ENVS 410/510: Data analysis and visualization

Dr. Lauren Hallett
Dr. Andrew Muehleisen
Lina Aoyama


PROF MACGYVER CONQUERS
REMOTE LEARNING
THANKFULLY TECHNOLOGY IMPROVED
imgflip.com

Overall goals

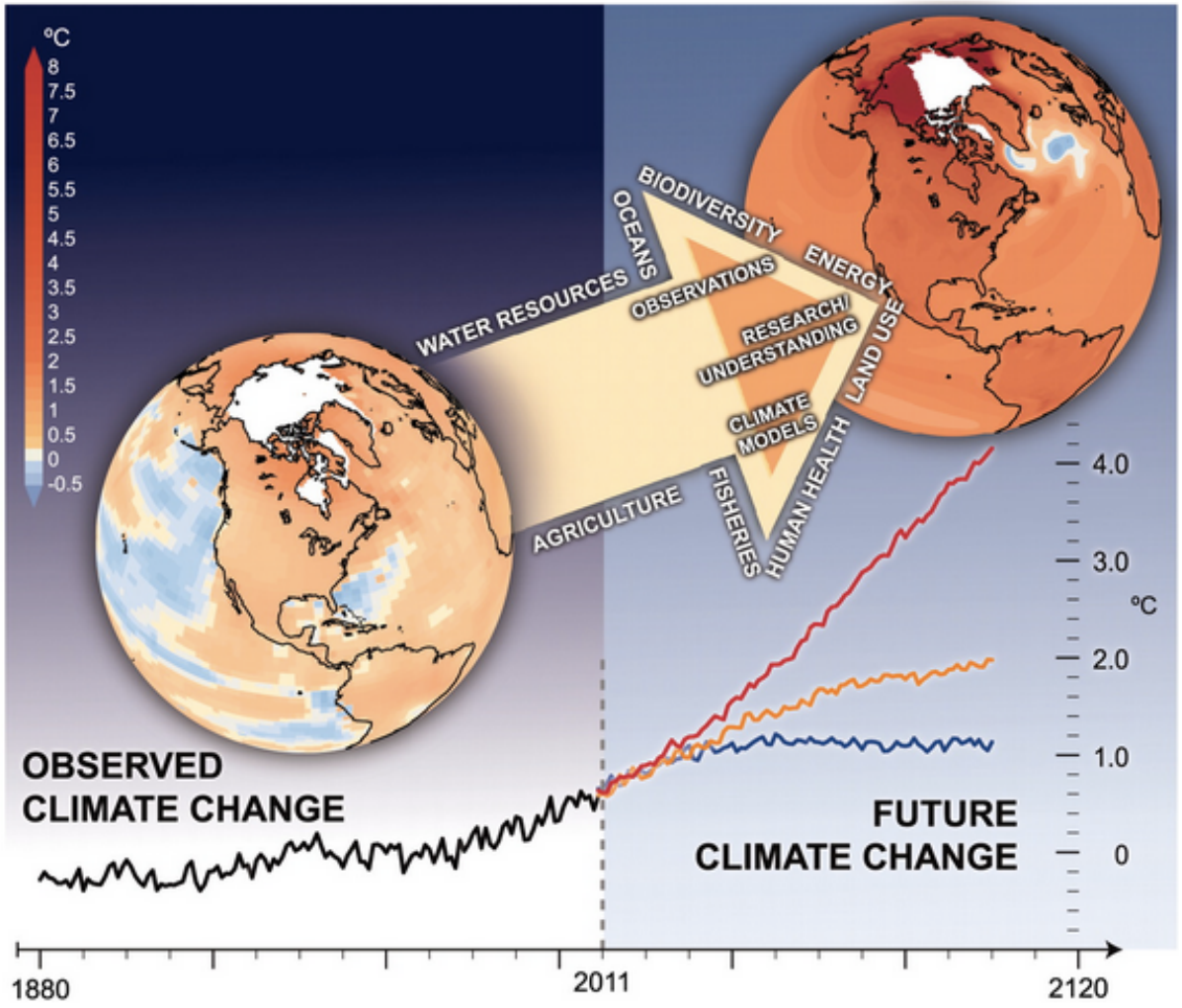
Your learning

- Understand and design steps that link raw data to communicated findings
- Interpret raw data as well as figures presented in scientific papers/the media
- Gain exposure to R and comfort with programming

