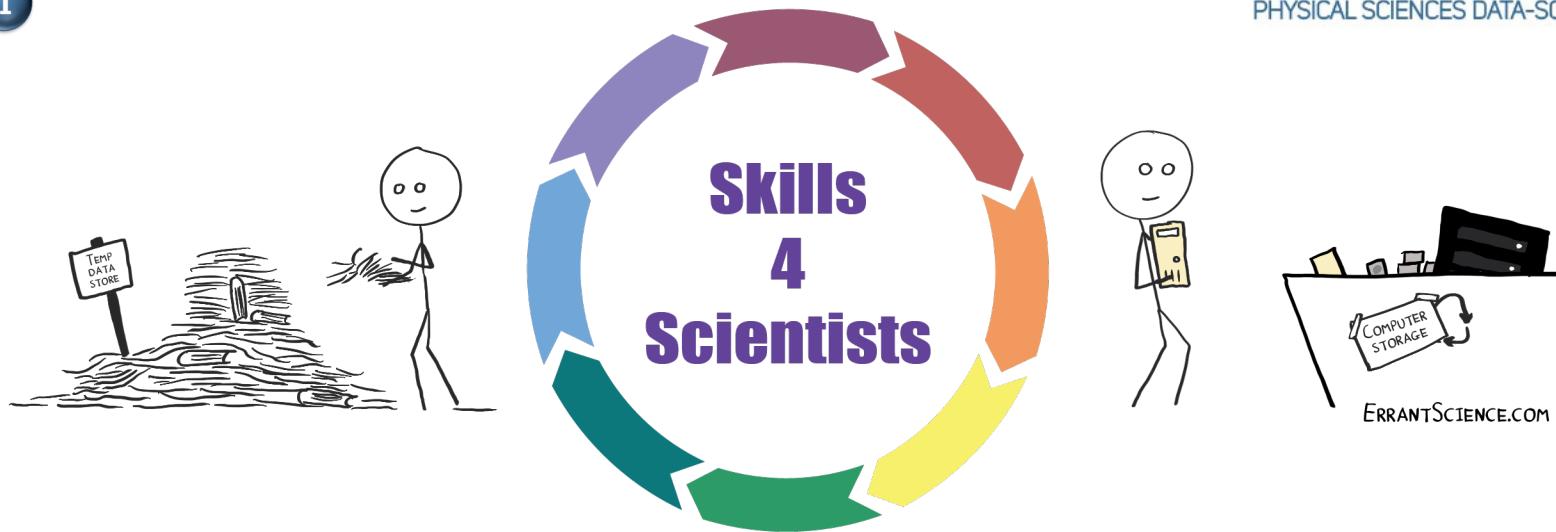


AI 4 Scientific  
Discovery Network<sup>+</sup>



# Collaborative Data Management

Dr Samantha Kanza

# Outline

- Interactive poll
- Step 1: Setup Communication
- Step 2: Identify Data Types
- Step 3: Identify Collaborative Tools
- Step 4: Project Management
- Tips & Tricks

# Collaboration Poll 1

- Have you worked on a collaborative group project before?
  - Yes
  - No



<https://www.kqed.org/mindshift/40797/inside-out-a-look-at-how-emotions-work-together-in-adolescents>

