≜UCL

Spatial Databases and Data Management

- Dr Claire Ellul
- c.ellul@ucl.ac.uk

** Data is an asset **

≜UCL

Spatial Databases

- Moodle
  - https://moodle-1819.ucl.ac.uk/course/view.php?id=1339
- No automatic enrolment
- Passcode: CEGE-ELL1

- I will upload slides before the lectures so that you can print them out if you like

---

≜UCL

Spatial Databases

- Overview
  - **Introduction to the Module**
  - Data and Database
  - What is Spatial Data

≜UCL

Spatial Databases

- What this module is about
  - Data: design, storage, management and query
  - Databases and structured query language (SQL)
  - Focus on spatial data ('can be mapped')
  - 2D and 3D spatial data, BIM will be mentioned
- What this module is not about
  - Big data (although you can store and analyse 100s of millions of records in a database!)
  - Analytics, Artificial Intelligence (specialist modules at UCL on these topics – this is a good foundation for those modules)

≜UCL

Spatial Databases

- Hours: 150 (all MSc modules)
  - Lectures 18 hours
  - Practical 12 hours (in the lab or at home)
  - Self-Guided Learning and Assignment 120 hours

≜UCL

Spatial Databases

- Week 1
  - Lecture 1:
    - Introduction to databases and spatial data
- Week 2
  - Lecture 2:
    - Database Design
      - Entity Relationship Diagrams

Spatial Databases

Spatial Databases

- Week 3
  - Lecture 3
    - SQL – Structured Query Language
      - DDL: Creating Tables and Structures
      - DML: Inserting, Editing and Deleting Data
  - Practical 1
    - SQL exercises
      - PostgreSQL
      - Creating Tables
      - Inserting Data

Spatial Databases

- Week 4
  - Lecture 4 - More SQL
    - Queries
    - Join Queries
  - Practical 2
    - SQL – queries and joins

Spatial Databases

- Week 5
  - Lecture 5 - Spatial Databases
  - Practical 3
    - Handling Spatial Data in PostgreSQL/PostGIS
    - Viewing Spatial Data in QGIS

Spatial Databases

- Week 6
  - Querying Spatial Data
  - Working with 3D Spatial Data
  - Practical 4
    - Querying Spatial Data
    - Creating and querying 3D data
    - Viewing Spatial Data in FME

Spatial Databases

- Weeks 7-10 – Advanced Topics (depending on time)
  - Emerging topics in 3D spatial data
  - Improving database performance
  - Spatial Data Management
  - NoSQL
  - Blockchain

Spatial Databases

- Practical Sessions
  - We are using free and open source software (FOSS), for example:
    - Quantum GIS
    - PostgreSQL/PostGIS
  - Also – FME (licensed at UCL)
    - You need to be on the UCL VPN to access FME: https://vpn.ucl.ac.uk

Spatial Databases

- Practical Sessions
  - All the software can be downloaded onto your computer or accessed via the UCL Desktop Anywhere system
    - ** may be some version issues **
  - This means:
    - You can do the practical sessions in the allocated cluster room or download the software yourselves and work elsewhere

---

But ..

4.3 Attendance Requirements

1. UCL's minimum attendance requirement is 70% of teaching and learning events. Departments may stipulate a higher percentage and additional requirements where appropriate.

2. Students whose attendance falls below the attendance requirements are ineligible for summative assessment.

3. Undergraduate students must meet UCL's minimum attendance requirements during term-time (designated teaching and examination weeks). Outside of term-time, there is no requirement for students to be present on campus or to apply for authorised absence.

4. Taught Postgraduate students must meet UCL's minimum attendance requirements during term-time (designated teaching and examination weeks). During the summer vacation, students are expected to be studying for their dissertation and to be based at UCL or, if not, to apply for Study Away from UCL. Departments may stipulate additional requirements where appropriate.

Spatial Databases

- I will be taking attendance during lectures
  - UCL rules state that you must be in class by 10 past the hour to count as present

- I will also check that you have completed the practical sessions
- If you are going to be absent for any reason let me know by e-mail

---

Spatial Databases

- It is up to you to keep up with the practicals - I will run regular checks on your database accounts and this may count as part of your 70% required attendance
  - If the practical is incomplete at the end of 2 weeks after the date, you will be marked as absent for that session

---

Spatial Databases

- If you do work elsewhere, make sure you keep up to date – you can post questions on Moodle
- A forum has been set up for each of the practical sessions
- NB: I will only answer questions about a specific practical up to 2 calendar weeks after the practical session is run.

