

CS15-319 / 15-619

Cloud Computing

Recitation 2

January 21st & 23rd, 2014

Piazza

- TA office hours are posted on Piazza
- Reasons to use Piazza
 - The ONLY support you can have a couple of hours before deadline
 - Really good questions raised by your classmates
- Try to ask a public question if possible
 - Your classmates will benefit from it
 - Reduce duplicate questions

Piazza Questions

- Choose “basic” support level when signing up for an AWS account
- Recitation online video, click [here](#)
- All deadlines are 11:59 PM Pittsburgh Time.
- Quizzes
 - can be saved and resumed before submission
 - 1 attempt to submit
 - Preset duration after start time (2 hours)
- Projects
 - 3 attempts
 - we take the highest score

Piazza Questions

- Primer Project
 - Read REALLY carefully because you WILL need it
 - There is a Deadline but NO submission required

Amazon Web Services (AWS) Account

- Students who are just joining us or who have not completed the AWS Account Setup:

=== ONLY IF YOU HAVEN'T DONE SO ALREADY ===

1. Setup an AWS Account
2. Complete AWS information Google form
3. Wait to receive Consolidated Billing Request email from Amazon
 - Manual process, waiting time varies
4. Click the link to verify the linked billing
 - Many students have not clicked on the link yet!
5. Start using EC2, S3, etc (beyond the Free Usage Tier)

Amazon Web Services (AWS) Account

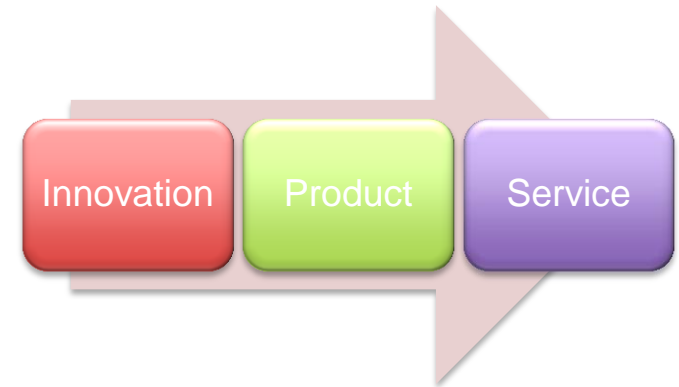
- **ALL STUDENTS:**
 - If you have created an AWS account and have not received notification that it was linked to our account
 - Your credit card on file may get charged!
 - We CANNOT reimburse you!!!

Last Week

- Project primer
 - Tagging
 - Checking expenditure on AWS
 - S3 Bucket
 - EC2 micro instance launch and remote login
- Introduction to cloud computing
 - Module1

UNIT 1: Introduction to Cloud Computing

- Definition
- Enabling Technologies
- Deployment Models
 - Private, Public, Hybrid
- Benefits and Risks
- Economies of Scale
- Building Blocks
- Service Models
 - IAAS, PAAS, SAAS
- Use Cases



