CS-E4640 Big Data Plaforms

How to succeed in Big Data: Experiences from The Course
The Real Job & Assignment Questions

Tri Nguyen
PhD student at Department of Computer Science
Researcher at AaltoSEA

Assignment

- **Build your own Big Data Platform**
  - Inspired by your projects or real applications
    - *What are the big data problems?*
    - *Macro-thinking*
  - Technologies:
    - *Hands-on Tutorials*
    - *Tech Radar*
    - *Your favor*
    - ➢ *Why do you use these technologies but not others?*

Assignment

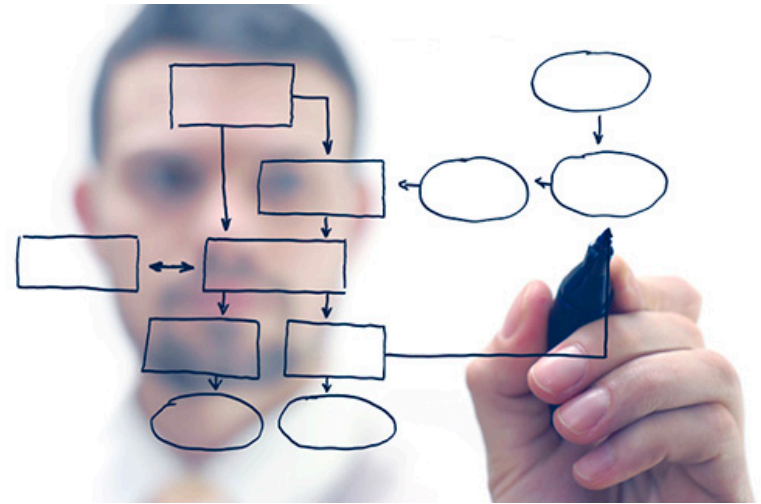- **Design & Implementation**
  - Clear Architecture
    - *Module*
    - *Function*
    - ➢ *Realistic, reasonable, scalable, extendable, …*

  - ➢ Tips:
  - *- Stick to the assignment instructions*
  - *- Try to make it reusable for later assignments*

Assignment

- **Experiments**

  - Again! be realistic, macro-thinking

    - *Workload? Why they are big data problems?*

    - *Monitor/measure multiple metrics, different scenarios, failures,...*

  ➤ Tips:

  - What are existing solutions, why/when they are applied?

  - Simulate big workloads to see potential problems, interesting points,...

Assignment

- **How do we evaluate your works?**
  - Based on the question outlines
    - *Answer one by one in order following the question outlines to help us save grading time without missing your points!*
    - *Be clear, concrete and complete!*
  - Demo
    - *Demo & F2F meeting*
  - ➤ Tips:
  - *- Stick to the question outlines, don't mess among questions!*
  - *- Search for existing solutions and show your understanding*
  - *- Don't copy others' works without citation!*

Thank you!

Any Question?

Aalto University
School of Science