Spatial Databases

- You are ineligible for assessment if you don't meet the 70% attendance requirement

UCL Advice on Recording Lectures - GDPR

- Due to the difficulty in guaranteeing that audio recordings will not contain personally identifiable data, it is safest to assume that all such recordings of spoken word will be subject to the Data Protection Act 1998. As a result, if this method of processing is used, then you must adopt an approach that ensures the security of the recordings and images together with any other identifiable data held at all times. Permission to audio record a lecture is usually granted prior to commencement of the session. Students with disabilities can usually obtain permission that overrules the need to obtain prior agreement, provided that everyone present is informed that a recording is being made.

- Any disclosure of the audio recordings should be treated in the same way as any written material in terms of the security and decisions made about possible disclosure. In most circumstances individuals who are responsible for arranging the recording, should ensure that consent is obtained in order to process the recording both fairly and lawfully.

UCL Advice on Recording Lectures

- So ..
  - If you want to record (audio or video) this lecture ask me first, so that I can ask the group

  - NEVER!! share the recordings – you'll be violating the general data protection regulations

Spatial Databases - Assignment

- There is **one assignment** required for this module
- The assignment builds on the work we do in class and in the lab
  - Submission Deadline: Monday 10th December at 12.00 midday

The assignment has been posted on Moodle and contains a week by week summary linking the tasks to the lectures.

Spatial Databases - Assignment

- Assignment will be introduced in class during the next few weeks
- However: all questions should be asked on Moodle so that everyone has the same information
- I do not organise a specific Q&A session
- I do give random hints during the lectures e.g. 'this is not required' 'this is a good example of what would be expected' 'pay attention here – this is something people sometimes get wrong'

How to pass this module

- Pay attention to detail
  - In the practical sessions
  - In the assignment
    - Make sure you submit EVERYTHING that is asked for
      - don't lose marks for reasons that could be avoided
    - There is a checklist to help you

- For this module, important updates and questions/answers happen on Moodle …
  - Check your e-mails!

http://phdcomics.com/comics.php?f=1790

How to fail this module

- Attend less than 70% of the lectures and practicals
  - Remember – more than 10 minutes late counts as absent
- You can't pass this module by doing all the assignment work during the last week – you need to work on the assignment every week!

https://scontent-lhr3-1.xx.fbcdn.net/v/t1.0-9/41368501_2279506795604438_1512630813852696576_n.jpg?_nc_cat=1&oh=8a2b70c9def49c7aaaf4d3ab0606df7d&oe=5BF02F17

Spatial Databases

- Overview
  - Introduction to the Module
  - Data and Databases
  - What is Spatial Data

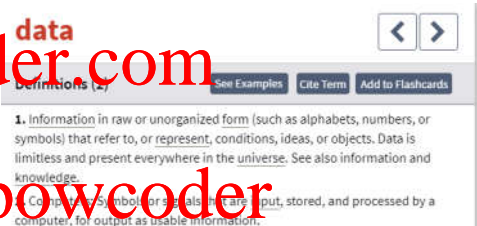## What is data?

https://dictionary.cambridge.org/dictionary/english/data

What is data?
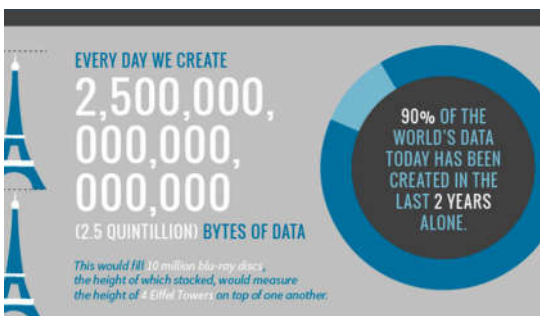

http://www.businessdictionary.com/definition/data.html

Examples of data

- A list of student names, contact details and addresses
- A list of assets held by a company
- A list of flights from an airport
- The level of noise in this room at various times of day
- Origins of the flights, and flight numbers into an airport
- Traffic counts along Gower Street
- Heart transplants in a hospital
- Road networks and speed limits
- Population counts in a district
- GPS track of where you walk every day
- And many, many, many more!

https://wonderfulengineering.com/this-is-the-amount-of-data-that-is-being-generated-in-the-world-every-minute/

How is data stored?

