BIAS AND REPRODUCIBILITY IN A COMPUTATIONAL NEUROBIOLOGY PHD'S JOURNEY

SUSANA ROMAN GARCIA, NEUROSCIENCE, CENTRE FOR BRAIN DISCOVERIES (PHD), UNIVERSITY OF EDINBURGH.

AN OVERALL QUESTION TO ASK OURSELVES WHEN REPORTING IS:

HOW WILL MY RESEARCH BE USED IN THE FUTURE?

4TH REPORTING:

PUBLISH OPEN ACCESS!

PUBLISHING UNDER AN OPEN LICENCE ALLOWS RESEARCH TO BE ACCESSED BY ANYONE FOR FREE.

SHARE AND LICENCE YOUR RESEARCH

ASSIGN DOI NUMBERS

FIND SUITABLE REPOSITORIES BY SUBJECT

Re3data.org
fairsharing.org

SHOW AND SHARE WITH EVERYONE ELSE...

POSITIVE RESULTS BIAS

NOT ALL IDEAS AND DATA GET PUBLISHED. IN THE CURRENT SCIENTIFIC CULTURE, NOVEL AND POSITIVE RESULTS ARE CONSIDERED MORE PUBLISHABLE THAN REPLICATIONS AND NEGATIVE RESULTS.

3RD DATA ANALYSIS:

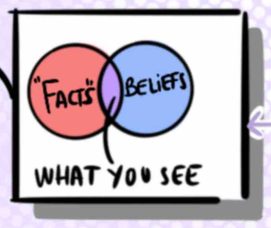
LOOK AT WHAT YOU COLLECTED...

PLAN FOR DATA ANALYSIS REPRODUCIBILITY BY:

- INSTALLABILITY
- CODE (RE)USAGE
- REPRODUCTION OF ENVIRONMENT
- STATISTICAL TRANSPARACENCY

DURING DATA ANALYSIS , THE % OF ASSUMPTIONS INCREASES AS WE GO AHEAD TO MAKE PREDICTIONS/TEST HYPOTHESES.

-WHAT'S MY CONFIRMATION BIAS?
-AND HISTORICAL BIAS?



WATCH OUT FOR:

WHO IS MY RESEARCH SERVING?

-AM I CONTINUING A PRE-EXISISTING BIAS THAT I
-HAD NOT THOUGHT ABOUT?
-RACIST, SPECIESIST, ABLEIST?

RESEARCHER DEGREES OF FREEDOM

RESEARCHERS CAN CHOOSE BETWEEN MULTIPLE WAYS OF COLLECTING AND ANALYSING DATA; THESE DECISIONS CAN BE MADE ARBITRARILY OR BECAUSE, PERHAPS, THEY PRODUCE A POSITIVE AND STATISTICALLY SIGNIFICANT RESULT.

WHEN CHECKING FOR BIAS IN DATA COLLECTION, YOU CAN ASK QUESTIONS LIKE:

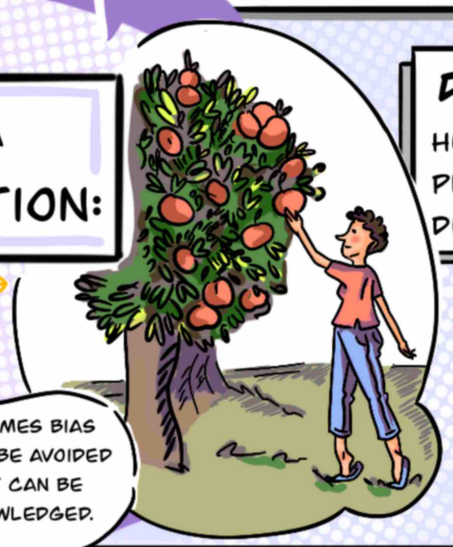
-WHERE IS MY DATA COMING FROM?

DATA IS NOT JUST NUMBERS. IT COMES FROM LIVING BEINGS THAT LIKELY GAVE THEIR LIVES FOR OUR BENEFIT.
THERE IS AN ETHICAL BIAS HERE.

THINK OF WHERE YOU'RE KEEPING THE DATA AND HOW WILL YOU SHARE IT LATER ON.

2ND DATA COLLECTION:

COLLECT THE FRUIT...



SOMETIMES BIAS CANNOT BE AVOIDED BUT IT CAN BE ACKNOWLEDGED.

DATA MANAGEMENT PLANS

HELP TO KEEP DOCUMENTS OUTLINING HOW YOU ARE PLANNING TO MANAGE YOUR RESEARCH DATA BOTH DURING AND AFTER YOUR RESEARCH PROJECT.

WHAT RESOURCES AM I USING TO COLLECT DATA? SUPERCOMPUTERS, MATHEMATICAL CALCULATIONS?

FINDABLE
ACCESIBLE
INTEROPERABLE
REUSABLE

BE TRANSPARENT WITH EACH STEP AND TYPE OF (META)DATA.

1ST DESIGN:

START HERE!

THIS PART OF THE PROJECT IS KEY TO LAY OUT *INITIAL BIAS* AND PREVENT FURTHER BIAS FEEDBACK. ASK QUESTIONS LIKE:

A PHD JOURNEY IS SIMILAR TO GROWING AND SHARING YOUR OWN FOOD...

-WHO IS MY DATA INCLUDING/EXCLUDING?
-WHAT AM I ASSUMING?
-HOW MUCH AM I SIMPLIFYING?

...START WITH FERTILE SOIL...

PLAN FOR REPRODUCIBILITY AND REPLICABILITY.

DATA AND CODE REPOSITORIES

DEPOSITING DATA AND CODE IN OPEN REPOSITORIES ONLINE ENABLES ANYONE TO ACCESS THESE RESEARCH ITEMS

- GitHub
- OSF
- Zenodo

ANALYSIS \ DATA	=	≠
=	REPRODUCIBLE	REPLICABLE
≠	ROBUST	GENERALISABLE

VERSION CONTROL IS YOUR FRIEND!

"CODE LITERACY"

WRITE CODE AS IF YOU WERE WRITING IT FOR SOMEONE ELSE TO REUSE LATER ON.

SCAN TO CONNECT AND SEE RESOURCES MENTIONED IN THE POSTER IN MORE DETAIL

THE UNIVERSITY OF EDINBURGH
MRS Funding a healthier future
MEDICAL RESEARCH SCOTLAND
The Alan Turing Institute
CERN