Your experience

- Resources to empower you to learn and explore
- Flexibility to adapt to challenging times
- Community to support and develop together

Global Change



Overpeck+ (2011) doi:10.1126/science.1197869

Remote sensors

credit: NASA

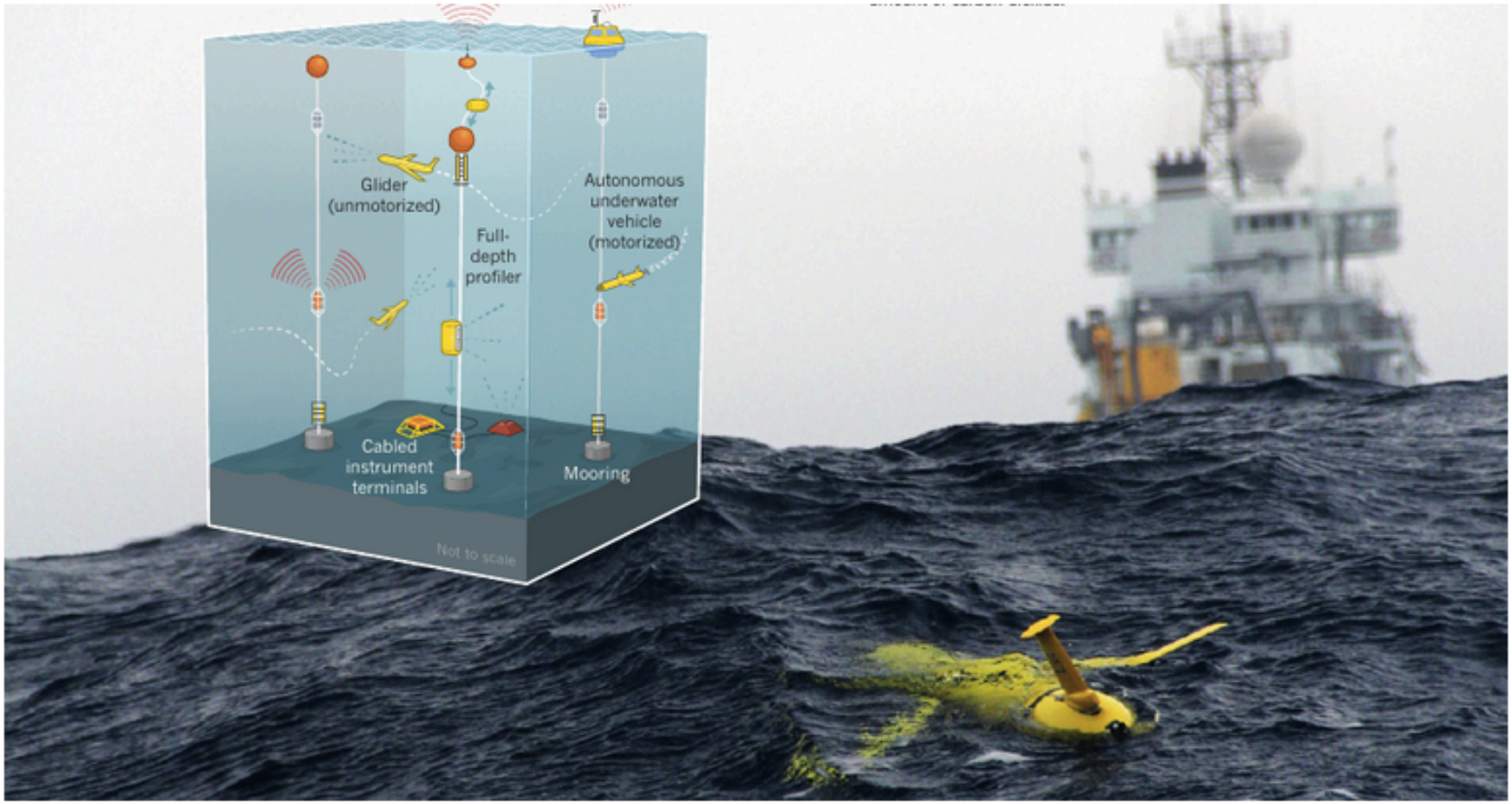
micro sensors

credit: NSF

NEON

credit: Hopkin (2006) doi:10.1038/444420a

OOI



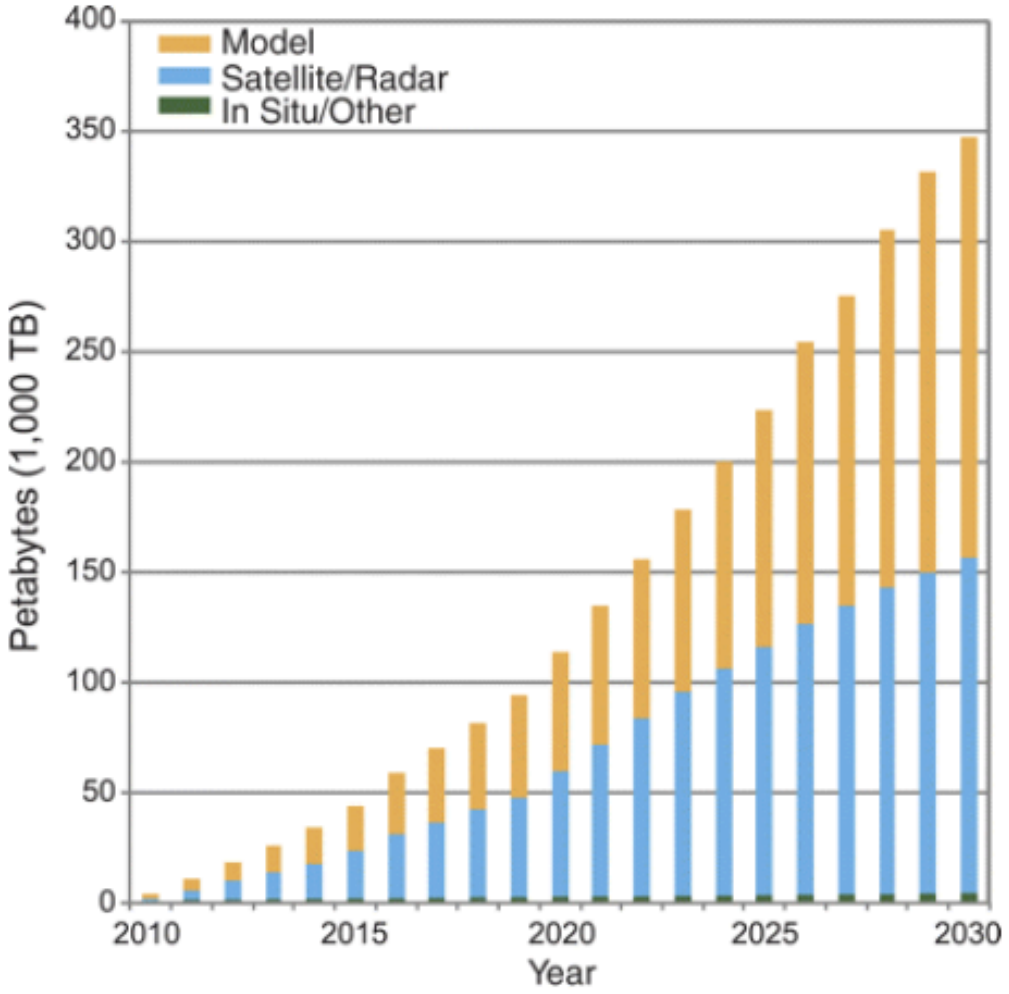
credit: Witze (2013) doi:10.1038/501480a