- On paper
- In a spreadsheet
- In a database management system (see next slides)
- In a PDF
- In a word document
- And many other formats ..

Database Fundamentals

- What is a database?
  - "A database is a collection of data used to represent information of interest to an information system" (Atzeni et al.)
  - Does not have to be computer-based

Database Fundamentals

- What is a DBMS?
  - DataBase Management System
  - A **software** system to manage collections of data
  - Tables, Primary Keys, Foreign Keys all stored in this software system

- NB: the terms database and DBMS are often swapped in practice – so when we say database we often mean DBMS

Database Fundamentals

- Characteristics of a DBMS (1)
  - Large
    - Millions of records
  - Shared
    - Millions of users
  - Persistent
    - Information not lost if computer is switched off
  - Reliable
    - Always gives same result for same question

Database Fundamentals

- Characteristics of a DBMS (2)
  - Efficient
    - No long wait for an answer
  - Secure
    - Multi-user access with varying privileges
    - Disaster recovery
    - Backup mechanism inbuilt
  - Minimises redundancy
    - Information only held once

Database Fundamentals

- Characteristics of a DBMS (3)
  - Self describing
    - You can find out what tables are in the database, what columns are in these tables, what data types the columns are using, the primary and foreign keys
  - Multiple views of the same data
    - Many programs can use the same data for different purposes
  - Has query language inbuilt
    - SQL (Structured Query Language)

Database Fundamentals

- Data Independence
  - Allows users to interact with the database without knowing how the database is physically structured
  - Equally, the data on disk can be moved to another location without the end user being aware

Database Fundamentals

- DBMS vs. File System (e.g. excel, word, pdf)
  - Central backup and recovery
  - Secure access, local and remote
  - File storage in controlled area, no files on users disks
  - Hides complexity of file system from users
  - Applies central rules base – constraints
  - Inbuilt query and reporting tools
  - Difficult to scale a single file
  - Data of interest to one piece of software may be replicated (perhaps in a different file format) for another application. Difficulties in maintaining copies consistent, current etc.

Advantages of a DBMS

- Co-operative approach to database design
  - Involvement of end users essential in designing database - therefore need an accessible framework
- Relational model very easy to understand
  - approachable and unimposing
  - simply a collection of tables - everybody understands tables
- The Structured Query Language (SQL - sequel) a language that (almost) anyone can understand
  - 'English like'
  - Non-procedural - specify what data you want rather than how to retrieve it
- Standard, consistent model of real world application

Advantages of a DBMS

- Data common resource, available to authorised members
- Central control = economy of scale
- Data independence favours flexible applications
- Data sharing implies less redundancy and improved quality
- Supports multiple users concurrently. Each user can benefit immediately from others' changes
- Different types of data in one central location which can be queried through one interface

Disadvantages of a DBMS

- Expensive and complex, training required, resource hungry
- May include extra services that are not required
- Not suitable for simple, single-user applications
- Operating costs may be high

Spatial Databases

- Exercise
  - We have the definition of a database …
  - The BBC quotes think-tank Demos as stating that:
    - "The average economically active individual in the developed world is on about 700 databases"
      - http://news.bbc.co.uk/1/hi/technology/7107975.stm
  - So, think about 5 places where you use a database in every day life …

Spatial Databases

⚐UCL

- Overview
  - Introduction to the Module
  - Data and Databases
  - **What is Spatial Data**

Spatial Data

⚐UCL

- CEGE0052 is a *spatial* databases module

- In this case *spatial* refers to any data that can be located somewhere on the earth's surface (or above or below the surface)
  - "data that can be mapped"