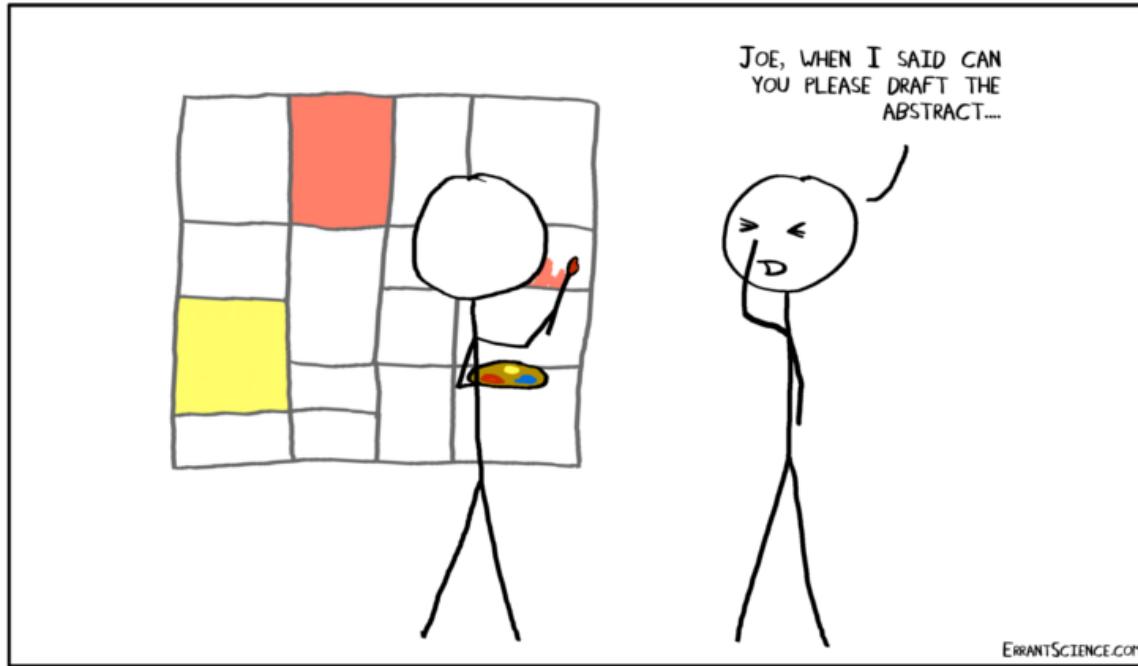
# Collaboration Poll 2

- Would you say that your research data (reports/code/data etc) were collaboratively managed well?
  - Yes
  - No

# Collaboration Poll 3

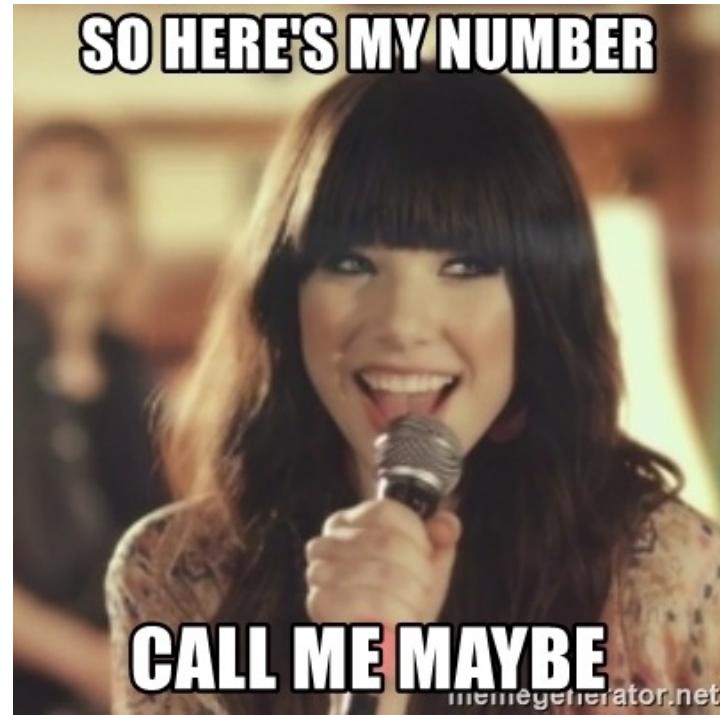
- Have you ever encountered any of these issues during your collaborative work? (Please select all that apply)
  - Being unable to get hold of a teammate
  - Being unable to find shared data/documents
  - Being unable to access shared data/documents
  - Different versions of files have been created and no-one knows which one is the latest one
  - Teammates emailing round documents despite them being on a shared system
  - Teammates making changes offline and not syncing them
  - Teammates forgetting who is doing what and multiple people working on the same job / forgetting other jobs
  - Other (feel free to put these in the zoom chat)

