

## 3DVC\_Green3\_Data

Scribe: Spencer Bliven

Big data often published, but small plots often only published as JPEG figure

Need data to be published, eg SBML, which can be imported into lots of packages.

Don't only need data standards, but also need to standardize data acquisition (or at least annotate the acquisition process)

Problem: pressure to have short methods section.

Journals should require detailed methods in supplemental section

Methods are continually developed and changed, but keeping up with changing acquisition methods is very tough..

Without a common way of describing data, how do we start? Many kinds of data (eg binary reactions vs timecourse vs etc in kinetic modelling)

Microscopy looks like animated simulations, but very different formats, so can't use same tools.

Evolving so fast: in past, never would have expected synapse-level neural images, but now trivial. Hard to be flexible enough to include all future data

Sometimes DON'T want to keep all data (eg simulation intermediates)

Some very basic annotation still lacking, eg units (eg mM for "millimolar" wrong)

How do we correct errors? 1) peer review, or 2) crowd source  
Journals should assume responsibility since they are paid for it

Is there a need for post-publication annotation of models? When models make incorrect predictions, how do we link that evidence? Can models be reused passed a single PhD thesis?

Kinetic models especially susceptible to version rot due to fast evolution

Since biologists usually responsible for simulation software, often use outdated software methods.

Need to use standard libraries for things like ODE solvers

NIH could provide API for basic simulation processes. Then everyone would use the most up-to-date simulator. Plus, could cache results, provide versioning, publication, etc.  
But sometimes don't need to save models, since specific to one condition  
For instance many basic systems like hemoglobin binding are general, but so old that they aren't in modern formats

Ideal: Siri, how does this change the cell?

Multi-scale modelling: How do we relate parameters at different scales?  
Use bayesian parameters, do sensitivity analysis

Relate parameters to intrinsic variability