Everything Happens Somewhere

⚐UCL

http://www.iotphils.com/wp-content/uploads/2014/07/Smart-Cities.png

Spatial Data

⚐UCL

- Used to understanding WHERE and WHEN things happen
  - And thus solve problems and provide useful services to people

http://www.directionsmag.com/images/newsletter/2011/01_week1/world_lg.jpg

Spatial Data is essential for:

⚐UCL

Smart Cities and the Internet of Things

Understanding and analysing Big Data

Spatial Data is essential for:

⚐UCL

Providing Services

⛌UCL

Spatial data is essential for:

Helping when disasters strike

⛌UCL

Spatial Data is essential for:

16%   28%   64%

Improving the Environment

Making the world more sustainable

Sustainable Community

⛌UCL

Spatial Data is essential for:

Routing and Navigation

Understanding, Predicting and Preventing Crime

HotSpot

⛌UCL

Spatial data is essential for:

Planning for the Future

⛌UCL

Asset Management – Built Asset Lifecycle

Plan
Design
Identify Need
Procure/Build
Asset Life Cycle (Value)
Decommission (Replace or Dispose)
Commission
Modify or Upgrade
Operate, Maintain, & Monitor Performance

⛌UCL

Spatial Data is essential for:
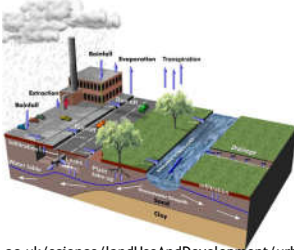## Planning (site selection)

http://www.cse.org.uk/projects/view/1164

Spatial Data is essential for:
Planning (site selection)

http://www.bgs.ac.uk/science/landUseAndDevelopment/urban_geoscience/SUDS/information.html

Spatial Data is essential for:
Design (what will I see)

http://www.ebay.co.uk/bhp/window-view-wall-stickers

Spatial Data is essential for:
Design (who will see me)

http://www.skyscrapercity.com/showthread.php?t=1067687

Spatial Data is essential for:
Design (do I need noise mitigation)

http://www.fhwa.dot.gov/environment/noise/noise_compatible_planning/federal_approach/land_use/qz06.cfm
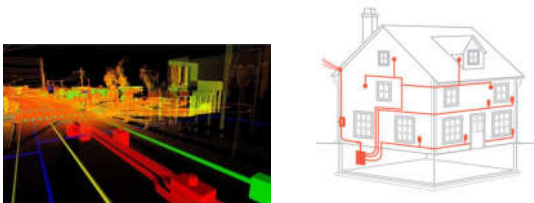

Spatial Data is essential for:
Procure and Build

http://www.ajgroupinternational.com/construction-2/construction-material/

Spatial Data is essential for:
commissioning

http://www.topfliteloftconversions.co.uk/planning-a-loft-conversion-let-your-neighbours-know/

Slide 1

UCL

Spatial Data is essential for:
commissioning

http://geospatial.blogs.com/geospatial/digital_cities/page/2/
http://www.thisoldhouse.com/toh/article/0,,1206502,00.html

Slide 2

UCL

Spatial Data is essential for:
Operation and Maintenance

http://forums.digitalspy.co.uk/showthread.php?t=1761849
http://www.allaboutmoney.com/debt-advice/news/incidents-of-council-tax-debt-up-by-27-pc-0-5649.htm

Slide 3

UCL

Spatial Data is essential for:
Operation and Maintenance

http://www.wifarer.com/hospitals

Slide 4

UCL

Spatial Data is essential for:
Decommissioning

- http://www.findlocales.com/bangalore/bcontacts/all/details/1248/an-yarab-traders.html

Slide 5

UCL

Spatial data is essential for:

Asset
Management –
Locating Assets

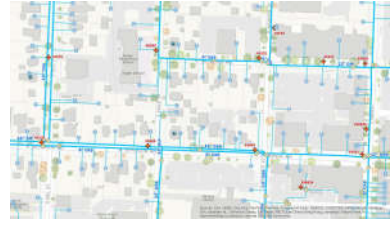Map of Britain's defibrillators 'could save thousands of lives'

https://home.bt.com/news/science-news/map-of-britains-defibrillators-could-save-thousands-of-lives-11364288032948

Slide 6

UCL

Spatial data is essential for:

Asset
Management –
Outage
Management

https://www.saimgs.com/imglib/other_pages/cmms/esri-water-distribution-map.jpg

Slide 1

Spatial data is essential for:

Asset Management – Maintenance and Safe Working

https://msl-ltd.co.uk/wp-content/uploads/2016/06/WAH.png
https://bestrane.com.au/wp-content/uploads/2014/11/bestrane_product_detail_847657.jpg

Slide 2

Spatial Data

- Spatial data includes **anything** that can be modelled using some form of *location* information!
  - i.e. where something is, referenced to a shared framework (could be a coordinate system, a map of London Boroughs, countries of the world, UK counties and many more)
- This referencing is called *geo-referencing*

Slide 3

Spatial Data

- The ability to create "maps" using spatial data can be found in:
  - Geographical Information Systems (GIS)
    - E.g. QGIS, ArcMap
  - Building Information Modelling (BIM)
    - E.g. Revit, Bentley Architecture, ArchiCAD
- In both cases, the maps can be 2D or 3D

Slide 4

Spatial Data

- We will be using GIS for mapping during this module as GIS software currently works best with databases –
  - GIS are also extensively used in Asset Management
  - ... BIM software is slowly becoming better at working with databases but isn't quite there yet ..

Slide 5

Geo-Referencing

- Can be direct:
  - E.g. a map that shows a building or another object, x/y coordinates, GPS coordinates
  - Using a coordinate reference system
- Or indirect
  - For example, a Post Code or a Street Address is an indirect geo-reference that can be used to link non-spatial data to a position on the map. A PDF file containing the specification of a water pipe can be linked to the location of that pipe.
- See later on in the module for more details about georeferencing

Slide 6

Spatial Data Formats

- In this module we will see how to store data (spatial and non-spatial) in a database
- However, you may also find some other formats for spatial data, including
  - dwg (Autodesk CAD)
  - rvt (Autodesk Revit)
  - shp (Esri ArcGIS 'shape')
  - tab (Mapinfo Professional)
  - dgn (Bentley Microstation)
  - kml (Google Maps and Google Earth)
  - geoJSON (geoJavascriptObjectNotation - usually used in web mapping but now increasingly common elsewhere)
  - gml (geographic markup language – data sharing standard set by the Open Geospatial Consortium, who are the ISO equivalent for geospatial data)

Spatial vs Geographic

- Geographic Information Systems (GIS) work with spatial data
  - Usually data used in a GIS has an official national or global coordinate reference system
  - Although they can handle data without an official coordinate reference system this is not what they were designed to do

---

The word spatial originated from Latin 'spatium', which means space. Spatial means 'pertaining to space' or 'having to do with space, relating to space and the position, size, shape, etc.' (Oxford Dictionary), which refers to features or phenomena distributed in three-dimensional space (any space, not only the Earth's surface) and, thus, having physical, measurable dimensions. In GIS, 'spatial' is also referred to as 'based on location on map'.

Geographic(al) means 'pertaining to geography (the study of the surface of the earth)' and 'referring to or characteristic of a certain locality, especially in reference to its location in relation to other places' (Macquarie Dictionary). Spatial has broader meaning, encompassing the term geographic. Geographic data can be defined as a class of spatial data in which the frame is the surface and/or near-surface of the Earth. 'Geographic' is the right word for graphic presentation (e.g., maps) of features and phenomena on or near the Earth's surface. Geographic data uses different feature types (raster, points, lines, or polygons) to uniquely identify the location and/or the geographical boundaries of spatial (location based) entities that exist on the earth surface. Geographic data are a significant subset of spatial data, although the terms geographic, spatial, and geospatial are often used interchangeably.

Geospatial is another word, and might have originated in the industry to make the things differentiate from geography. Though this word is becoming popular, it has not been defined in any of the standard dictionary yet. Since 'geo' is from Greek 'gaya' meaning Earth, geospatial thus means earth-space. NASA says 'geospatial means the distribution of something in a geographic sense; it refers to entities that can be located by some co-ordinate system'. Geospatial data is to develop information about features, objects, and classes on Earth's surface and/or near Earth's surface. Geospatial is that type of spatial data which is related to the Earth, but the terms spatial and geospatial are often used interchangeably. United States Geological Survey (USGS) says "the terms spatial and geospatial are equivalent".