UNIT 1: Introduction to Cloud Computing

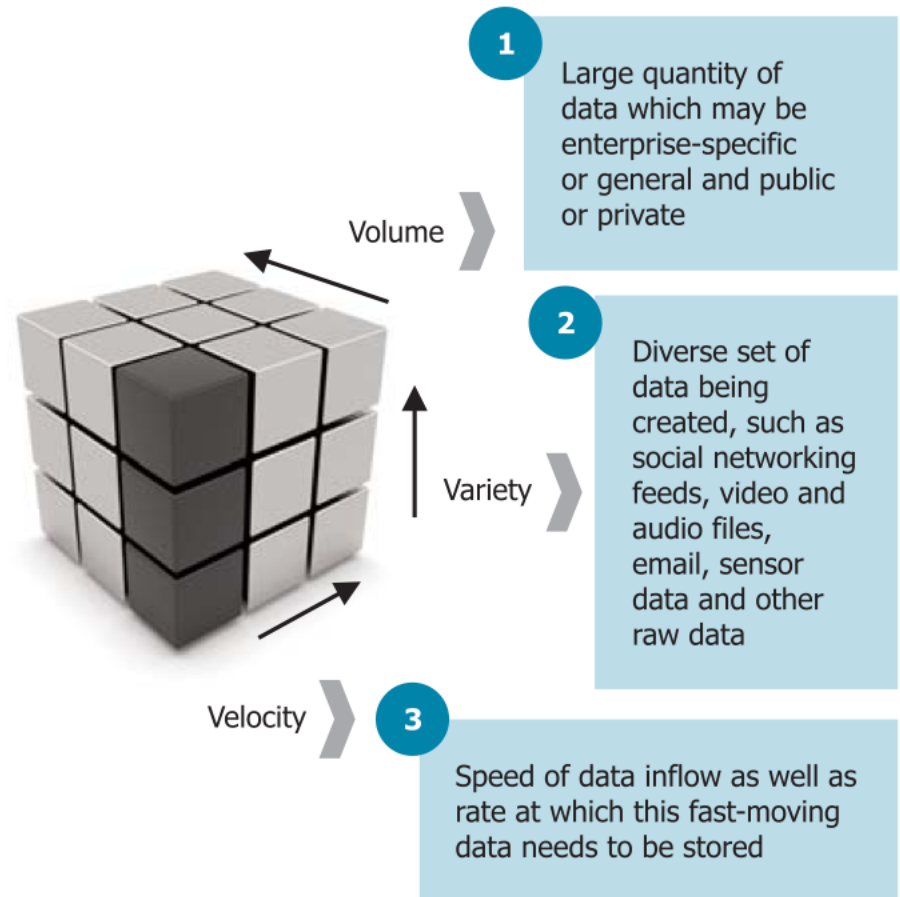
- Complete Unit 1 by Thursday
 - Read all pages in modules:
 - Module 1: Introduction to Cloud Computing
 - Module 2: Building Blocks and Service Models
 - Complete activities on each page
 - In-module activities are not graded but for self-test
 - Take Checkpoint Quiz 1
 - (Deadline, Thursday, January 23, 11:59pm EST)

Quiz 1

- Quiz 1 is open
 - This assignment is timed. You will have 120 minutes to complete the attempt once you begin.
 - Deadline for completion is Jan 23, 11:59 PM Pittsburgh
 - Late submissions are NOT accepted
 - You may not start the assignment after the deadline has passed.
- You only have 1 attempt

