

Guest Presenter Prof. Gupta Topic Spatial Data Processing

Write down a major point or a question with the answer for each 15 minute interval of the guest presentation. Use the table below to record the time you make the note

Interval	Time	NOTE/Question	Answer, if a question
9:30-9:45	9:37	Geoparsing: what geoparsing entails & geolocation, Becons)	Yes. All are Geoparsing
9:45-10:00	9:51 9:55	(latitude & longitude for full location : same efficiency with data structure.	
10:00-10:15	10:06 10:12	Blockchain: calculate hash mining (contributes to transaction efficiency point made for GPU?) R-Trees - Store Coordinates.	Clustering Algorithms - Group similar things together GeoJSON - ID
10:15-10:30	10:15 10:22	JSON spatial data. databases: PostgreSQL (PostGIS) or MySQL (mmDB) or MongoDB Shape files -	Shape files MongoDB or JSON Objects ArcGIS, QGIS
10:30-10:45	10:35	AI/ML Spatial data processing DB (charts visualizing data) Geopandas	Dataframes, Karray

Guest Presenter Dr Griffin Topic Spatial Data Intro

Write down a major point or a question with the answer for each 15 minute interval of the guest presentation. Use the table below to record the time you make the note

Interval	Time	NOTE/Question	Answer, if a question
9:30-9:45	9:38	Geo fencing, types of data available	What is big data ① What is offsite ②
9:45-10:00	9:48 9:50	Data structures to store spatial data ④ ③ 9:55 ② 9:58	
10:00-10:15	10:06	① ⑤ 10:08 ⑥ ⑦ 10:12 ⑧ 10:13 ⑨	What is the difference between a tree and a graph ③
10:15-10:30	10:21	⑩ ⑪ ⑫ 10:26	
10:30-10:45	10:36 10:44	⑬ ⑭ went through some data Visualization	How much data processing is involved in AWS ⑬

① Big data is data that starts spilling out of a computer system

② Obfuscate: makes code less human readable. Make it more difficult for humans to get the full picture

④ Finding different ways to store spacial data

⑤ RTrees, ^{quad} trees, KD trees

How you would store spacial data to look up fast

⑦ Block chain: adding hashes to a chain that follows a coin persistently

① R trees: one of the first ways to store 2d coordinates and have a viable way to look them up

look up: binary trees
n tree

⑤ Quad Trees: Good for Collision detection

⑥ KD tree: k dimensional trees

③ A tree has no cycle but a graph does

- ③ DBScan: clustering algorithm.
K mean
Groups ~~non~~
non-spatial data.
Visualizing tool

④ Data Formats

CSV, JSON, GeoJSON, data bases,
shape files

- ① GeoJSON: ^{↓ super set of JSON} storing geo location
data in JSON

JSON → key value pairs

- ② cs.msu.texas.edu/~griffin
for data that can be
processed

⑤ Databases:

→ PostgreSQL (PostGIS) that can
handle spatial data

→ MySQL (MariaDB): relational
Databases

→ MongoDB: stores JSON directly

- ③ Shape files: a file type
that stores 2d data

→ QGIS is ~~an~~ another one that
is free

- ④ Depends on ~~what~~ what ~~you~~
you want. They have several
options like SQL, etc.

AWS is similar to digital
Ocean

⑥ Tools:

→ ArcGIS (so not free)

→ QGIS (Free!)

→ Plotly

→ ~~Mapbox~~ Mapbox (Plotly)

→ Xarray

→ Geobinder

① 03 javascript

* Custom charts
* Data visualization

② xarray 3D ~~arrays~~ arrays

Guest Presenter Dr. Grith Topic Spatial Data Intro

Write down a major point or a question with the answer for each 15 minute interval of the guest presentation. Use the table below to record the time you make the note

