14-848 Cloud Infrastructure

SPRING 2023

MOHAMED FARAG

FARAG@CMU.EDU

Agenda

- Welcome and Introductions
- Focus Areas of this Course
- Course Syllabus & Course Communication
- Expectations
- TA Introductions
- Next Steps

Why is this course Important?

In recent years, the demand for Cloud computing jumped significantly.

The most in-demand hard skills of 2021 (according to the New York Institute of Finance) emphasizes on the knowledge of cloud computing:

- Business Analysis
- Analytical Reasoning
- Affiliate Marketing
- Sales
- **Cloud Computing**

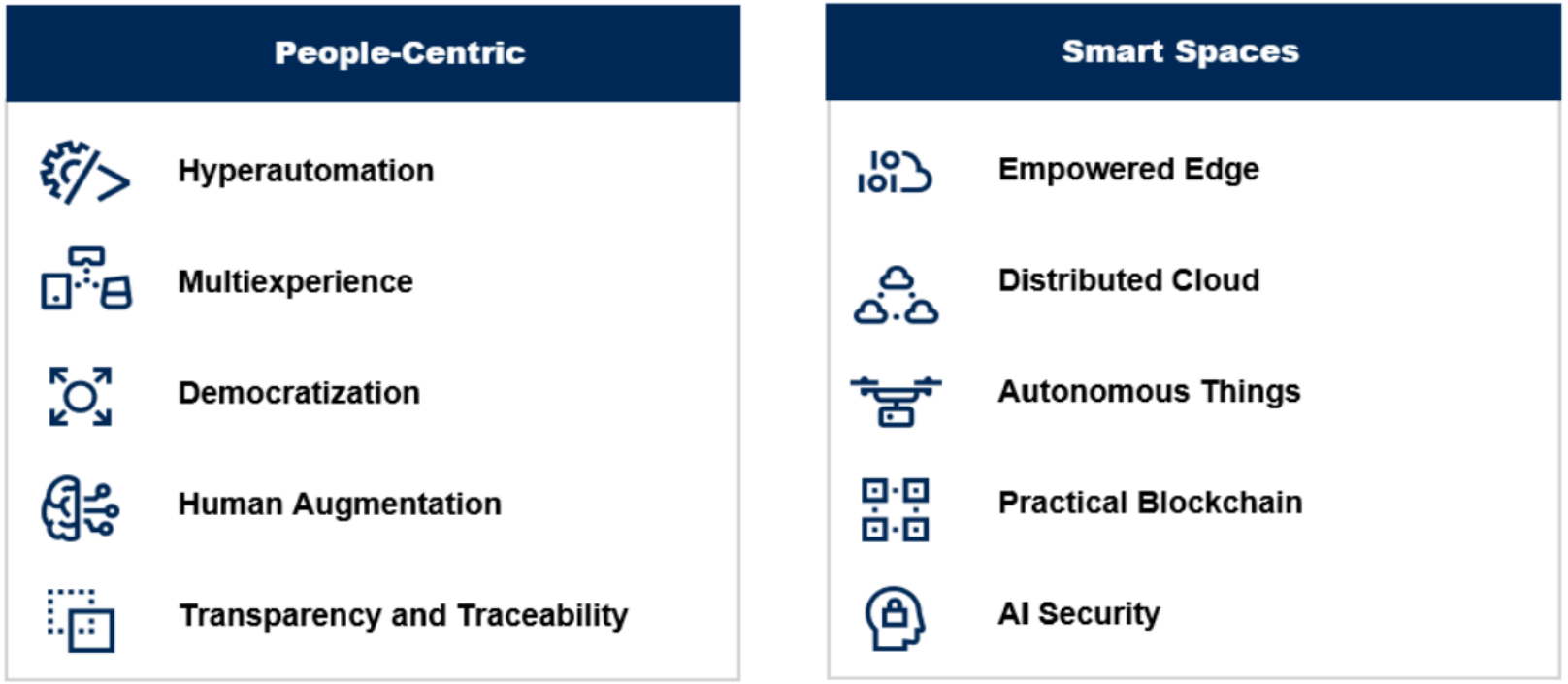
A recent article by Coursera, an e-education platform, lists the essential IT skills in 2021 as the following:

- Security
- Programming
- Systems and networks
- Data analysis
- **DevOps**
- **Cloud Computing**
- Machine Learning

Why is this course Important? – Cont'd

And the demand is expected to grow!

Notice the Top strategic Technology trends for 2020



Source: Gartner (October 2019)
ID: 432920

Focus Areas of the Course

Cloud Technologies

Cloud Technologies from the following vendors:

- Amazon Web Services
- Google Cloud Platform
- Microsoft Azure

Big Data Processing Platforms

- Apache Hadoop
- Apache Spark

NoSQL Database

- DynamoDB

DevOps

Deployment Scaling and Orchestration:

- Docker
- Kubernetes

Cloud Infrastructure

- Metaverse
- Data Centers
- Edge Computing & Fog Computing

Why Do We Care about Cloud Computing?

