

3DVC_WorkingGroup_Infrastructure_Full

Scribe: Spencer Bliven

Chris McClendon
Andrew McCulloch
Tim Clark
Tony Hunt
Wilfred Li
Larissa
Tom B
Yolanda Gill
Steven Peltier
Anita
Jim S
Blake
Fritz Herberg
Susan Tayler
Maryann Martone
...

- Infrastructure should unify experiment and theory (simulation)
 - Impossible to find all parameters for complex models, but often can estimate them just by physical constraints from geometry, physics, etc
- How to make modeling part of the experiment
- Need to go beyond cytoscape
- Structures and models as data - errors persistence
- Models need to be portable between formats, annotatable and already documented (interoperable, literate)
- Reproducibility - does it matter?
 - Replication (exact rerun) vs reproducibility vs inspectability
 - reproducible is less stringent but more important
 - version control infrastructure important
 - not all models equally valuable in perpetuity (need to discard data eventually)
 - Just OS changes and version rot obsoletes some data
 - But while not replicable, should be reproducible
 - Expensive studies generally not repeated, but may be reproduced
 - Should expect reproducibility for some time period (eg 5 years)
 - Community should describe timeframe
 - Longest time frame for your data is your lifetime
 - Modulated by cost of data storage
 - Where data not stored, should at least be reproducible
- A coat rack where people hang their data
- Continuum of scales - enter at any point and understand concept and content in both directions

YG: What are CS challenges?

TB: Frustrated by tiny steps

Many different user groups, many of whom prefer UI to compiling code, despite need for OSS

We should look for problems where its 1964 and we can reach the moon. Or where sputnik just flew over and we need to get moving.

TB: "Scientists have 2 important traits: curiosity and persistence"

AM: With existing tools and theory, what problems should we have solved but haven't?
"Entire cell" may take away from biology problem, since most biologists really care about some specific problem. Need to map specific problems against existing technologies

JS: but whole cell interesting since includes localization problems

CM: Get info about molecular pathology

YG: What existing baby steps are out there?

TB: See Terry's talk this morning

MM: Lots of things change when scale changes (coordinate system)

We have some spacial knowledge, but it's incomplete

B: Determining localization of cell components could be a **tractable first step**.

Just putting model together can show holes in knowledge. How much space in the cell is unknown?

TH: biology is about behaviour. Infrastructure should enable adding behaviour.

AM: Biologists are interested in what happens and why.

L: Virtual cell should allow virtual experiments. Predict data

MM: Very difficult to find parameters (literature search), and no progress on that
Impossible to experimentally get parameters

TB: Useful to visualize protein structures in many ways at many resolutions.

With 3D part of 3D virtual cell can make many fundamental predictions

CM: Want to do coarse simulation, then take a snapshot to initialize fine simulation

Tim & Tom: Infrastructure should support "virtual witnessing" (steven chapen?)—doctors

would sit in operating theater and all observe the experiment/operation.

SB: how differs from Openness?

Witnessing = Simultaneous observation