Interval	Time	NOTE/Question	Answer, if a question
9:30-9:45	9:39 9:44	What is geofencing? Big data is a large data set over time. Another definition is when data starts spilling out of a computer but there is not enough room.	The distance between two x,y points... alerts you if you are in a certain range of something
9:45-10:00	9:46	Obfuscating: takes all javascript data, readable functions Companies like Airbnb do not give you exact location & they obfuscate. What apps need spatial data? What are some data structures to store spatial data?	Graphs, R Trees, Quad Trees Binning trees, KD tree, DBScan, Kmeans Games need spatial data structures for collision detection
10:00-10:15	10:00	Blockchain makes crypto currency secure everytime a trade happens two parties create a receipt, calculate the hash and ensures that is correct. The hash created is added to a coin and everybody knows. It is hard to steal and is verifiable.	R trees are one of the first data structures to store two dimensional x,y coordinates - can be extended to three dimensional - can store lines and polygons
10:15-10:30	10:16 10:23	Spatial data is well formed and very visual. Databases provides fast access to store and retrieve things. What are some popular databases?	PostGIS allows you to store spatial data. MongoDB - stores JSON really well MySQL (Maria DB) & PostGIS are free and available relational databases
10:30-10:45	10:36	What are some good library tools for spatial data? Start learning to write data driven charts all about visualizing data. They are interactive and clear.	plotly: charting for python pyplot geopandas & array

Guest Presenter Dr. Griffin Topic Special Data

Write down a major point or a question with the answer for each 15 minute interval of the guest presentation. Use the table below to record the time you make the note

Interval	Time	NOTE/Question	Answer, if a question
9:30-9:45	9:40	What is big data?	Large data sets over a terabyte. Data is considered large after it starts spilling out of one computer.
9:45-10:00	10:05	Data/Information is concealed and code transforms to something. It's readable. This is obfuscation.	
10:00-10:15	10:15	B-trees are first data structure to store 2D coordinates. Quad trees are good for collision detection.	
10:15-10:30	10:20	Geo json is an ideal way to store and represent spatial data. If data is structured, it can be presented in a way that is visually appealing.	
10:30-10:45	10:37	DB clients are concerned with visualizing data. I should learn to use data frames just in case.	

Guest Presenter Griffin Topic Data Structures

Write down a major point or a question with the answer for each 15 minute interval of the guest presentation. Use the table below to record the time you make the note

Interval	Time	NOTE/Question	Answer, if a question
9:30-9:45	9:41	What is big data considered as?	Data that is so large or complex that traditional methods are nearly impossible to process them.
9:45-10:00	9:55	QuadTrees and RTrees are what are typically used for any spacial data or storing points because they are efficient	
10:00-10:15	10:10	Trees and graphs are both similar but trees have a root and do <u>NOT</u> have cycles, which is possible within a graph.	
10:15-10:30	10:29	What is the difference/benefits between csv and json.	csv files are smaller and easier to use while json formats and holds data better/more for various/different needs
10:30-10:45	10:44	Geoson/Geojson is a form of json which holds spacial data (lat, long I believe) and holds locations almost like needles on a map. Also github will directly display the file if uploaded.	

Guest Presenter Dr. Griffin Topic Spatial Data

Write down a major point or a question with the answer for each 15 minute interval of the guest presentation. Use the table below to record the time you make the note

Interval	Time	NOTE/Question	Answer, if a question
9:30-9:45	9:39	Geo fencing example: going to the gym and your phone automatically opening the gym app.	
9:45-10:00	9:53	Without spacial data games will basically be out the door.	
10:00-10:15	10:09	R Trees were one of the first to store 2D coordinates.	
10:15-10:30	10:20	GeoJSON is a structured way to store data. Geojson files show in github (pretty cool)	
10:30-10:45	10:40	Pretty interesting topic "Spatial Data"	

Guest Presenter Dr. Griffin Topic Spatial Data

Write down a major point or a question with the answer for each 15 minute interval of the guest presentation. Use the table below to record the time you make the note

Interval	Time	NOTE/Question	Answer, if a question
9:30-9:45	9:45	What is obfuscation?	Scrambles and restructures code.
9:45-10:00	9:58	Blockchain in crypto, what's its purpose?	Track and verify data.
10:00-10:15	10:05	R-trees, One of the first data structures to store 2D coordinates.	
10:15-10:30	10:30	Databases reinforce data integrity and the ability to rollback.	
10:30-10:45	10:38	Data visualization -D3.js	

Guest Presenter Dr. Griffin Topic Partial Data

Write down a major point or a question with the answer for each 15 minute interval of the guest presentation. Use the table below to record the time you make the note