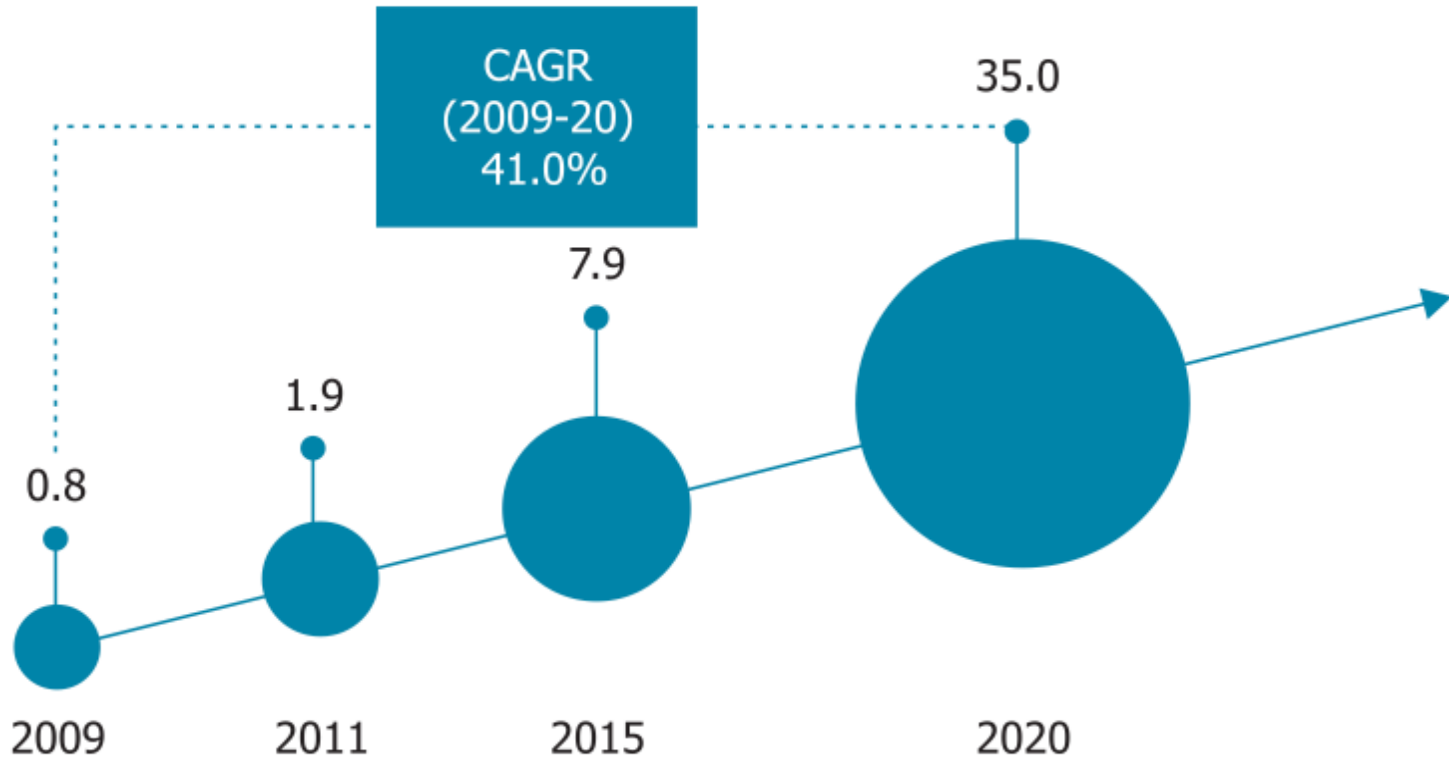
Big Data

- What is Big Data?
 - It is the proliferation of data that floods organizations on a daily basis .
 - It is high volume, high velocity, and/or high variety information assets.
 - It requires new forms of processing to enable fast mining, enhanced decision-making, insight discovery and process optimization.