Field-based study

Computer simulations

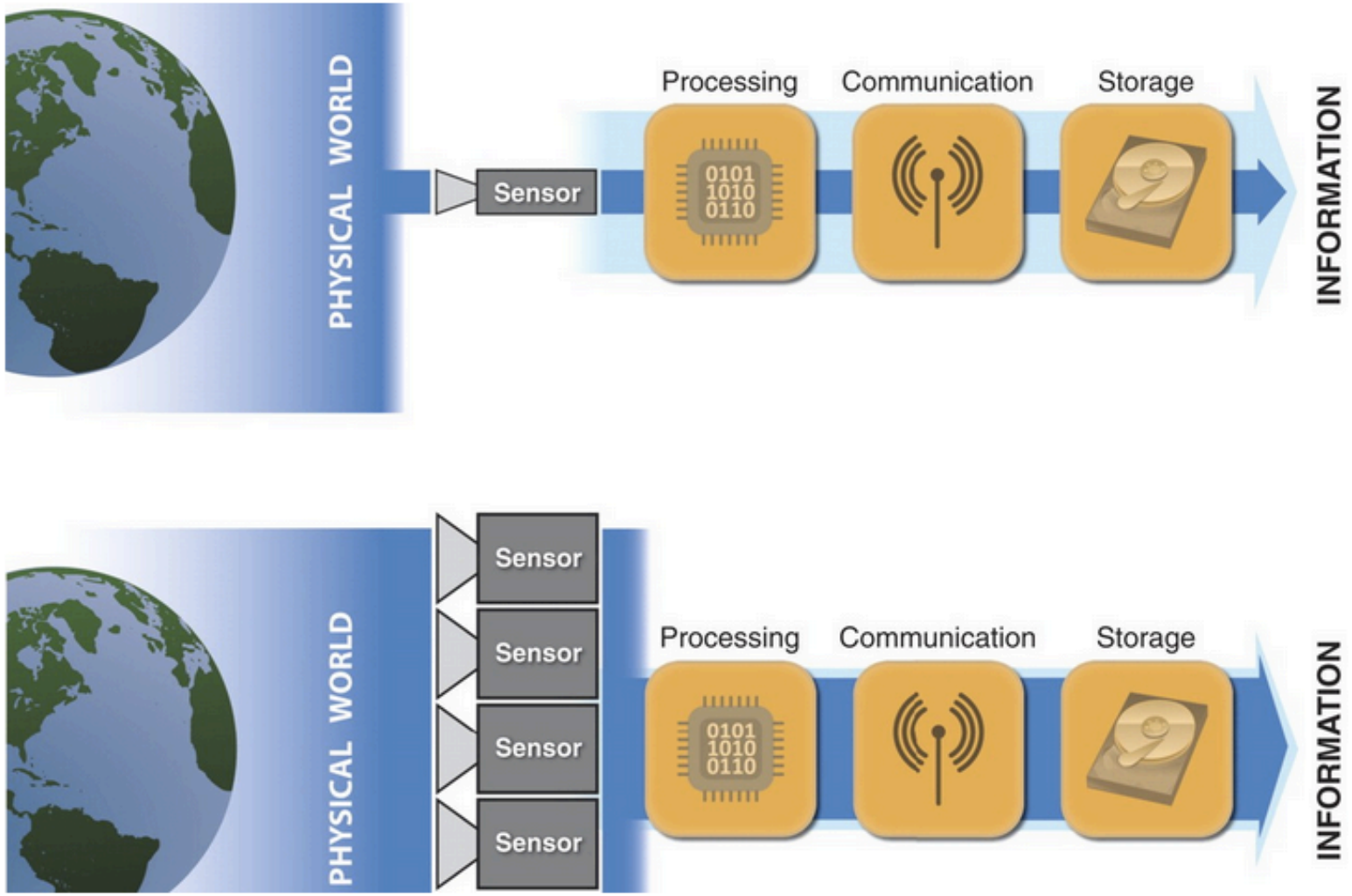
credit: NSF Cyverse / Jetstream