Interval	Time	NOTE/Question	Answer, if a question
9:30-9:45		Find what geographic thinking	Geographical data has two useful traits. First geographic data is ubiquitous. Everything has a location in space-time and this location can be used directly to make better predictions or inferences.
9:45-10:00		explain Geographic Process	Geography Process are represented using <u>objects</u> , <u>fields</u> and <u>networks</u> . objects are discrete entities that occupy a specific position in space and time.
10:00-10:15		what Data Structures & Algos use?	graphs RTree Quad trees BTrees KD Tree
10:15-10:30		Data formats (what data formats?)	The following various file formats in which spatial data is typically csv, json, geojson, database (mongoDB, MySQLDB)
10:30-10:45		Tools	ArcGIS (this is not free tool) QGIS (free) Plotly Mapbox Leaflet Xarray Geopandas

use www.bboxfinder.com

Guest Presenter Griffin Topic Spatial Data Intro

Write down a major point or a question with the answer for each 15 minute interval of the guest presentation. Use the table below to record the time you make the note

Interval	Time	NOTE/Question	Answer, if a question
9:30-9:45	9:38	What are the useful traits of Geographical Data?	Geographic data is ubiquitous - everything has location in space and time - can be used to make better predictions
9:45-10:00	9:46	What is obfuscating in coding?	Taking all readability and erases it and renames it and crams it all together much like obfuscated javascript behind a website
10:00-10:15	10:00	* RTree First implemented Data Struct What types of Structures used to implement spatial data?	- these implemented in Structures like Rtrees, Btrees, QuadTrees, KDtrees etc - without spatial data, then things such as games and bitcoin would not be a thing.
10:15-10:30	10:17	What does clustering algorithms do for us? - Data formats for spatial data → CSV, JSON, GeoJSON	- Clustering can take non coordinate data and project so can visualize and see items b/c close together
10:30-10:45	10:38	Visualizing data → ID3 ← javascript Data Driven Charts (helps visualizing Data C D3JS)	Plotly in python Map box Layers Charting library for python that webbased integrates

→ all these use datastructure called "dataFrame"

→ xarray ← 3D dataframes

→ what is data shaving? → example Plotly with time, x, y coordinate and visualizes data

Guest Presenter Dr. Griffin Topic Spatial Data Processing

Write down a major point or a question with the answer for each 15 minute interval of the guest presentation. Use the table below to record the time you make the note

Interval	Time	NOTE/Question	Answer, if a question
9:30-9:45	✓ 9:35 ✓ 9:40 ✓ 9:44	What is spatial data processing? What is geo fencing? What is big data?	Manipulating data that has a direct or indirect reference to a specific location or geographical area. The use of GPS or RFID tech. to create a virtual geographic boundary, enabling software to trigger a response based on location.
9:45-10:00	✓ 9:50 ✓ 9:53	What is obfuscation? Always have a burner phone.	Extremely large data sets that maybe analyzed computationally to reveal patterns, trends and associations, especially relating to human behavior and interactions.
10:00-10:15	✓ 10:00	R-Trees, one of the first data structures to store geo spatial points. (Stores them through boundary boxes.)	The action of making something obscure, unclear, or unintelligible. (Done to code so it is not easily interpretable.)
10:15-10:30	10:17	Quad Trees: good for collision detection. KD Trees: you can set how many dimensions you want to store.	
10:30-10:45	10:31 10:40	Data formats: CSV, JSON, geojson, databases, Shape files. Chart visualizers are awesome!	

Guest Presenter: Dr. Griffin Topic: Spatial Data Processing

Write down a major point or a question with the answer for each 15 minute interval of the guest presentation. Use the table below to record the time you make the note.

1) Time: (9:35)

Question/Note: What is spatial data processing?

Answer, if a question: Manipulating data that has a direct or indirect reference to a specific location or geographical area.

2) Time(9:40)

Question/Note: What is geofencing?

Answer, if a question: The use of GPS or RFID technology to create a virtual geographic boundary, enabling software to trigger a response based on location.

3) Time(9:44)

Question/Note: What is big data?

Answer, if a question: Extremely large datasets that may be analyzed computationally to reveal patterns, trends and associations, especially relating to human behaviour and interactions.

4) Time(9:50)

Question/Note: What is obfuscation?

Answer, if a question: The action of making something obscure, unclear, or unintelligible. (Done to code so it is not easily interpretable)

5) Time(9:53)

Question/Note: Always have a burner phone.

Answer, if a question: N/A

6) Time(10:00)

Question/Note:

R-Trees, one of the first data structures to store geospatial points; stores them through boundary boxes. R-trees are tree data structures used for spatial access methods, i.e., for indexing multi-dimensional information such as geographical coordinates, rectangles or polygons. The R-tree was proposed by Antonin Guttman in 1984 and has found significant use in both theoretical and applied contexts.