# Step 1: Setup Communication



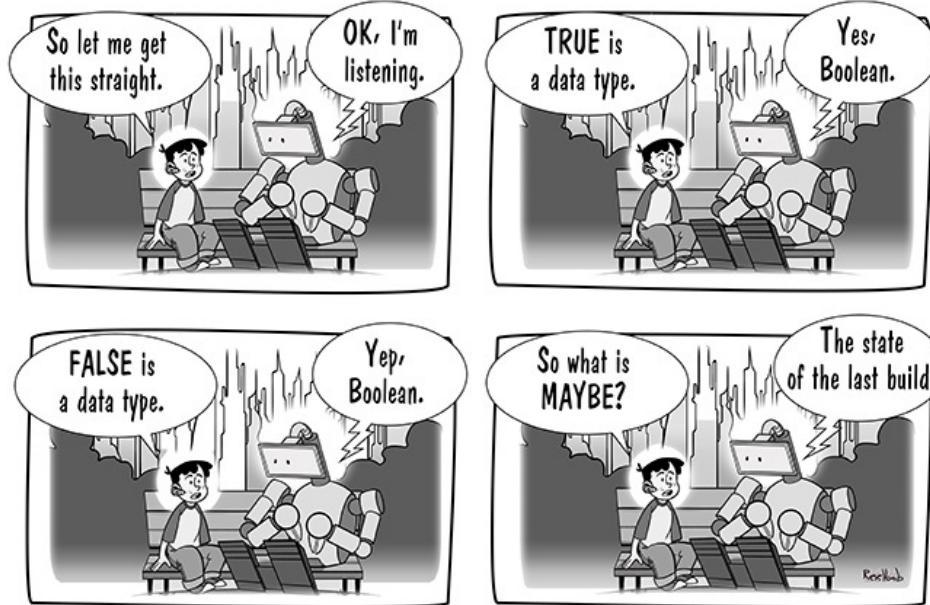
# Communication

- Communication method
  - Teams
  - Slack
  - Facebook Groups
  - Whatsapp Groups
- Exchange Contact details



<https://memegenerator.net/instance/52557245/carly-rae-jepsen-call-me-maybe-so-heres-my-number-call-me-maybe>

# Step 2: Identify Data Types



<https://devops.com/data-types/>

# Data Types

- Reports/Presentations
- Journal Papers/Reports with References
- Datasets
- Code
- Other

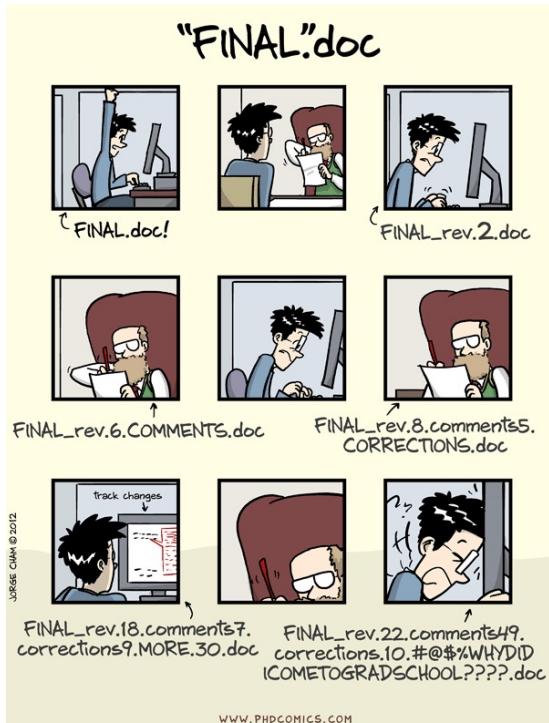
# Step 3: Identify Collaborative Tool(s)



<http://www.thecomicstrips.com/subject/The-Collaboration-Comic-Strips.php>

E-mail: ThavesOne@aol.com  
©2011 Thaves. Dist by Univ Uclick for UFS, Inc.

# Collaborative Tools: The Benefits



- Streamline communication
- Eliminate messy document & email chains
- Real-time editing for multiple users
- Helps share workload
- Speeds up collaboration

[http://phdcomics.com/comics/archive\\_print.php?comicid=1531](http://phdcomics.com/comics/archive_print.php?comicid=1531)

# Collaborative Tools

- Data Storage
  - MS Teams/One Drive, Google Drive, Dropbox, GitHub
- Collaborative Editing
  - Office 365, Google Apps, Dropbox Paper, Overleaf
- Version Control
  - GitHub
- Reference Management
  - Zotero, Mendeley, EndNote

