Enhanced Sampling Techniques

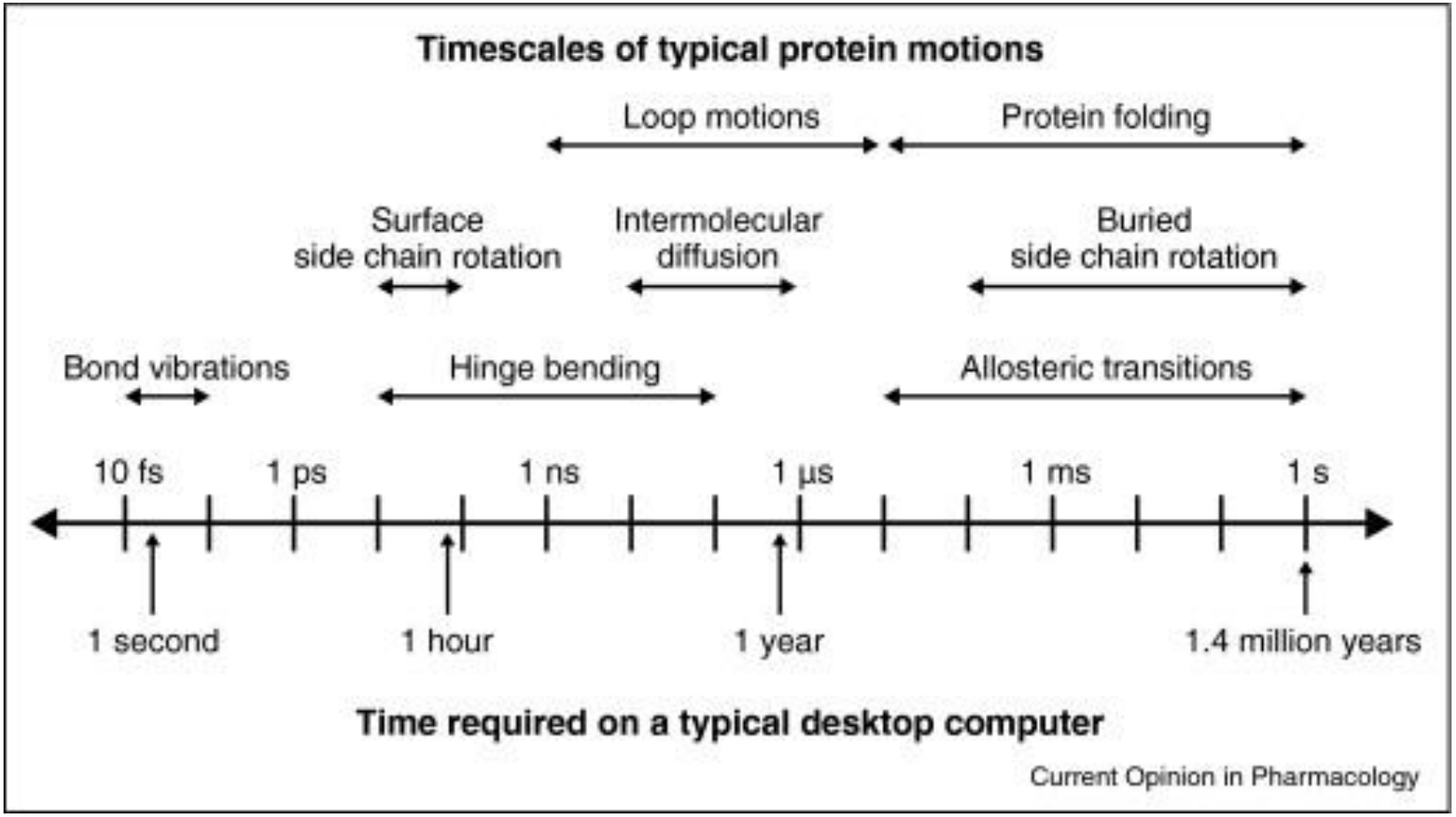
https://wkkrqsriursybsor-zippykid.netdna-ssl.com/wp-content/uploads/2014/10/costco-samples.jpg

