

Managing Data Science Projects

Brendan Herger, hergertarian.com
Slides: <https://bit.ly/2Zhdonp>

THIS IS YOUR MACHINE LEARNING SYSTEM?

YUP! YOU POUR THE DATA INTO THIS BIG
PILE OF LINEAR ALGEBRA, THEN COLLECT
THE ANSWERS ON THE OTHER SIDE.

WHAT IF THE ANSWERS ARE WRONG?

JUST STIR THE PILE UNTIL
THEY START LOOKING RIGHT.



Intro

Software engineering management

Data science project management

Pro tips

Recap

Intro

Opportunities

- **Value:** Data science projects are incredibly valuable
- **Scope:** Can be closer to research: Open ended, changing scopes
- **Repeatability:** No culture of 'assembly line data science'

Software engineering management

**Easy things should be easy,
and hard things should be possible.**

– Larry Wall,
Father of Perl programming language

Software engineering

- Agile or Scrum project management
 - Gather requirements
 - Create milestones & specs
 - Implement
- Designed to be pretty linear

Data Science project management

Data Science

- “It’s hard to scope”
 - Gather requirements
 - “Best first search”

Data Science Opportunities

- Consistency
- Standardization & automation
- Time boxing / Pomodoros

Pro tips

Pro tips

- **Scoping:** Scope outcomes, not methods
- **Steps:** Treat data gathering, data QA as explicit steps
- **Overestimating:** Include buffer
- **Re-scoping:** Build in clear go / no go cutoffs
- **Flexibility:** Be honest with stakeholders about non-linearity

Recap

Agenda

- Intro
- Software engineering management
- Data science project management
- Pro tips
- Recap

Takeaways

- DS is not software engineering (yet)
- Leave room for milestones to change
- Be explicit with granular tasks

Thanks!

Brendan Herger, hergertarian.com

Slides: <https://bit.ly/2Zhdonp>