Open by Design: Planning a Bioinformatics Project

Caleb Kibet

19th July 2022

Twitter: @calkibet @openke #OpenScienceKE

OpenScienceKE

“If You Fail to Plan, You Are Planning to Fail” — Benjamin Franklin

What comes to mind when you hear the statement 'project planning'?

What contributes to a successful project?

Project Management

- **project management** is the discipline of **planning**, **organising**, **securing** and **managing** resources to bring about the successful completion of specific project goals and objectives
- a **project** is a temporary endeavour, having a defined beginning and end

Developing a new project

- Gap, aims, expectations from the research
- Earliest version of project plan
    - What do I need to know to achieve the research goals?
    - What resources are needed?
    - What are the appropriate funding sources?
    - An idea of constraints and bottlenecks.
- Create a potential timeline
- Project Proposal (For funding or University Approval)
- Project Plan

https://www.jstor.org/stable/j.ctt6wp816

Project goal or Aim

A research aim describes the **main goal** or the overarching **purpose** of your research project.

Why…What…How

SMART

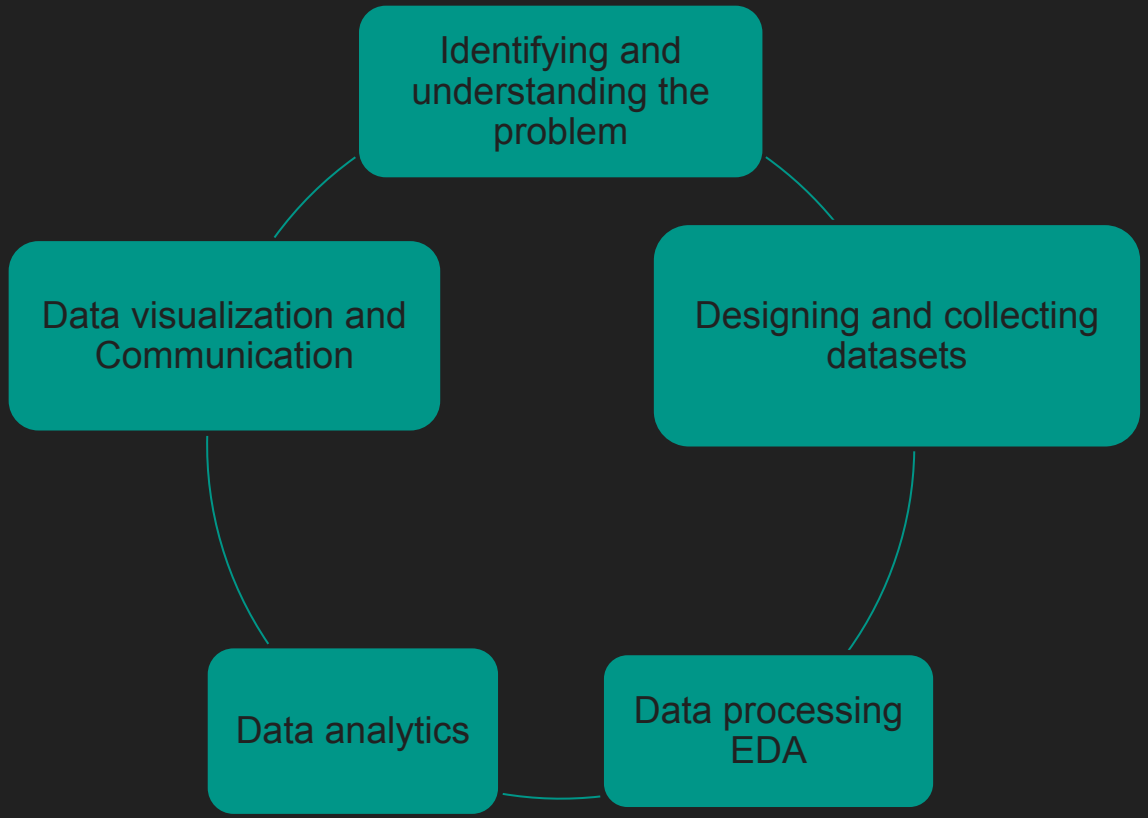
Some Main Concepts

- Plan for how much time and money it will cost you to publish your results and data such that it is freely available right from the start
- Ask for these costs in your grant proposals
- Publishing both your data and final research products openly increases trust
- Provide good documentation and metadata so that others may reuse and remix your data, further increasing your citation rate
- Pre-register your research project to avoid “scooping” and to further increase trust