# Collaborative Tools Comparison

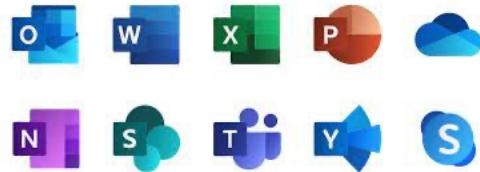
Data Type	MS Teams / Office 365	Google Drive / Google Apps	Dropbox / Dropbox Paper	Overleaf/LaTeX	GitHub / Other Version Control
Presentations	✓✓✓	✓✓✓	✓✓✓	✓✓	✓✓
Reports	✓✓✓	✓✓	✓✓✓	✓✓	✓✓
Journal Papers	✓✓✓	✓✓✓	✓✓✓	✓✓	✓✓
Datasets	✓✓✓	✓✓✓	✓✓	✓	✓✓
Code	✓	✓	✓	✓	✓✓
Media (Images/Videos/Sound)	✓	✓	✓	✓	✓✓
Links with Reference Manager	Yes	Limited	Limited	Export	Export

- ✓ Store
- ✓ Collaboratively edit in real time
- ✓ Versioning / Version Control

# WYSIWYG – Common Features

- Multiple platforms (online, desktop & mobile)
- Simultaneous Editing
- Import & Export from multiple formats
- Integration with reference managers
- Comment & track changes
- Version control / History

# WYSIWYG Tools – Office 365



- Available to students through office in education
- Powerful editing features
- Allows checkout to desktop programs and syncing to your hard drive
- Simultaneous editing can be temperamental
- Integration between office 365 programs including storage & project management tools
- Links with MS Teams
- Comments/Chats/Conversations

# WYSIWYG Tools – Google Drive



Docs



Sheets



Slides



Forms

- Accessible to anyone with a google account
- Also allows online & offline editing
- Best for simultaneous editing
- Doesn't have all the features that may be available in Office
- Has its own native format
- **Not EEA Compliant!**

# WSYIWYG Tools – Dropbox Paper



- Accessible to anyone with a dropbox account
- Also allows online & offline editing
- Can link with a number of apps and file types
- Doesn't have all the features that may be available in Office
- Has its own native format
- Inbuilt project management tools

# Other Tools

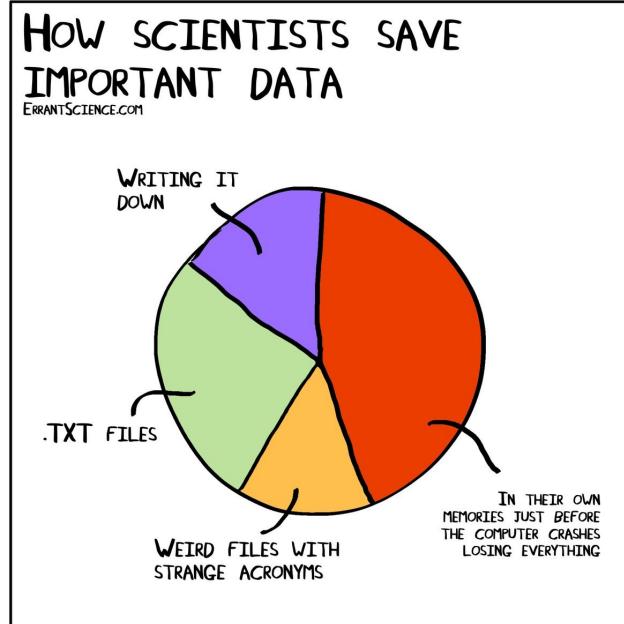
- Reference Managers
- Overleaf
  - Collaborative LaTeX editor
- GitHub
  - Version Control
- Tune in on the 21<sup>st</sup> to find out more!

# Presentations/Reports/Journal Papers

- Do you want to use a WYSIWIG or LaTeX?
- If this is a journal paper, how are you submitting it?
- Do you need to reference literature?
- Is your reference manager/format compatible with your tool of choice

# Datasets

- What formats are the data in?
- Is the data auto generated or team generated?
- Will the data be regularly edited?



# Code

- USE VERSION
- CONTROL!

COMMENT	DATE
CREATED MAIN LOOP & TIMING CONTROL	14 HOURS AGO
ENABLED CONFIG FILE PARSING	9 HOURS AGO
MISC BUGFIXES	5 HOURS AGO
CODE ADDITIONS/EDITS	4 HOURS AGO
MORE CODE	4 HOURS AGO
HERE HAVE CODE	4 HOURS AGO
AAAAAAA	3 HOURS AGO
ADKFJSLKDFJSDFKLJ	3 HOURS AGO
MY HANDS ARE TYPING WORDS	2 HOURS AGO
HAAAAAAAAANDS	2 HOURS AGO

AS A PROJECT DRAGS ON, MY GIT COMMIT MESSAGES GET LESS AND LESS INFORMATIVE.

