EDS 411B: MEDS Capstone Week 4

Instructor: Ruth Oliver
Email: rutholiver@bren.ucsb.edu

This quarter

Deadline	*Item*
April 14 (Week 2)	Technical Documentation Outline
May 5 (Week 5)	Technical Documentation Draft & Repo Draft
May 17 (Week 7)	Shiny deployed for Public Presentations
June 2 (Week 9)	Public Presentations (at Bren)
June 9 (Week 10)	Final Project Materials due

Documentation feedback

Design and Implementation Plan vs. Technical documentation

	Design and Implementation Plan	Technical Documentation
audience:	Your team	Future employers, students, faculty, and users
purpose:	To PLAN!	Articulate project objectives and guide future users
access:	Private	Public on the Bren website

Document with a purpose

- **Who** will consume this document?
- **Why** do they need this document?

Stakeholder	Primary Interest	Common Artifacts
ML and data professionals	How the model works What data is used	Code comments Model cards README files User stories
Software engineers	How the system runs Service level agreements	Code comments Runbooks README files User stories
Business stakeholders Product owners	Use cases Business impact / ROI	Slide decks User stories Product roadmaps Cost benefit analyses
End users	How to use the system	User guides
Impacted individuals	Key decisions that impact me	Touchpoints such as emails or push notifications
Regulators	Regulatory compliance Data privacy	Compliance audits
Project team	How can we efficiently deliver the project	Project plans User stories Design documents
Security professionals	Data privacy System security	System audits Data usage reports
Quality assurance professionals	System reliability	Code comments Test use cases User stories

https://www.datascience-pm.com/documentation-best-practices/

Document with a purpose

Data:

- What data is being used for the model?
- Why was this data selected (and other data excluded)?
- How was the data obtained?
- What are the known issues in the data?
- What does the data look like? (mean, median, volume, etc.)
- How did you alter the data? (imputations, transformations, cleaning)
- Where is the data located?
- Is the data publicly available?

Document with a purpose

Model:

- What are the inputs and outputs?
- What assumptions did you make?
- What was the control/test split?
- What did you use as a validation set?
- What approaches did you try, but abandon and why?

Document with a purpose

User documentation:

- Who is the intended user?
- Where is code/data stored?
- How do I control visualizations?
- What are the definitions for key measures and dimensions?
- How can the code be updated

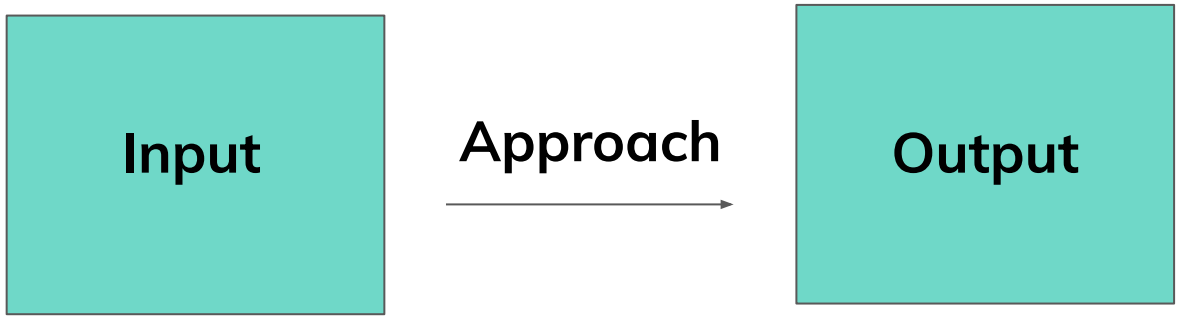
Celebrate your contributions with your team, but remove most references to your effort (e.g. “we made,” “we collected,” etc.).

“We created a repo with all product materials including x, y, z at github.io/wheresyourproject.”
“All project materials including x, y, and z are archived at github.io/wheresyourproject.”

“We collected and wrangled 32 datasets, which are now hosted on the Taylor server…”
“Cleaned datasets from the 32 sources (Appendix B) are archived in the NAME folder on the Taylor server.”

Focus on the “why” and support with the “how”.

“We used bioacoustic and remote sensing data to create linear mixed effects models.”
“Species diversity was modeled using linear mixed effects models based on bioacoustic and remote sensing data.”



Priorities!

(modified) Manifesto for Agile Software Development

We have come to value:

Individuals and interactions over processes and tools

Working software over comprehensive documentation

Responding to change over following a plan

That is, while there is value in the items on
the right, we value the items on the left more.

Work time until 3:40pm

Check-ins

- FireFutures
- iMPAct
- AquaFire
- steelTracker
- mosaiks
- wattmaps
- kelpGeoMod
- PYFOREST