A common real-world usage for an R-tree might be to store spatial objects such as restaurant locations or the polygons that typical maps are made of: streets, buildings, outlines of lakes, coastlines, etc. and then find answers quickly to queries such as "Find all museums within 2 km of my current location", "retrieve all road segments within 2 km of my location" (to

display them in a navigation system) or "find the nearest gas station" (although not taking roads into account). The R-tree can also accelerate nearest neighbor search for various distance metrics, including great-circle distance.)

Answer, if a question: N/A

7) Time(10:17)

Question/Note:

Good for collision detection.

Quad Trees: Quadtrees are trees implemented to efficiently store data of points on a two-dimensional space. In this tree, each node has a maximum of four children.

We can build a quadtree from a two-dimensional area implementing the following steps

- The current two dimensional space is divided into four boxes.
- If a box consists of one or more points in it, build a child object, storing in it the two dimensional space of the box.
- If a box does not contain any points, do not build a child for it.
- Perform recursion for each of the children.

Quadtrees are implemented in image compression, where each node consists of the average colour of each of its children.

The deeper we visit in the tree, the more the detail of the image.

Quadtrees are also implemented in searching for nodes in a two-dimensional area. For instance, if we wanted to compute the closest point to given coordinates, we can do it by implementing quadtrees.

Answer, if a question: N/A

8) Time(10:17)

Question/Note:

KD Trees:In computer science, a k-d tree (short for k-dimensional tree) is a space-partitioning data structure for organizing points in a k-dimensional space. k-d trees are a useful data structure for several applications, such as searches involving a multidimensional search key (e.g. range searches and nearest neighbor searches) and creating point clouds. k-d trees are a special case of binary space partitioning trees.

Answer, if a question: N/A

9) Time(10:31)

Question/Note:

Data formats: csv, json, geojson, databases, shape files

Answer, if a question: N/A

10) Time(10:40)

Question/Note: Chart Visualizers are awesome! (e.g D3.js it is a javascript chart visualizer)

Guest Presenter T6 Topic Spacial Data

Write down a major point or a question with the answer for each 15 minute interval of the guest presentation. Use the table below to record the time you make the note

Interval	Time	NOTE/Question	Answer, if a question
9:30-9:45	9:34	what is big data	when data is too large for single computers to handle
9:45-10:00	9:54	What is a block chain	List of transactions
10:00-10:15	10:06	R trees can store pairs	
10:15-10:30	10:15	Geojson is a superset of json, can store spacial data	
10:30-10:45	10:36	DD3 can visualize data from your documents	

Guest Presenter Griffin Topic Spatial Data

Write down a major point or a question with the answer for each 15 minute interval of the guest presentation. Use the table below to record the time you make the note

Interval	Time	NOTE/Question	Answer, if a question
9:30-9:45		Geo-Fencing: crossing a certain location & getting notified of a person's location	
9:45-10:00		Data Structures & Algorithms for spatial data: Graphs, RTrees, QuadTrees, KDTree, DBSCAN, Kmeans. QuadTrees store spatial data for collision detection	
10:00-10:15		File Formats in which spatial data is found: CSV, JSON, GeoJSON, PostGIS, & Shape Files	
10:15-10:30		RTrees were the first data structure to store coordinates. Put the coordinates in a bounding box & place it in an RTree	
10:30-10:45		Tools to use for processing spatial data are: ArcGIS, QGIS, Plotly, Mapbox Layers, Xarray, & Geopandas	

Guest Presenter Dr. Griffin Topic Spatial Data Processing

Write down a major point or a question with the answer for each 15 minute interval of the guest presentation. Use the table below to record the time you make the note

Interval	Time	NOTE/Question	Answer, if a question
9:30-9:45	9:40	Geographic data has two useful traits: ① Everything has a location ② it is ubiquitous	
9:45-10:00	9:46	An obfuscator turns readable function names, renames everything, sets id of while and shoves everything together to	make it obscure
10:00-10:15	10:08	R trees are first data structure to store coordinates.	
10:15-10:30	10:28	Databases are fast, standard places to hold data that maintain data integrity	
10:30-10:45	10:39	Data Frames are crucial for data visualization. Xarray, plotly, geopandas, D3,	