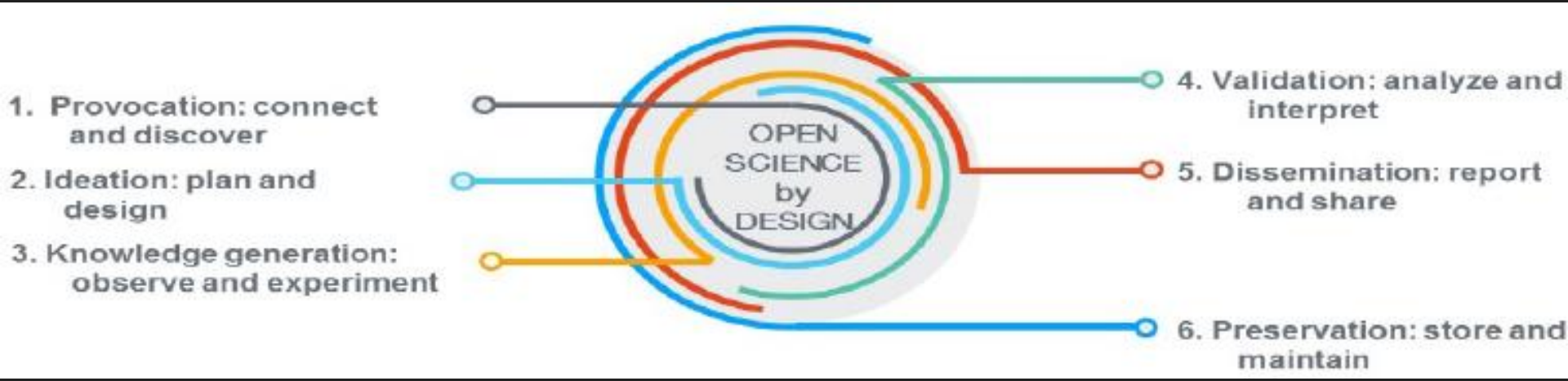
”From the very beginning of the research process, the researcher both contributes to open science and takes advantage of the open science practices of other members of the research community”

What do you include in a project proposal?

Data Analytics Life Cycles



Phases of Open Science by Design in the research life cycle



SOURCE: https://www.nap.edu/read/25116/chapter/2#4

1. **Ideation**: develop and revise research plans and prepare to share research results and tools under FAIR principles

- Test hypothesis with existing data
- DMP: how data, workflow and code generated will be shared
- ?

2**. Knowledge generation**: collect data, conduct research using tools compatible with open sharing, and use automated workflow tools to ensure accessibility of research outputs.

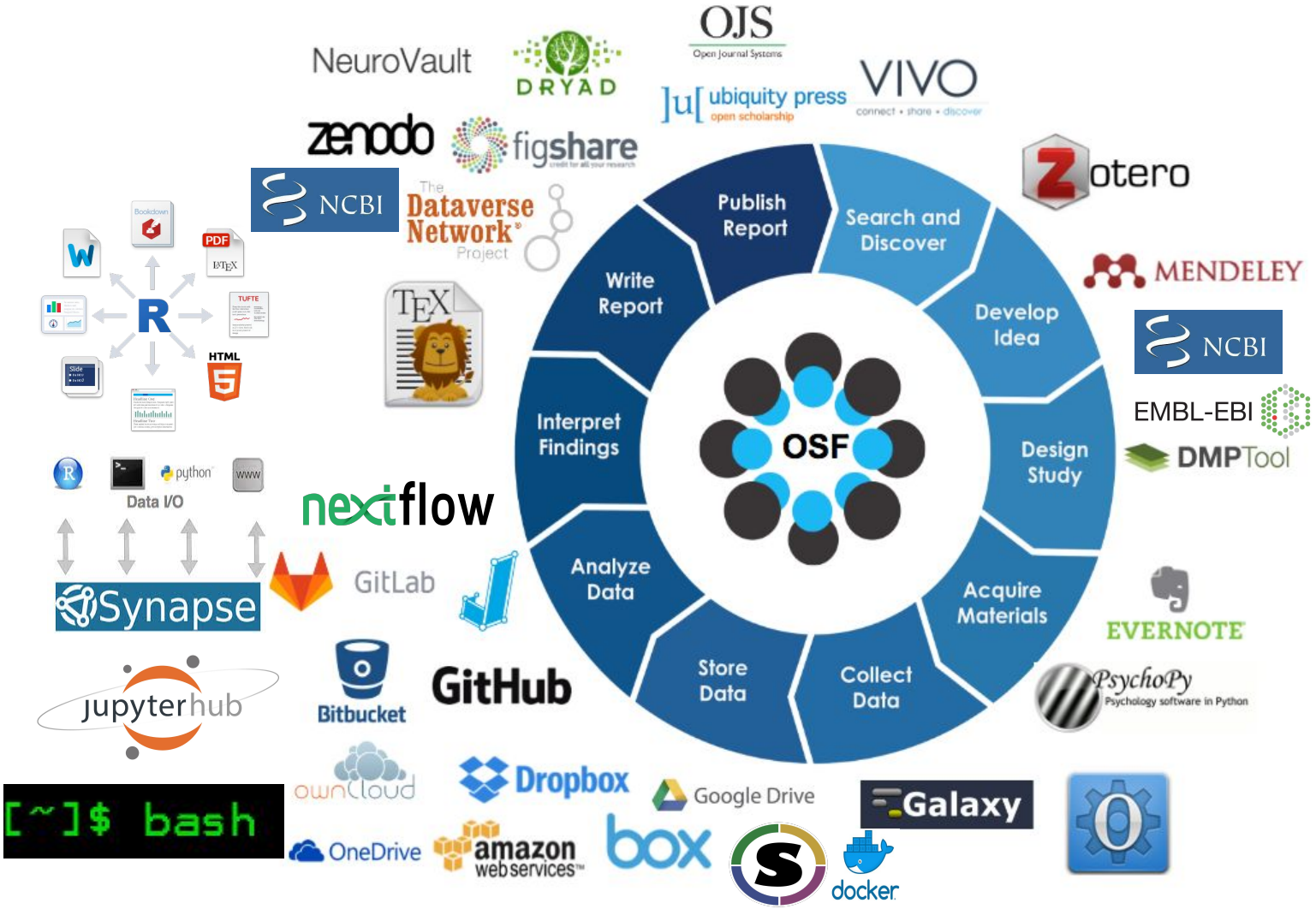
- Lab Notebooks
- ?

