

Medical Informatics

Lecture 1: Introduction

Dr Areti Manataki



Nanjing Medical University

Welcome!

- Medical Informatics: a gentle introduction to data science in biomedicine and healthcare
- Dr Areti Manatakis
 - Lecturer at the University of St Andrews, United Kingdom
 - Health informatics research: artificial intelligence methods to improve the delivery of healthcare services
 - Extensive experience of teaching health data science to a variety of audiences.
 - <http://homepages.ed.ac.uk/amanatak/>

Welcome!

- Medical Informatics: a gentle introduction to data science in biomedicine and healthcare
- Our aim: to equip you with the **key foundations** and **data skills** that are needed for the data-intensive medicine of the future
- In line with recommendations in the Topol Review and by the National Academy of Medicine

Data is everywhere!

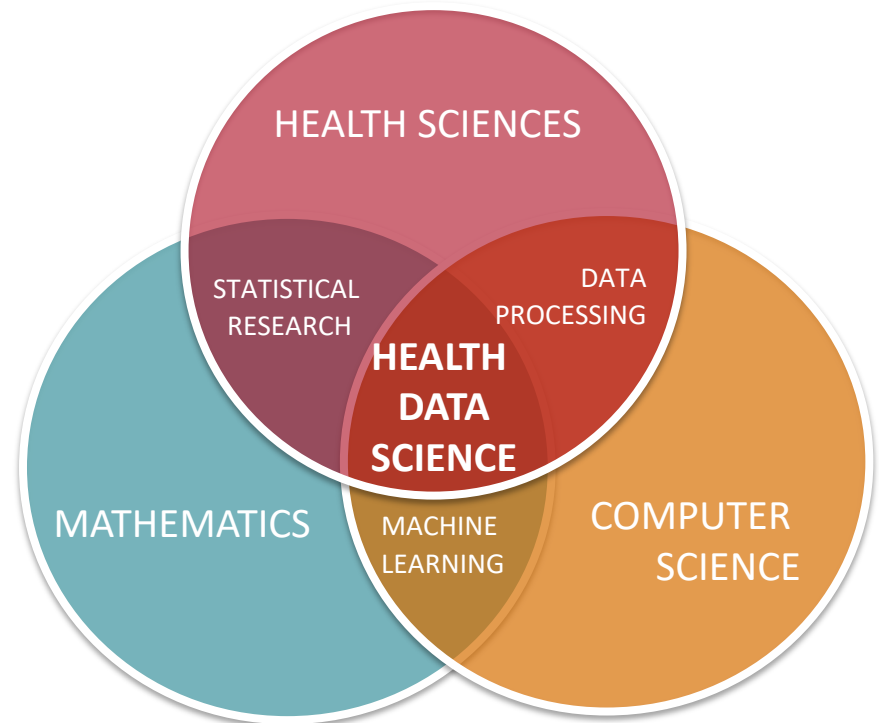
- Genomic data
- Electronic patient records
- Medical images
- Clinical notes
- Wearable devices
- Social media posts
- Bank card transactions
- ...and much much more

[illegible]

[Images by Louis Philippe Lessard and NurseRecord on Wikimedia, licensed as CC BY-SA 3.0]

Demystifying Health Data Science

- **Health Data Science** is a discipline that combines mathematics, computer science and health sciences to study different types of health problems using data.
- It provides the tools to manage and analyse very large amounts of different datasets across our healthcare systems.



*[Image adapted from Shelly, P. 2015.
Data Science for the C-Suite]*

Opportunities brought by Health Data Science

Data and its analysis is revolutionising how medicine is understood, how biomedical research is conducted and how healthcare is delivered.

- Better understand disease
- Earlier and improved diagnosis
- Prevention of disease
- Enhanced predictions (e.g. in epidemiology)
- Safer and more effective treatments
- More effective integrated care pathways
- Driving clinical research
- Precision medicine

Challenges for Health Data Science

- Volume of data
- Data quality
- Data complexity
- Data sharing and linkage in a complex environment
- Digital maturity of health systems
- Culture in national health systems
- Data ethics
- Training the medics of the future

This course

How do we represent and interpret data?

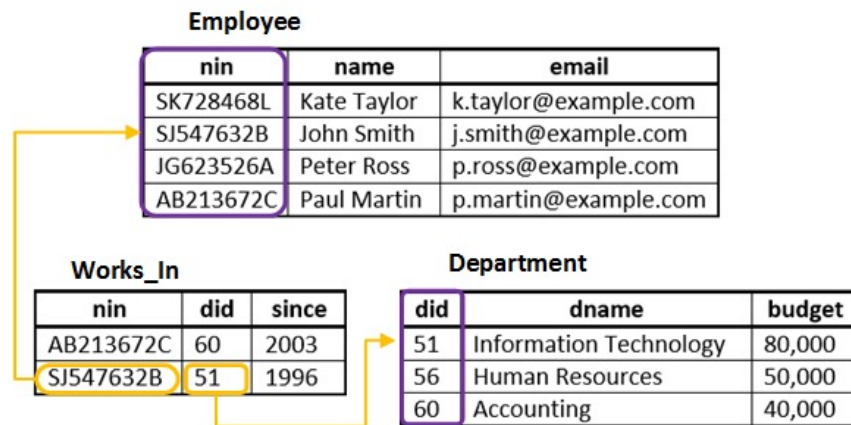
- Variety of approaches and methods available
- We'll have a look at:
 - Relational databases for medicine and healthcare
 - Medical ontologies and semantic web data
 - Advanced topics in Medical Informatics

