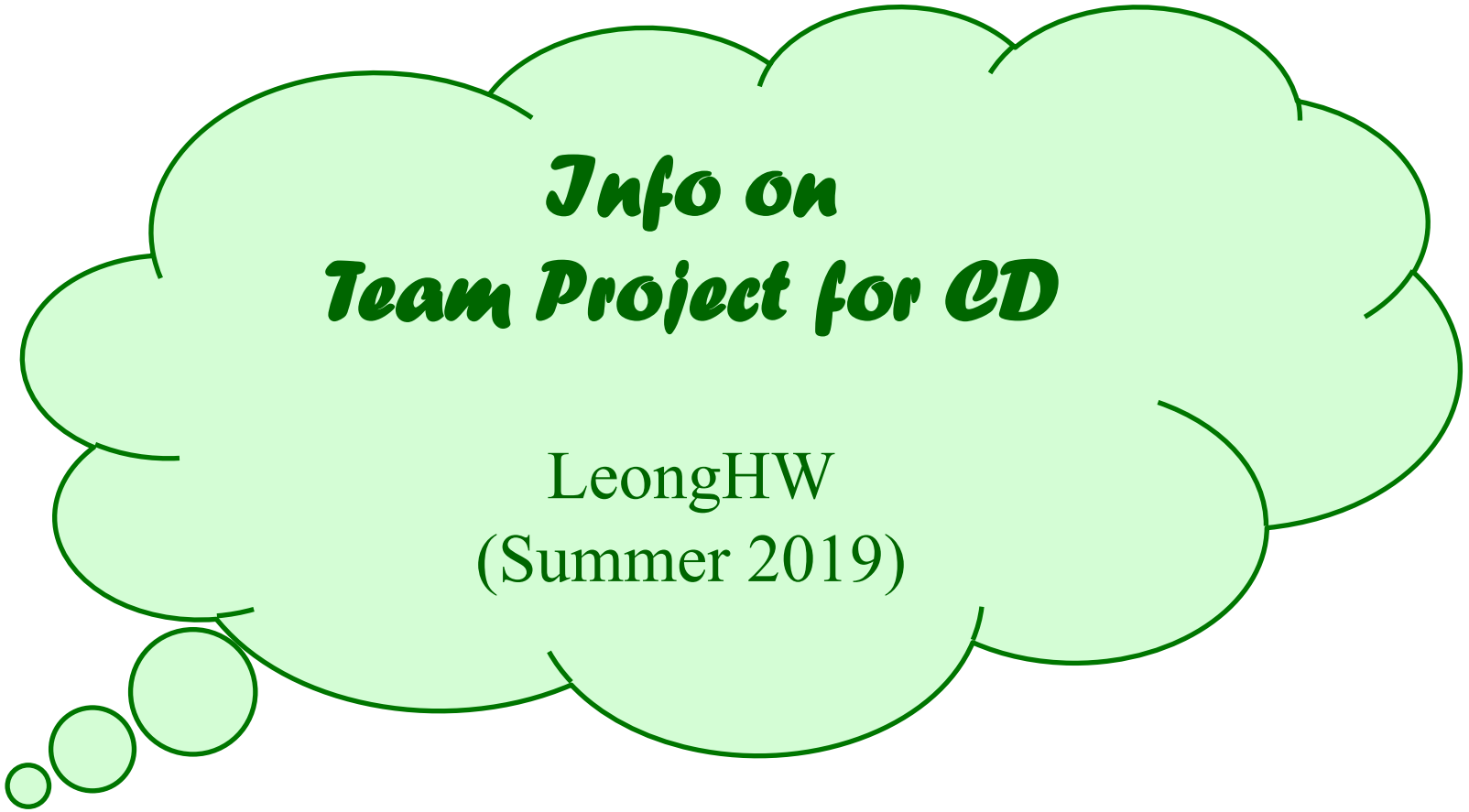
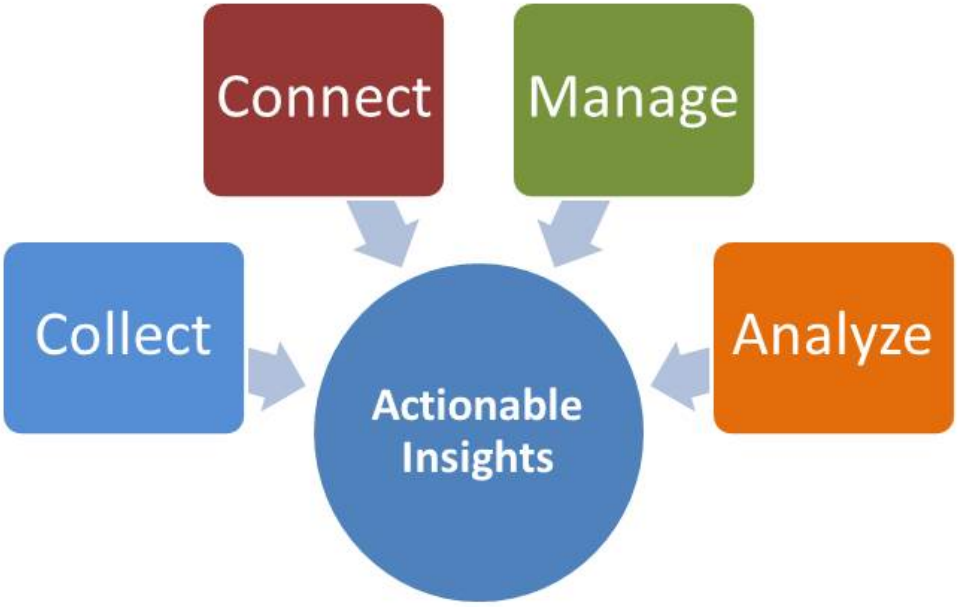
CD: Project