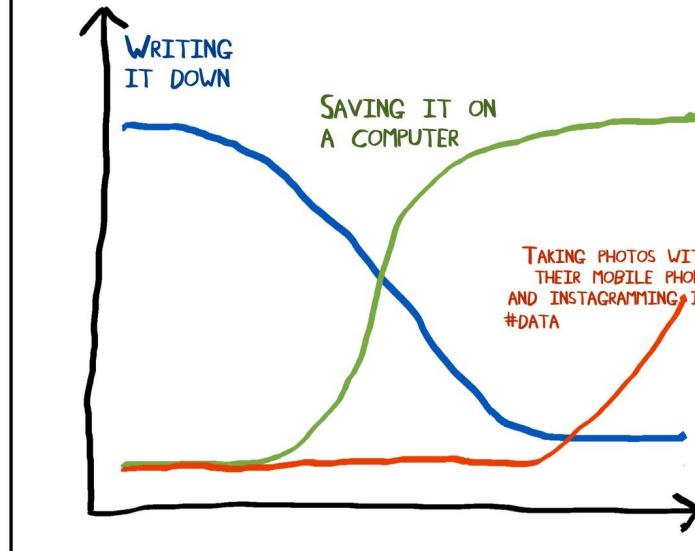
<https://www.lpalmieri.com/posts/2018-09-14-machine-learning-version-control-is-all-you-need/>

# Media

- Store with the relevant files!