Expectations for Incoming Students

- ***You are expected to know Python or are willing to learn it.***
    - If you need assistance with Python, a tutorials session is offered on September 8th. More details to be released.
- ***You are expected to have a basic understanding of Computer Virtualization.***
    - If you don't know about virtualization, check out this reading: https://www.vembu.com/blog/physical-server-vs-virtual-machine-choice-open/

Logistics

- We meet on **Tuesday/Thursday 03:30 - 04:50PM ET/12:30 - 01:50PM PT** for lectures. Lectures are offered in-person only
- Instructor's Office hours are conducted via Zoom on **Wedbesday 11:30AM - 1PM ET/8:30-10AM PT**
- TAs' Office hours to be announced soon
- Lecture slides are delivered via TopHat during the lecture. Sign up for a free account and join the course with the following code: **168828**
- For some assignment, you may need to run Python on Jupyter Notebooks.
- Students who have approved accommodation shall contact the course instructor to figure out how the instructor can meet their needs

Logistics – Cont'd

- Use Course Piazza to ask asynchronous questions that require instructor and/or TA help
- Use Zoom Office Hours with the instructors and TAs to get 1-1 help
- Use the Student Space Slack Channel to communicate among yourselves (No instructor or TA help is offered there)
- **This course doesn't offer live remote attendance via Zoom**
- **If you have to miss the lecture, email me and I'll send you a copy of the lecture recording after it becomes available**

Course Assessment

Final Exam	Project	Assignments	Quizzes
15%	20%	40%	25%

- **Final Exam: open-notes. Exam will be offered via Lockdown Browser and no knowledge exchange is allowed among students during the exam. Exam is offered during the last lecture (on April 27th).**
- **Course Project:** students are expected to deliver one course project. This project leverages most of the topics and practices that are covered throughout the semester. Course details are released in Week-4. You can select 1 peer for the project.
- **Homework Assignments:** there will be 8 homework assignments provided throughout the semester covering the practical aspects of the class. There will be good learning curve that students will have to take on their own.
  - For 3 days after the deadline, you may submit the homework with a penalty of 20%. No submissions are accepted afterwards
- **Quizzes:** there will be 1 quiz published on Canvas after each lecture with a specific access code. The access code will be revealed during the lecture to the students who are attending.
  - Quizzes will start next lecture.
  - You will receive two excused absences from Quizzes for emergencies, sickness, etc.
  - If you need to attend remotely for extended time period, please request accommodation from the disability office.

Course Grade Scheme

+/- are used to provide granularity

Grade	Percentage Interval
A	[85-100%]
B	[70-85%)
C	[55-70%)
D	[40-55%)
R (F)	Below 40%

General Information about Communication

- All office hours will be conducted remotely
- We will have a mixture of Zoom and Piazza OHs

Course Syllabus

Date	Topic	Notes
Week-1	- Introduction & Syllabus - Introduction to Cloud	- GitHub homework released
Week-2	- Virtualization	- GitHub homework deadline. - Virtualization homework released
Week-3	- Containerization	- Virtualization homework deadline - Docker homework released
Week-4	- Deployment Orchestration - Pokémon Go Case Study	- Docker homework deadline. - Course project details are released. - Kubernetes homework is released

Course Syllabus – cont'd

Week-5	- Cloud Computing Concepts	- Kubernetes homework deadline
Week-6	- Cloud Computing Paradigms - Microservices lab	- Microservices homework is released
Week-7	- Cloud Data Storage	
Spring Break		

Course Syllabus – cont'd

Week-8	- Introduction to Hadoop - Hadoop HDFS	- Microservices homework deadline - NoSQL DB homework released.
Week-9	- Hadoop MapReduce	- NoSQL DB homework deadline. - Course Project Checkpoint
Week-10	- Hadoop MapReduce Cont'd - Big Data Algorithms	- MapReduce on GCP homework is released.
Week-11	- Introduction to Spark	- MapReduce on GCP homework deadline

Course Syllabus – cont'd

Week-12	Spark (Cont'd)	- Apache Spark homework released - Course project submission
Week-13	Cloud Security & Privacy	- Apache Spark homework deadline
Week-14	Edge Computing and Fog Computing Final Exam	

Course Delivery and HW Notes

- Lecture materials will be released on Canvas prior to the lecture.
- All HW assignments will be submitted via GitHub classroom (except the first Homework).
- First HW assignment focuses on learning/refreshing your knowledge about GitHub. It's released this Thursday and will be available until next Thursday January 26th, 11:59PM ET / 8:59PM PT.

Expectations down the Road!

- For some cloud platforms like Azure, you will have to use the Free Trial version which includes a step for you to add your Credit Card information. You will not be billed if you used the cloud platform correctly.
- For the final exam, we will use the Lockdown browser and you will have access to the lecture PDFs via the browser. You will receive few training attempts on the final exam environment before the exam date.

TA Introductions

- Dheeraj Nayak
- Yuqi Mao
- One more TA to be announced soon.

Next

- Complete Pre-semester Survey:
  - https://forms.gle/jTZKVSQTPtitGbnG8
- Submit your 12-digit AWS Account ID by tomorrow noon ET
  - This video will help you: https://youtu.be/gU1kjzgb-gA
- Sign-up for the course on TopHat.
- Download Anaconda on Your System
  - https://www.anaconda.com/
- Join the Course Piazza
- Join the Student Space on Slack

Waitlisted Students

- All materials will be uploaded here