**Capstone: Checklist**

* EDA: by April 7
* Prophet Time Series (multivariate): by April 7
  + Applies to all Subreddits
  + Cross-Validation
* Tensorflow Modelling: by April 7
  + If possible, try PyTorch later
* LDA: by April 7
* Results: by April 17
* Paper: by April 21 ([draft](https://docs.google.com/document/d/1FLXNjLhrWxDaQ6zLLM4xbK2V45C__sR-mwUIQYLY26I/edit#heading=h.bjbvngh5kjx))

(12-15 pages single spaced; formatted like a research paper/conference submission)

* + Introduction
  + Data Collection and Transformation
  + Ethical Considerations
  + Methodology
    - EDA
    - Prophet Time Series
    - BERT Model
    - LDA - Topic Modeling
  + Results
  + Conclusion
  + Future Directions
  + References
* Presentation (15 mins): by April 30
  + Slides
  + Notes
* Talk to Maksim about paper/presentation: April 24

Conferences:

* AI in Social Science : Due May 1st

<https://bfi.uchicago.edu/event/call-for-papers-ai-in-social-science-conference-2024/>

* Social Implication on AI: Due May 24th

<https://attend.ieee.org/istas-2024/>

Preparation for paper

EDA:

* Data collection
* General go over dataset
* Central tendency and residual
* Log transformation - correlations
* Descriptive statistics
* measurement

Model - bert, lda, prophets

* Bert - accuracy, model
* LDA topic modeling - wordcloud
* Prophets analysis model - post engagement, post trends