Growth of climate data by type



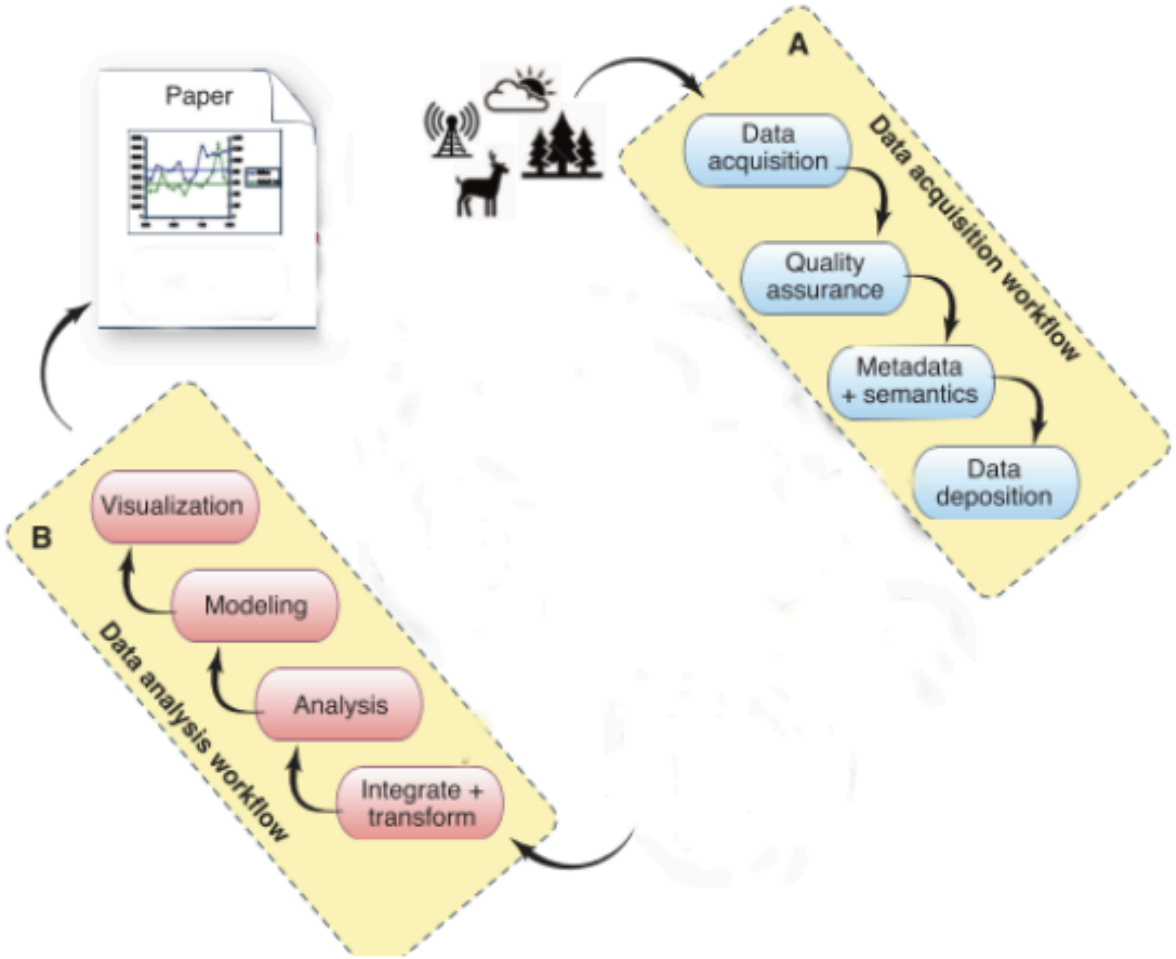
Overpeck+ (2011) doi:10.1126/science.1197869