NUS National University of Singapore
School of Computing



Data analytics workflow…

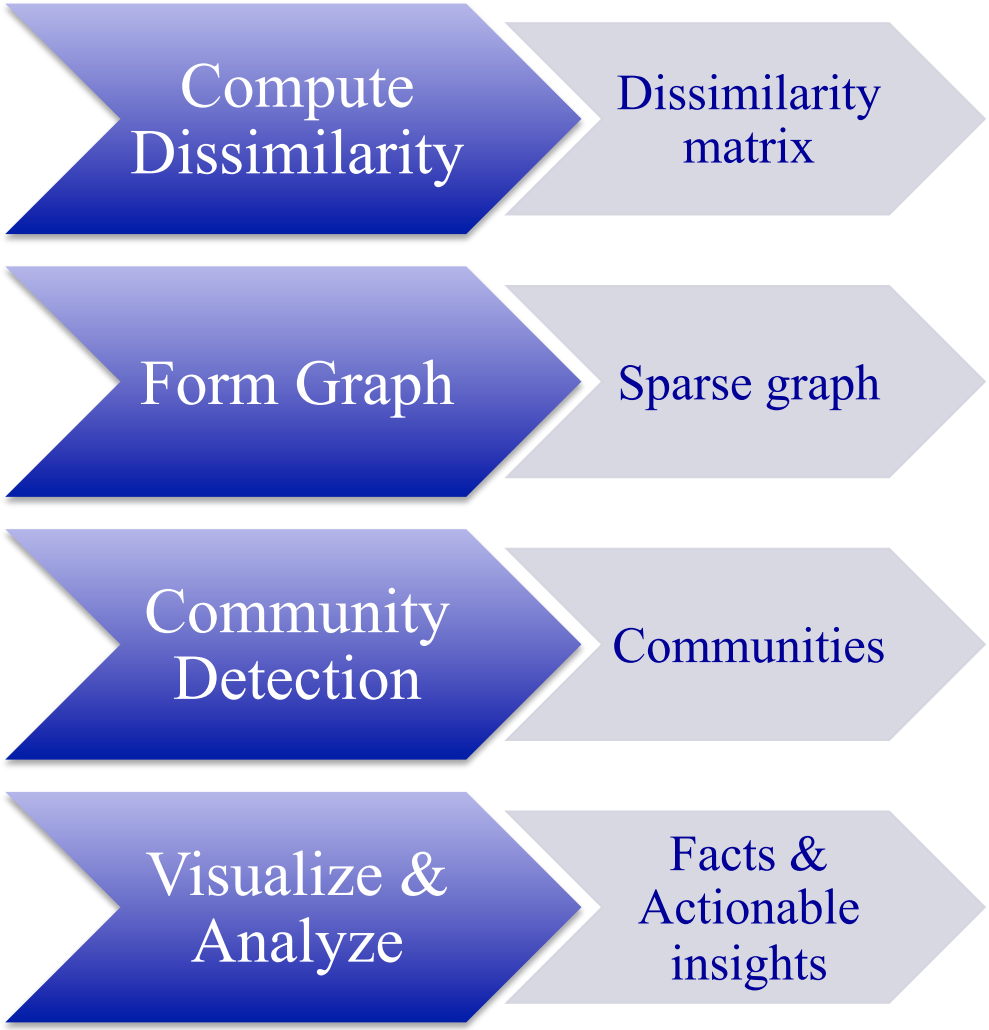


Dilemma in Big Data analysis

- Know what we know
- Know what we don't know
- Don't know what we know
- Don't know what we don't know

What can we do?
How can CD help?

CD-Proj Workflow…



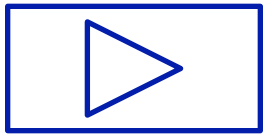
Project GOALS

1. Big data gives us big graphs
2. Detect communities in these graphs
3. Analyze the communities detected
   - ❖ Derive/confirm known facts / relationships

- ❖ Derive new insights
  (betw. community & relationships)
- ❖ Can turn into competitive advantage?

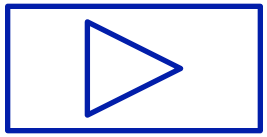
Projects from Summer 2018 (1)

1. Evolution of Architectural Genre
2. Analysis of Monkey Species
3. Gossip Arts (analysis of paintings/painters)
4. CD of Chinese Movie Actors
5. Nerd Gallery (CD on Painting Styles)
6. Fashion Miner (CD on Runway Fashion)



Projects from Summer 2018 (2)

7. Music Genre Analysis (Music Video Scenes)
8. The Next City (CD of Cities)
9. Relationship Discovery of Wikipedia Entries
10. Target: James Bond
11. Influence of CNN Model on CD Algorithm



Feedback Summary (2018)

- **M1: Don't take too long with project proposal**
  - Choose images quickly (prepare/trim them)
- **For M2**
  - Quickly go through *one complete cycle*
  - Then even if you change data, you can do it QUICKLY
- **For M3:**
  - Check the "quality" of your communities
  - What are known/expected insights?
  - Can you find new/unexpected insights?

Feedback

- Communities are *the end* of your project
- Communities are *starting points* of analysis
- Did you discover any *new insights*?
- Turn *new insights* into ***actionable insights***

Thank you.

Q & A

School of Computing