## A TIMELINE OF RESEARCHERS RECORDING DATA

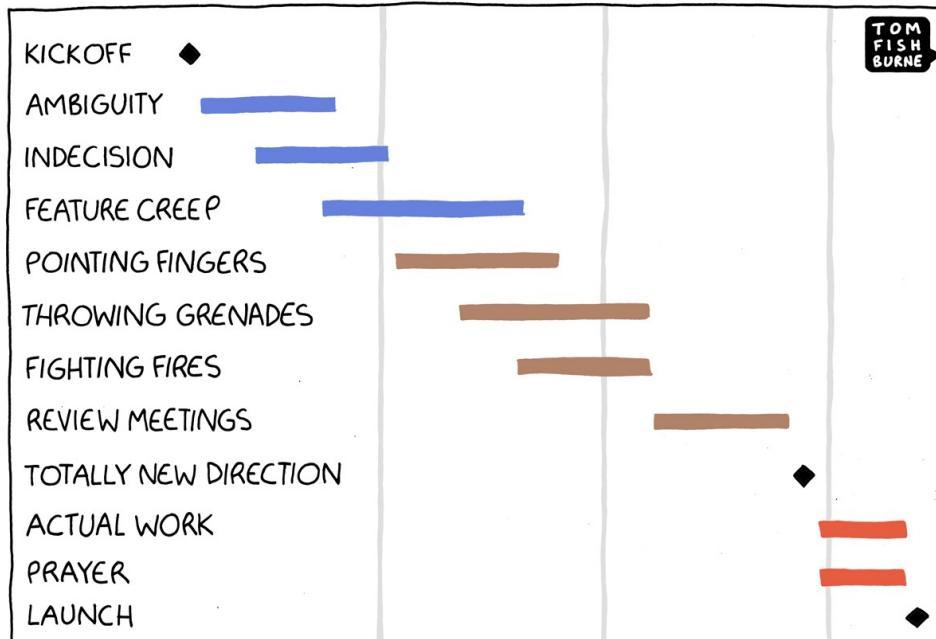
ERRANTSCIENCE.COM



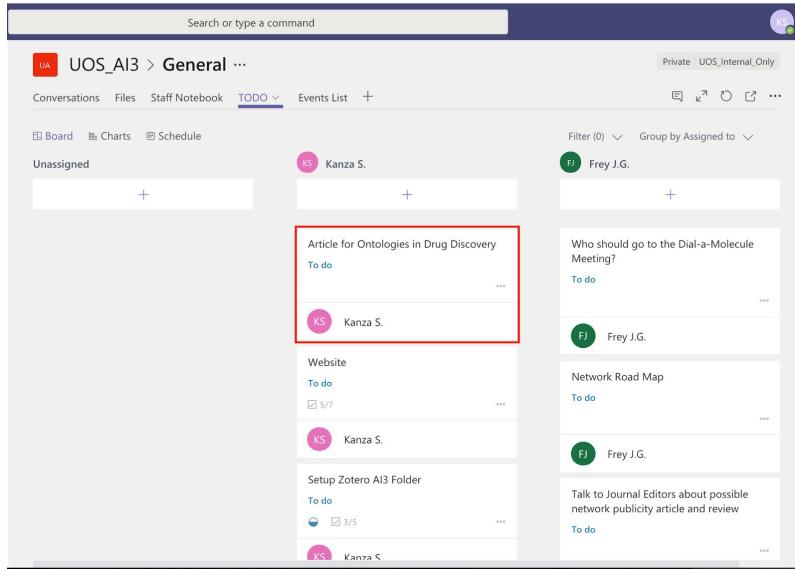
# Collaborative Tools: Recommendation

Data Type	MS Teams / Office 365	Google Drive / Google Apps	Dropbox / Dropbox Paper	Overleaf/LaTeX	GitHub / Other Version Control
Presentations	✓	✓	✓		
Reports	✓	✓	✓	✓	
Journal Papers	✓			✓	
Datasets	✓	✓	✓	✓	✓
Code					✓
Media (Images/Videos/Sound)	✓	✓	✓	✓	✓

# Step 4: Project Management



# Task Management – MS Teams



Search or type a command

UOS\_AI3 > General ...

Conversations Files Staff Notebook TODO Events List +

Board Charts Schedule

Unassigned

Kanza S. Article for Ontologies in Drug Discovery To do

Kanza S. Who should go to the Dial-a-Molecule Meeting? To do

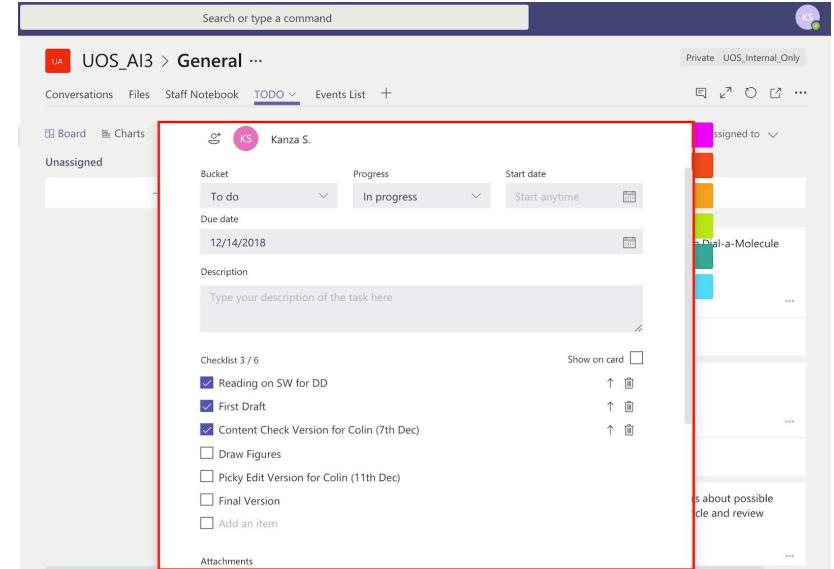
Frey J.G. Network Road Map To do

Frey J.G. Talk to Journal Editors about possible network publicity article and review To do

Website To do 5/7

Kanza S. Setup Zotero AI3 Folder To do 3/5

Kanza S.



Search or type a command

UOS\_AI3 > General ...

Conversations Files Staff Notebook TODO Events List +

Board Charts

Unassigned

Kanza S. Article for Ontologies in Drug Discovery To do

Kanza S. Who should go to the Dial-a-Molecule Meeting? To do

Frey J.G. Network Road Map To do

Frey J.G. Talk to Journal Editors about possible network publicity article and review To do

Website To do 5/7

Kanza S. Setup Zotero AI3 Folder To do 3/5

Kanza S.