Engineering bottlenecks



Baraniuk (2011) doi:10.1126/science.1197448

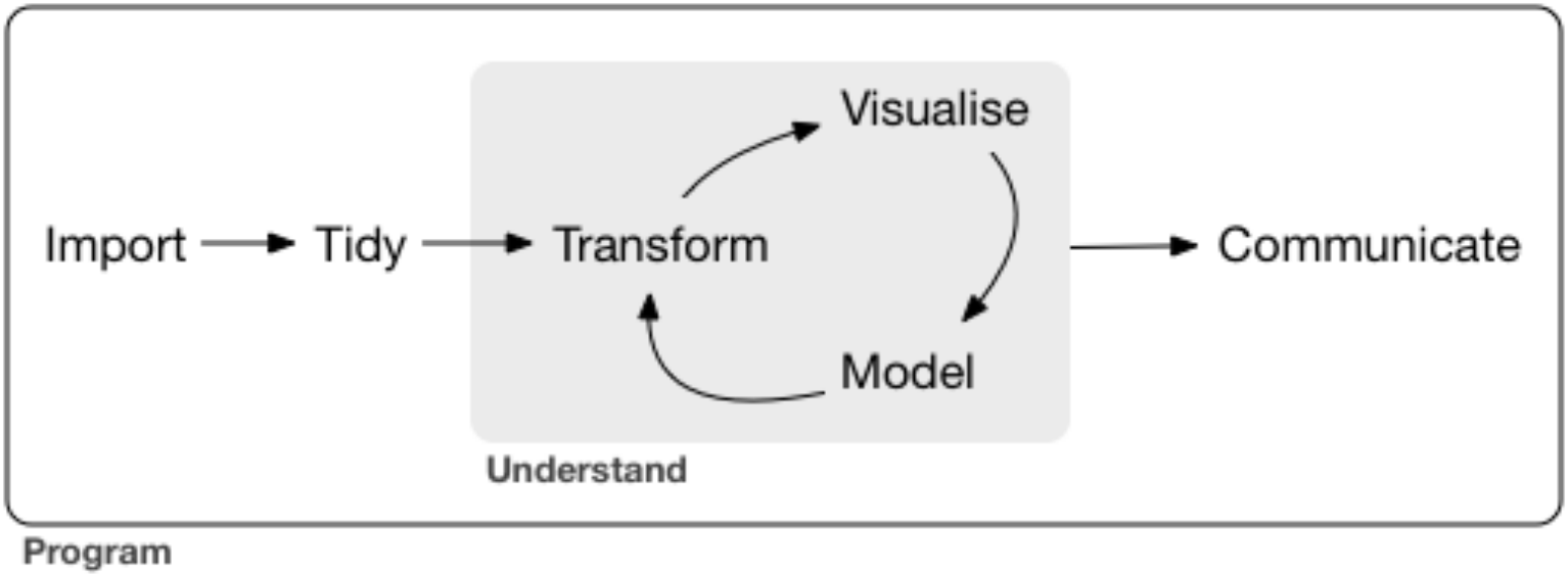
Science bottlenecks



adapted from Reichman+ (2011) doi:10.1126/science.1197962

Our focus is on these
Science bottlenecks

Course content: Developing workflows for the data life cycle



Wickham and Grolemund R for Data Science

This course:

Real data, real tools

- Things will break
- Things will change

This course:

Some things will be in keeping with good pedagogy:

- We'll start with graphs first
- Short videos and active learning

Some things aren't:

- We'll jump forward and back at times
- Starting at the deep end

How we'll do it:

- Different modules that transition from demonstration to independent coding
- Modules can be completed fully asynchronously, but we recommend you watch the videos and then finish the code with us in class
- Frequent practice via four problem sets
- An independent project in which you go from raw data to a reproducible final report
- A free online text to help you along! https://r4ds.had.co.nz/

Language: Why R?

- Scripted languages allow reproducibility
- Not about technical attributes
- Pick your language based on what people in your area speak/code
- Rstudio, Rmarkdown and other tools for communication and sharing
- A great open-source community (on that note - thanks to Carl Boettiger and Jenny Bryan for many of todays slides)

Imposter Syndrome

Before the first class

- Check out materials in the “START HERE” module to learn the nuts and bolts of course structure and introduce yourself
- Watch the videos associated with the “Week 1” module
- Follow the computer set-up instructions to install R and RStudio