3. **Validation**: prepare data and tools for reproducibility and reuse and participate in replication studies

- Presentations
- Pre-prints

4. **Dissemination**: use appropriate licenses for sharing research outputs and report all results and supporting information (data, code, articles, etc.).

5. **Preservation**: deposit research outputs in FAIR archives and ensure long-term access to research results.


NeuroVault
DRYAD
OJS
Open Journal Systems
ubiquity press
open scholarship
VIVO
connect • share • discover
zenodo
figshare
NCBI
The Dataverse Network Project
Zotero
MENDELEY
NCBI
EMBL-EBI
DMPTool
Publish Report
Search and Discover
Develop Idea
Design Study
Acquire Materials
Collect Data
Store Data
Analyze Data
Interpret Findings
Write Report
OSF
R
Bookdown
W
PDF
LaTeX
TUFTE
HTML
Slide
TEX
Data I/O
python
www
nextflow
Synapse
GitLab
EVERNOTE
PsychoPy
Psychology software in Python
jupyterhub
Bitbucket
GitHub
[~]$ bash
owncloud
Dropbox
Google Drive
Galaxy
OneDrive
amazon web services
box
docker

What is your take home?

Book Recommendations

1. Planning and Managing Scientific Research by Brian Kennett, ANU Press, https://www.jstor.org/stable/j.ctt6wp816.6

2. National Academies of Sciences, Engineering, and Medicine. 2018. *Open Science by Design: Realizing a Vision for 21st Century Research*. Washington, DC: The National Academies Press. https://doi.org/10.17226/25116 .


THANK YOU
GRACIAS
ARIGATO
SHUKURIA
JUSPAXAR
BOLZÏN
MERCI
BÏYAN
DANKSCHEEN
TASHAKKUR ATU
SUKSAMA
EKHMET
MEHRBANI
GRAZIE
GOZAIMASHITA
EFCHARISTO
KOMAPSUMNIDA
MAAKE
PALDIES
YAQHANYELAY
SHUKRIA
TINGKI
SPASSIBO
SNACHALHUYA
NUHUN
CHALTU
WABEEJA
MAITEKA
YUSPAGARÀTAM
HUI
GUI
HÀTUR
UNALCHÉESH
DHANYAVAAD
AHHIA
ATTO
MERSI
SPASIBO
DENKAUJA
NENACHALHYA
EKOJU
SIKOMO
MAKETAI
MINMONCHAR
SAINCO
MERASTAWHY
GAEJTHO
LAH
TAVTAPUCH
MEDAWAGSE
AGUYJE
FAKAAUE
BAIKA