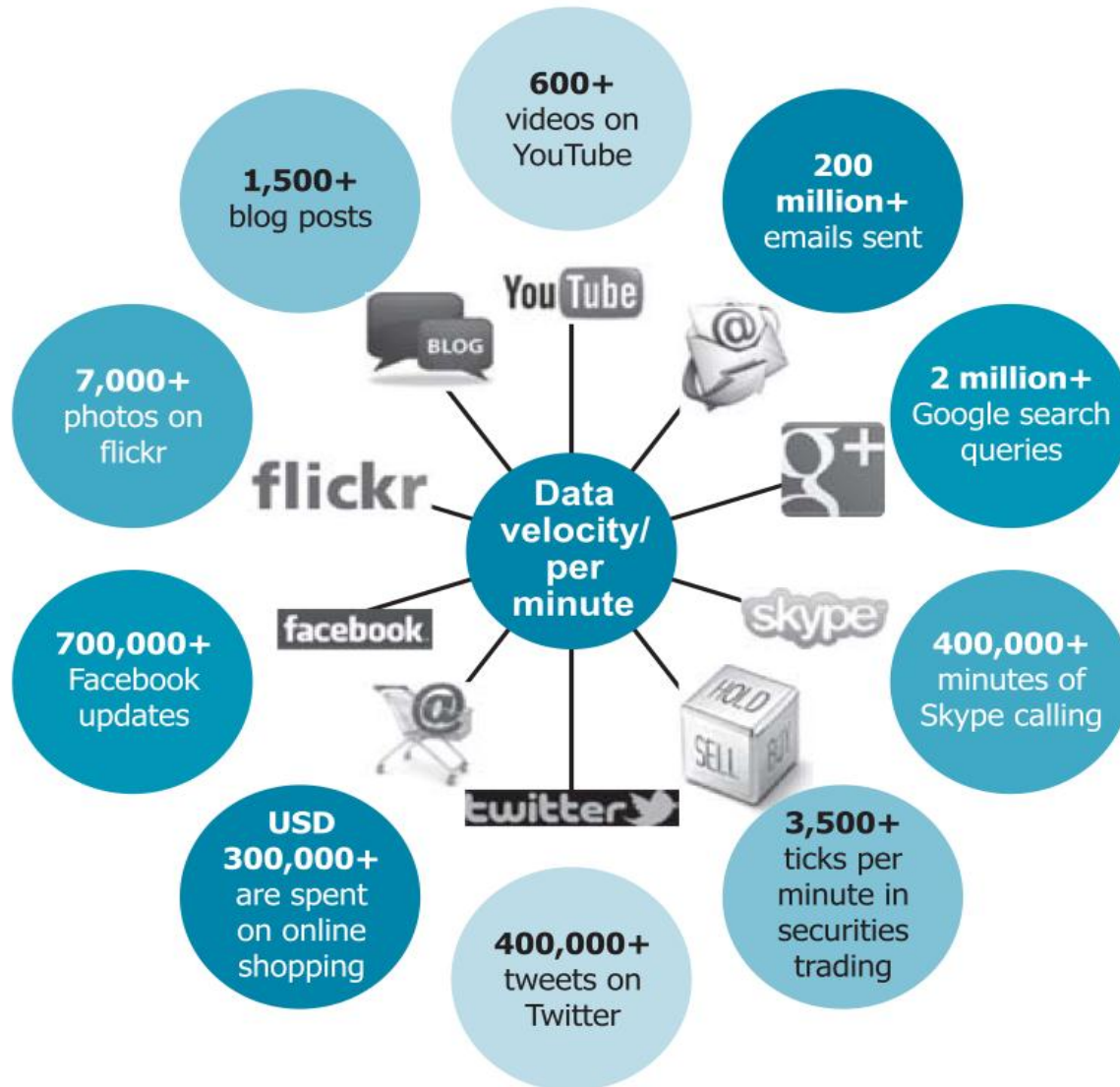
Big Data

(Growth of global data, 2009-20)
Zettabytes



*CAGR – Compound Annual Growth Rate

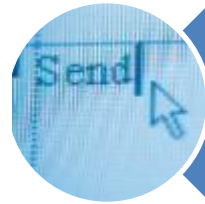
Big Data



What Do We Do With Data?



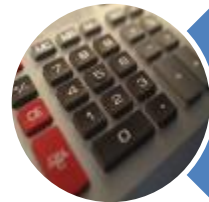
Store



Share



Access



Process



Encrypt



.... and
more!

We want to do these seamlessly...