- https://gis.stackexchange.com/questions/34733/spatial-data-geodata-geographic-data-geospatial-data
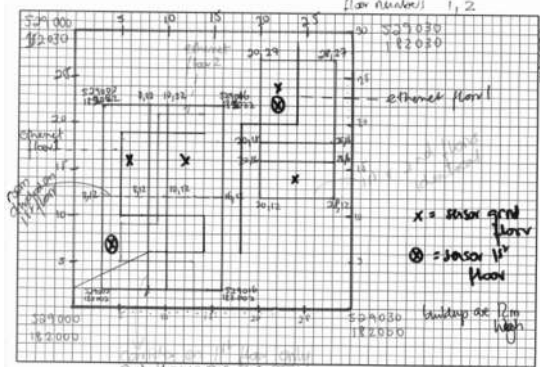- From 2012

Spatial vs Geographic

- For this module, work on the basis that:
  - All geographic data is spatial
  - Not all spatial data is geographic
    - Geographic – data relates to a location on the Earth's surface
    - i.e. data in a local reference system is not geographic as you don't know where on the Earth's surface it relates to

---

At Home Exercise - 1

- Read the 'UCL Facilities Management' briefing document
  - We will use this example next week so make sure you are familiar with it
  - This is also an example of the sort of system specification I'm expecting for your assignment (more about the assignment in a later lecture)

At Home Exercise - 2

- Draw a 2D sketch map (plan) of two adjacent room at UCL or in your home, using a local reference system
  - What objects (assets) will you include on the map?
- If you don't have graph paper, you can print some out e.g. from https://incompetech.com/graphpaper/

Spatial Databases

- Overview
  - Introduction to the Module
  - Data and Databases
  - What is Spatial Data

Further Reading

- If you're new to spatial data and GIS, a good book to read is: **Geographic Information Science and Systems** Paul A. Longley, Michael F. Goodchild, David J. Maguire, David W. Rhind

Further Reading

- "Everything happens somewhere"... test your understanding of spatial data by completing the 'What is spatial data' quiz on Moodle

https://moodle-1819.ucl.ac.uk/mod/quiz/view.php?id=711009

Further Reading – GIS and Asset Management

- https://www.esri.com/library/whitepapers/pdfs/a-practical-guide-to-gis-in-asset-management.pdf
- https://www.esri.com/library/whitepapers/pdfs/enterprise-gis-utilities-performance.pdf
  - From the perspective of the leading GIS vendor

Further Reading – GIS and Asset Management

- https://theiam.org/knowledge/Knowledge-Base/the-anatomy/
  - From the perspective of Asset Management standards
  - GIS mentioned in 6.4.2 and 6.4.4 - note that PAS 55 has been recently superseded by ISO 55000 but the principles described here still apply

Further Reading – GIS and Asset Management

- https://www.novarageo.com/industries/utilities/

Spatial Databases Video

- Directions magazine video: http://www.directionsmag.com/webinars/register/discover-the-secret-of-scaleable-extensible-and-secure-geodata-managem/458927
  - (NB: does not play in Chrome, use Edge or another browser)
- (with Hans Viehmann)

GIS applications

- A useful list of applications that GIS and spatial data can be used for can be found here:
  - https://grindgis.com/blog/gis-applications-uses

Where to find spatial data

Ordnance Survey is the National Mapping Agency of Great Britain

- Mastermap Topographic Data - map of GB – includes buildings, roads (as polygons), railway, water, green space etc
- Mastermap ITN intelligent transport network – road network – as lines

Where to find spatial data

- Mastermap data is available free to students:
  - Edina Digimap Service - http://www.edina.ac.uk/ (lots of other spatial data there too)
- You can also find other Ordnance Survey spatial data here: https://www.ordnancesurvey.co.uk/opendatadownload/products.html

Where to find spatial data

- London Data Store -https://data.london.gov.uk/
- https://data.gov.uk/search?filters%5Btopic%5D=Mapping (UK only, lots of local and central government data)
- https://www.ons.gov.uk/methodology/geography/ukgeographies (UK administrative boundaries)
- http://download.geofabrik.de/ (worldwide map data, captured by the crowd)