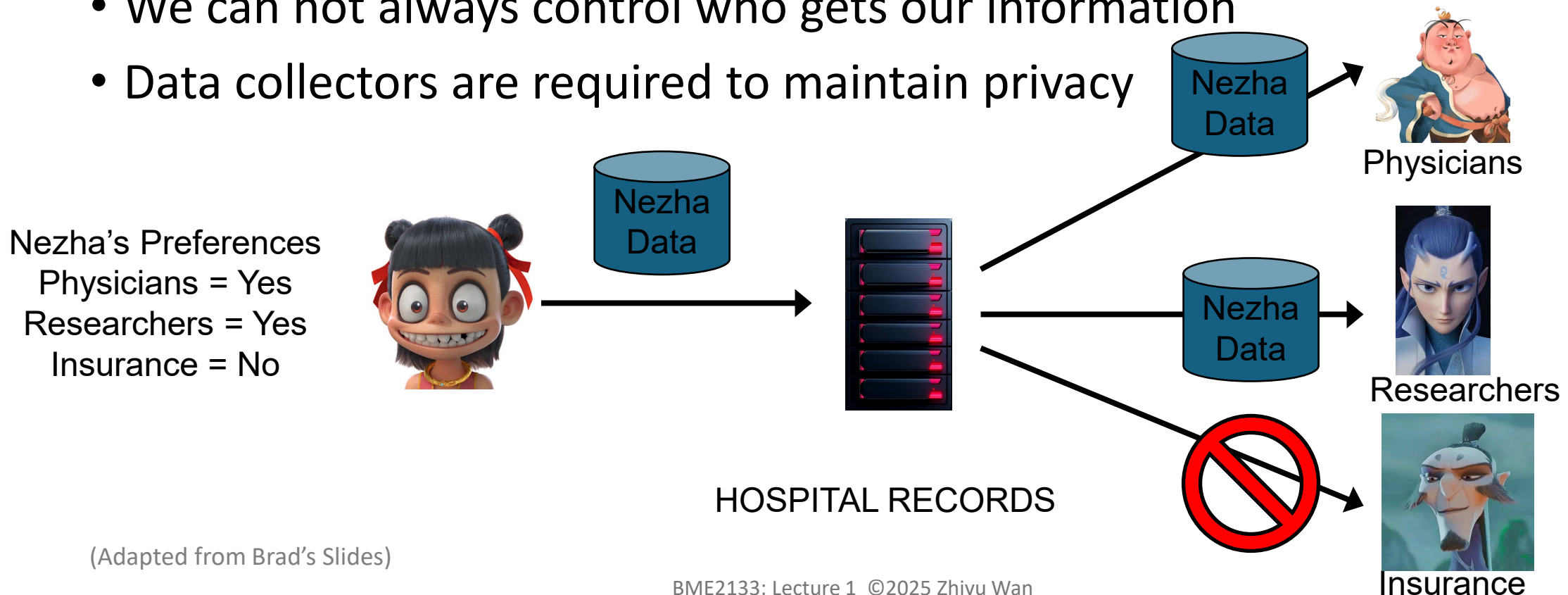


A Visual Perspective



Data sharing, Privacy, & Policy in Healthcare

- Individuals want control over who can – AND CAN NOT – view their health-related records
- We can not always control who gets our information
- Data collectors are required to maintain privacy



(Adapted from Brad's Slides)

Biomedical Information



- Not quite in the public
- But... information is shared for various purposes in various contexts
- How do you protect privacy of corresponding individuals?



(Adapted from Brad's Slides)

Definitions



- **Ethics:** The systematic set of principles or guidelines that govern conduct within a particular context, often derived from philosophical theories and professional standards. It deals with questions of what is right and wrong, and what individuals ought to do in various situations.
- **Values:** Individual or collective beliefs about what is important, desirable, and worthwhile. They reflect personal or cultural priorities and inform decisions and behaviors.
- **Morality:** The principles or rules of behavior that individuals or societies believe are right or wrong. It is more focused on practical, everyday conduct and often carries a strong emotional and social component.

Etymology of Ethics

- Etymologically, the English word “ethics” (ethica in Latin) can be traced back to the ancient Greek noun, (ethos), which denotes a “habit” or “custom”.
- Ethics is a practice discipline that refers to human action with the purpose of being morally good.
- In Chinese, “伦” means “Order” which represents relationships; “理” means “Rule” which represents principles.
- “Morality” is more subjective. “Ethics” is more objective.

Readings for the Next Lecture

- None.

- Optional

- ☐ 《信息科学技术伦理与道德》 Chs.8-10.
- ☐ 《工程伦理》 Ch.12.
- ☐ 《信息科学技术伦理与道德》 Chs.3&4.
- ☐ 《工程伦理》 Ch.10.
- ☐ 《信息科学技术伦理与道德》 Chs.5&7.