Using Diverse Interfaces & Devices



Computers



Mobile Devices



Consumer Electronics



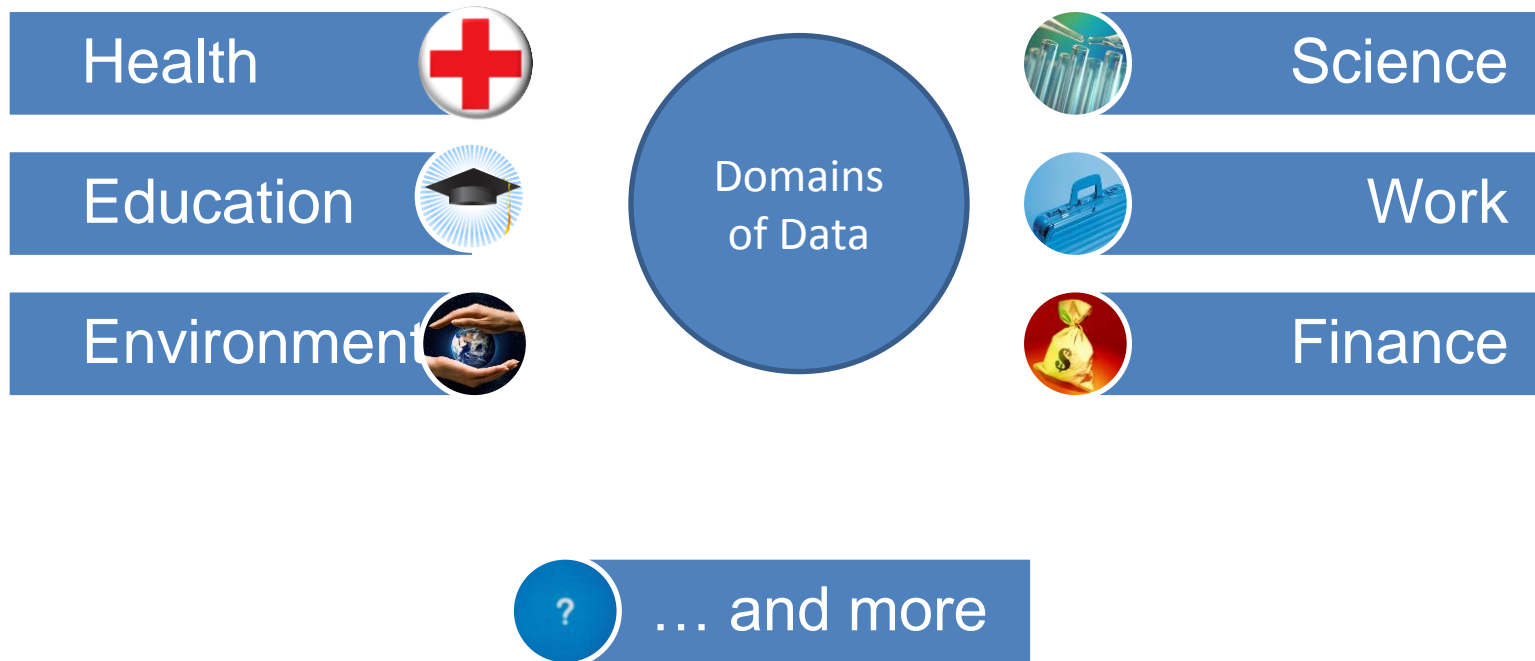
Personal Monitors and
Sensors



...and even appliances

We also want to access, share and process our data from all of our devices, anytime, anywhere!

Data Becoming Critical to Our Lives



Project 1, Module 1

- Wikimedia data set
 - Wikipedia page views
- Explore the data set to learn the format

<project name> <page title> <number of accesses> <total data returned in bytes>
- You need to parse and filter the data
 - As described on OLI
- Sort the data and save the output to a file

Project 1

1. Log on to OLI and complete the checkpoints in OLI modules:

- **Introduction to Big Data**

- Sequential Analysis

- **Deadline: 11:59 PM Pittsburgh Jan 26 (Sunday)**

Demos

- Demo 1:
 - Provision a micro instance
 - Preparation: Key Pairs
 - Dashboard & Security group settings
- Demo 2:
 - SSH to a provisioned instance
- Demo 3:
 - Build S3 bucket folder structure
 - Download File from Amazon S3
 - access key/secret key != Key Pairs

Demo

- S3 folder structure
 - Set Permission Policy
 - Only upload your source code file (.java, .py, .sh...)
OR
source code files in a single zip/tar/tgz/gz...
 - DO NOT submit .jar file

Discussion

- Your questions...

Upcoming Deadlines

- **Quiz 1:** Introduction to Cloud Computing
 - Checkpoint Available Now
 - Due 1/23/14 11:59 PM
- **Project 1:** Introduction to Big Data Analysis
 - Sequential Analysis
 - Checkpoint Available Now Due 1/26/2014 11:59 PM