**Kanza S.**

Bucket: To do Progress: In progress Start date: Start anytime

Due date: 12/14/2018

Description: Type your description of the task here

Checklist 3 / 6

- Reading on SW for DD
- First Draft
- Content Check Version for Colin (7th Dec)
- Draw Figures
- Picky Edit Version for Colin (11th Dec)
- Final Version
- Add an item

Attachments

# Task Management – Trello

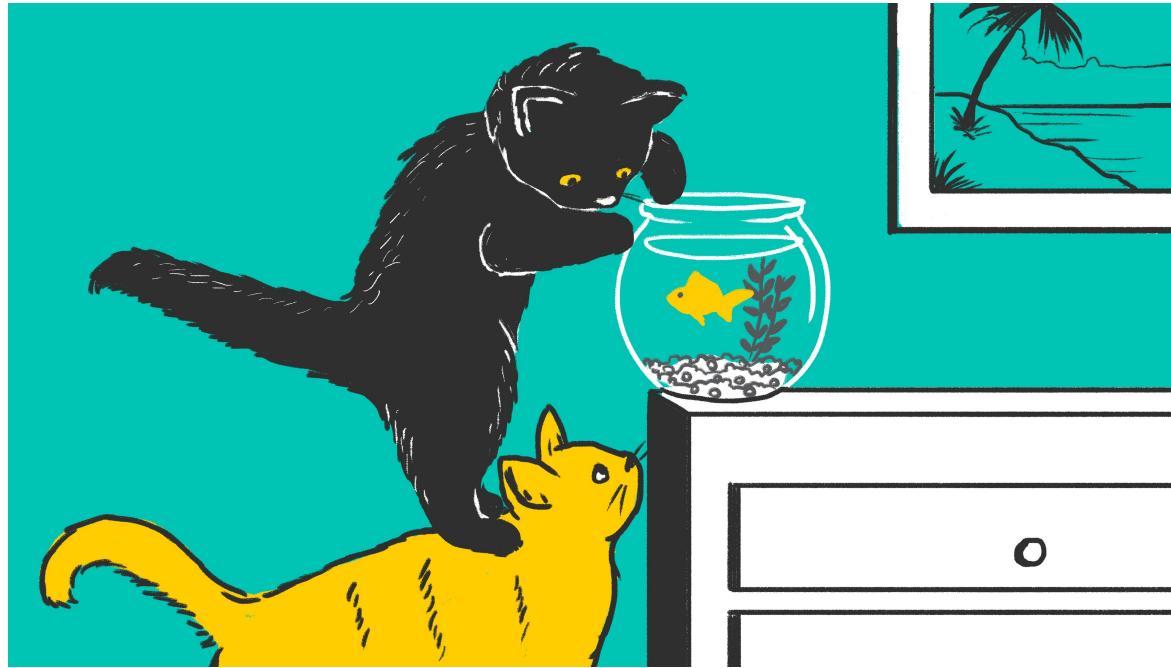
**ToDo**

- Evaluation Section
- New Intro
- Conduct Focus Groups
- Add safety Info
- Add new coding
- Transcribe and Analyse Focus Groups
- Enter a title for this card...

**Done**

- + Add another list
- Re-do outline from Feedback
- Ethics for Focus Group
- Introduction
- Literature Review
- Tak2Lab Background
- Tak2Lab System
- Figures

# Considerations in Collaborative Work



<https://www.thedesigngym.com/skip-collaboration-collaborations-sake/>

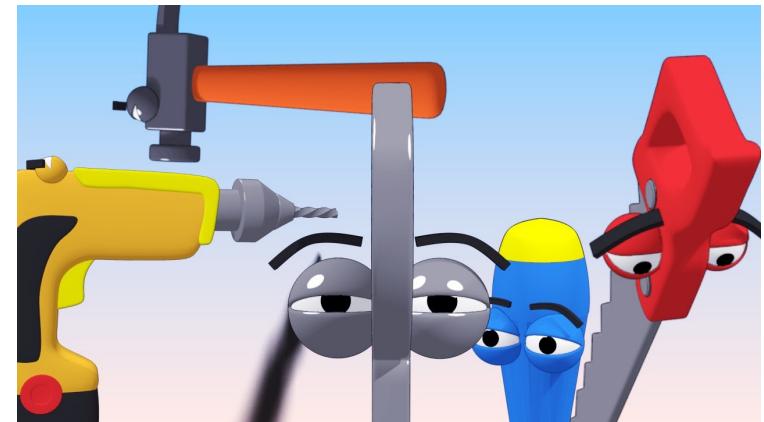
# Task Management

- Organise yourselves!
- ToDo Lists / Assign Tasks
- Agree small manageable deadlines
- Regular Meetings
- COMMUNICATION IS KEY!