Course syllabus

- Relational databases for medicine and healthcare
 - Design & Representation: ER model, relational model
 - Querying: SQL
- Medical ontologies and semantic web data
 - Concepts: ontologies, linked data
 - Representation: RDF
 - Querying: SPARQL
- Advanced topics in Medical Informatics
 - Artificial Intelligence in medicine and healthcare
 - Imaging and text data in medicine and healthcare

Relational databases for medicine and healthcare

- Tables with columns and rows and links between them

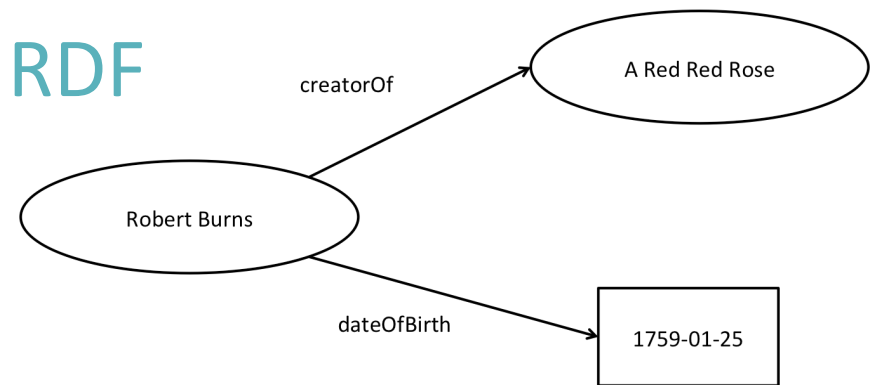


- How to design them: **ER model**
- How to specify tables: **DDL**
- How to query the data: **SQL**

Medical ontologies and semantic web data

- Ontologies: class hierarchy, relations, axioms, etc.

- How to represent data: **RDF**



- How to query the data: **SPARQL**

Copyright

This course is based on the Medical Informatics course from The University of Edinburgh, which is licensed under a Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0) license.

Unless otherwise stated, all content of the current course, including lecture slides, quizzes, tutorial and lab worksheets, is also licensed under a CC BY-SA 4.0 license.



Course structure

- The course runs over 10 days across 3 weeks.
- It is offered fully online.
- Learning activities:
 - Lectures
 - Tutorials
 - Labs
 - Coursework

Lectures

- Delivered live online by Dr Manataki.
- Active participation is expected.
- Three themes covered:
 1. Relational Databases for Medicine and Healthcare (lectures 2-7)
 2. Medical Ontologies and Semantic Web Data (lectures 8-12)
 3. Advanced Topics in Medical Informatics (lectures 13-14)
- Lectures are focussed on the technical side of things and are using examples from everyday life. The application to biomedicine and healthcare is achieved through tutorials and coursework.

Tutorials

- Tutorials will give you the opportunity to:
 - Apply what you've learnt in the lectures to particular scenarios and examples
 - Reflect on topics covered in the lectures
 - Clarify any questions you may have
- The main focus of tutorials will be practical exercises.
- Tutorial sheets will be released one week before the corresponding tutorial.
- You are expected to attempt all exercises in advance. There will also be some questions for group discussion in class.
- Tutorial sessions take place online and are facilitated by course staff.

Labs

- Labs will demonstrate how you can use particular software for representing and manipulating data in the different data models covered in the course:
 - MySQL
 - Apache Jena ARQ
 - WebProtégé
- Delivered live online by Dr Manataki.
- During the labs we'll go through practical examples of working with data.
- You will need to install the software in advance.

Optional textbooks

No compulsory textbooks for this course. If you are interested in reading further, we recommend the following optional textbooks:

- Relational databases for medicine and healthcare:
 - Raghu Ramakrishnan and Johannes Gehrke. Database Management Systems. McGraw-Hill, 3rd edition, 2003.
 - S. Sumathi and S. Esakkirajan. Fundamentals of relational database management systems. Springer, 2007.
- Medical ontologies and semantic web data:
 - Dean Allemang and Jim Hendler. Semantic Web for the Working Ontologist: Effective Modelling in RDFS and OWL. Morgan Kaufmann, 2nd edition, 2011.
 - Tom Heath and Christian Bizer. Linked Data: Evolving the Web into a Global Data Space. Morgan & Claypool Publishers, 2011.

Assessment

- No exams for this course.
- Your final grade will be based on the final course quiz, consisting mostly of multiple choice questions.
- The quiz will cover both theoretical and practical knowledge.
- Key dates for quiz:
 - Quiz released on 12th November 2021.
 - Quiz submission deadline on 26th November 2021.
- Coursework is individual. No groupwork is allowed.

People

- Course lead: Dr Areti Manataki A.Manataki@st-andrews.ac.uk
<http://homepages.ed.ac.uk/amanatak/>
- Course assistants:
 - Dr Cheng Wan chengwan@njmu.edu.cn
 - Dr Wei Feng 544299055@qq.com
 - Ms Huiting Sun 826064962@qq.com

Conclusions

- Medical Informatics: a gentle introduction to data science in biomedicine and healthcare
- Fundamental question: How do we represent and interpret data?
- Over the next few weeks, we'll be looking at:
 - Relational databases for medicine and healthcare
 - Medical ontologies and semantic web data
 - Advanced topics in medical informatics
- Bring your questions and ideas to class!