The Portrait of the Hackers as Young Men or Things Fall Apart

Andrew Higginbotham, Tyler Pease

Starting out

- Mostly chaos
- Assembling into groups
- Then not working well as a team
- Figuring out we don't know enough JS

The Plan

- Run a search
- Get/store text
- parse text
- concatenate text
- display text

Initial Success

- Can pull and store text from a page
- learning about python
- promising search ideas

Initial Problems

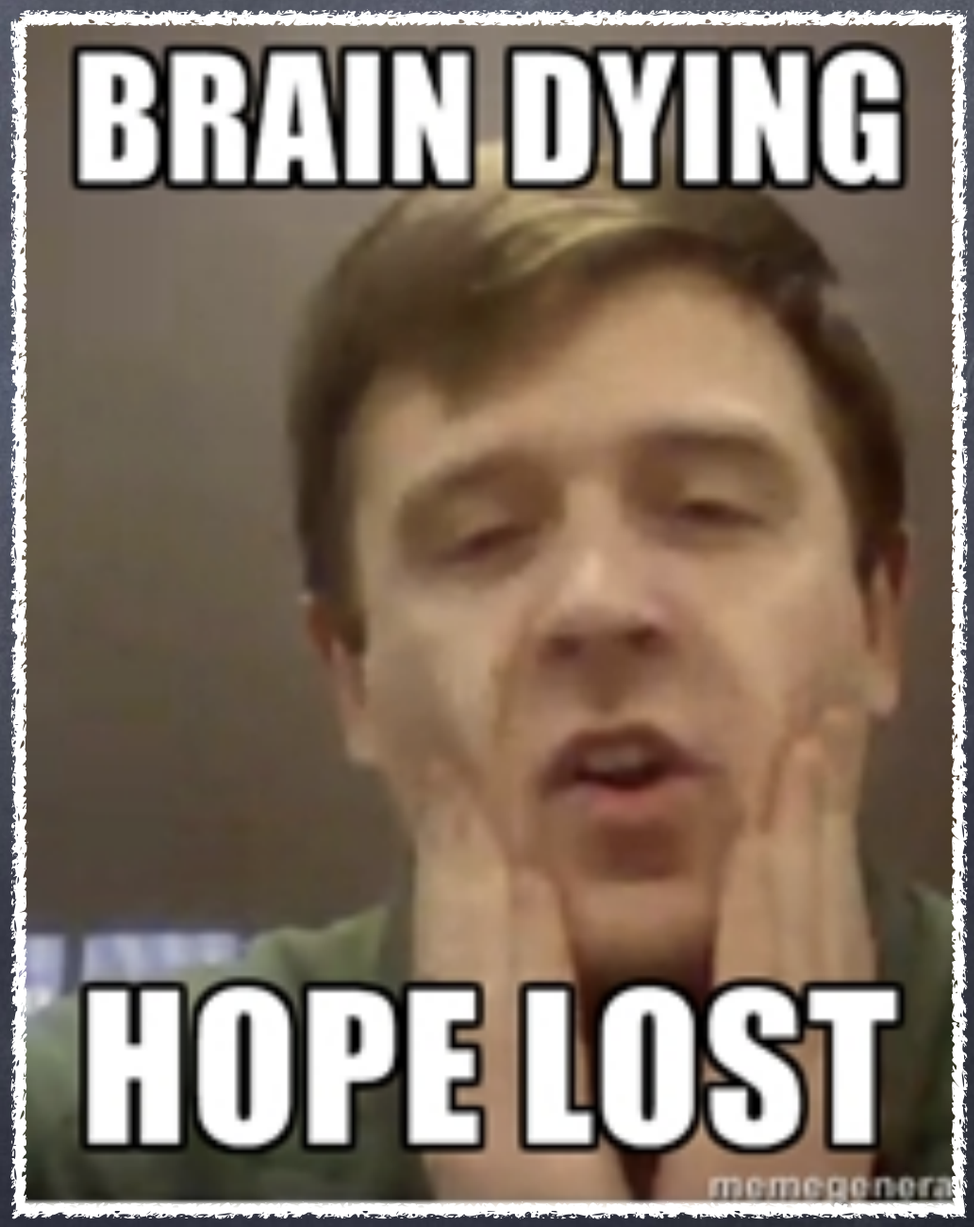
- Google search API deprecated
- Research into others went nowhere
- Dead end with no sources
  - unless hard-coded

Scrapy

- Python web-crawler
- Use on Google Scholar
- crawl 1-2 levels vs. use 2 spiders
- 3 hours of research...
- WON'T CRAWL ANYWHERE!!!!

Problem List

- hard coded sources
- no parser
- no front end



Parser

- Attempted custom C-parser to strip all HTML tags/content
- 3 hours with mild success
- Found Python HTML2Text
- 5 minutes and minimal success
- Finally decided to use HTML as it was

End Result

- A wild, jumbled hack
- multiple pieces
- different languages
- disjointed and going nowhere

What We Learned

- Hard to learn languages in a night
- Working better as a team
- Working through adversity
- Explaining the process of failure
- Making last minute PowerPoint to have something to show in class