# Don't be afraid to use more than one tool

- But only when its necessary!
- Document this!



[https://www.youtube.com/watch?v=BaKcwU1lt3c&ab\\_channel=AApV-VidsForKids](https://www.youtube.com/watch?v=BaKcwU1lt3c&ab_channel=AApV-VidsForKids)

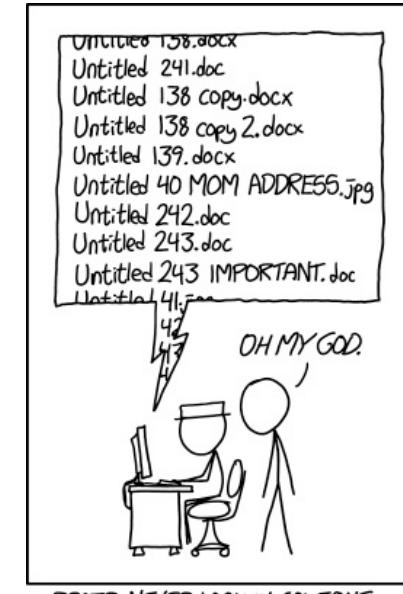
# Choose tools you can all use!



<https://www.insider.com/the-lion-king-differences-between-remake-and-original-animated-movie-2019-7>

# Project/Document Structure

- Sensible File/Folder Structure
- Naming Conventions
- Template Folders
- Template Documents
- [www.ai3sd.org/documents](http://www.ai3sd.org/documents)



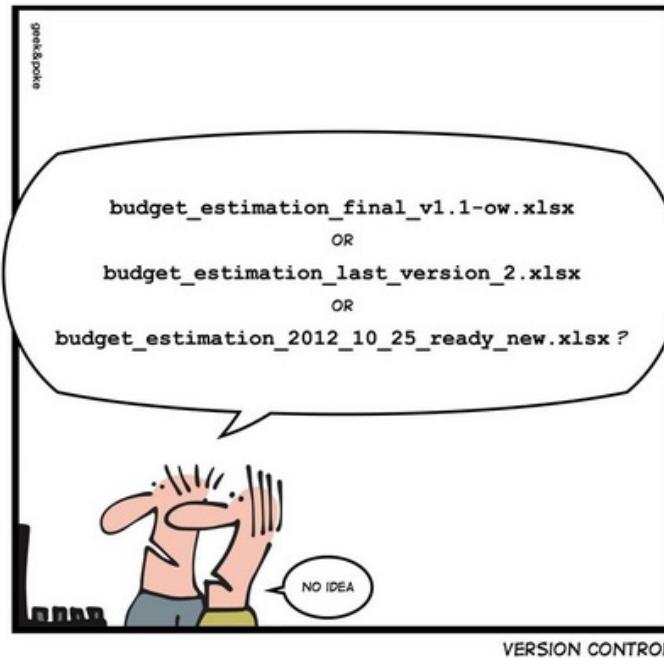
<https://sites.allegheny.edu/lits/tutorials/file-naming-conventions/>

# Access/edit permissions



# Version Control / Tracking Changes

## SIMPLY EXPLAINED



<https://www.groovecommerce.com/ecommerce-blog/guide-to-version-control-for-magento-using-git-and-beanstalk/>

# Conclusions

- Communication & Organisation is key
- Use the tools that:
  - Work for your data types
  - Can be used and understood by all the team
- Don't be afraid to sometimes combine tools
- Data isn't just data

# Conclusions

- Fostering a collaborative attitude is just as important as the tools!



<https://thegouldstandardnyu.files.wordpress.com/2015/11/group-project.jpg>

# Useful Links

- <https://drive.google.com>
- <https://support.google.com/drive>
- <https://www.office.com/>
- <https://support.office.com/>
- <https://products.office.com/en-us/microsoft-teams/group-chat-software>
- <https://www.dropbox.com/>
- <https://www.dropbox.com/paper>
- <https://www.overleaf.com/>
- <https://trello.com/en>

# Skills4Scientists!