QMMM Study Group
October 5th, 2018
Anthony Bogetti

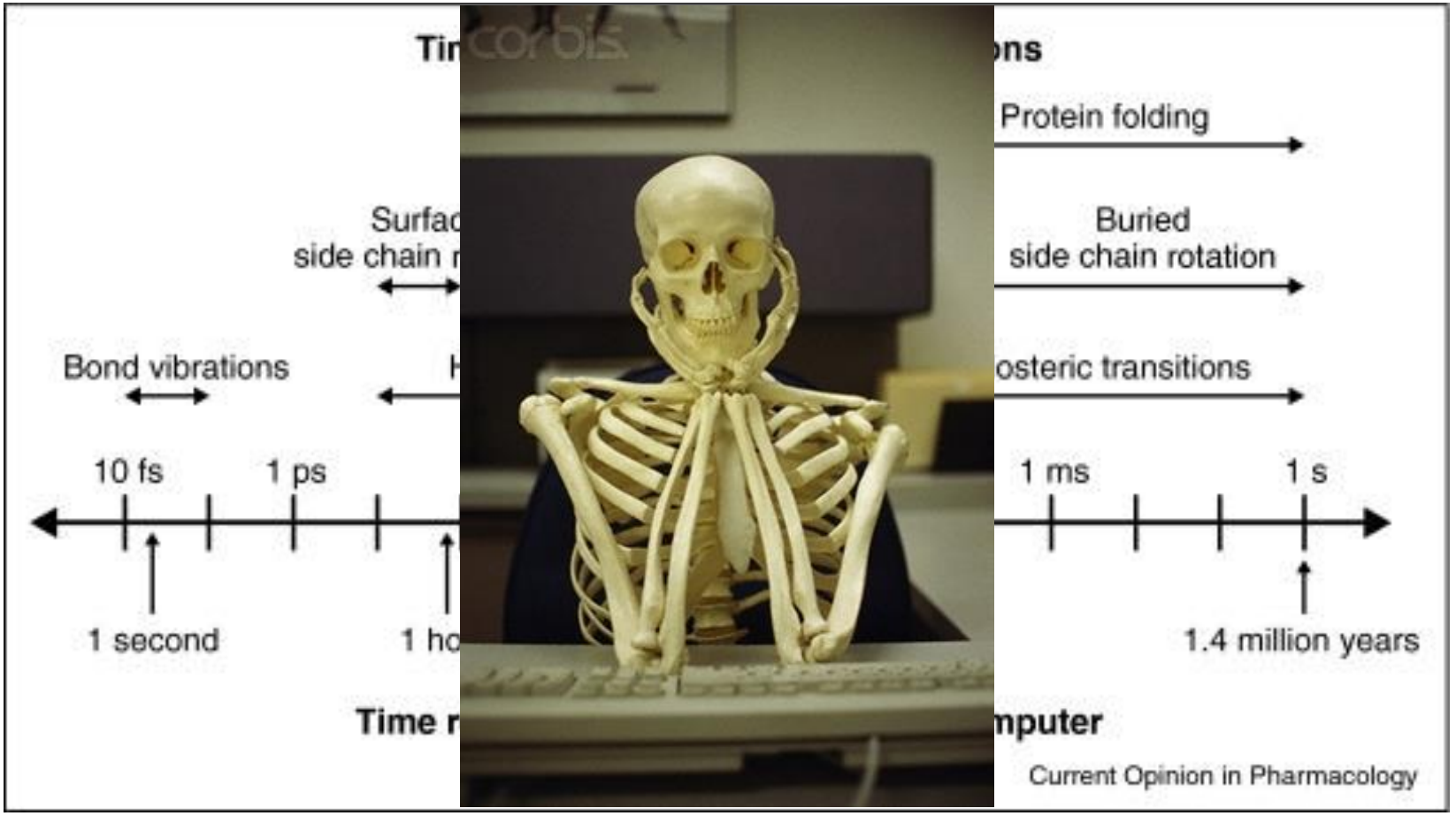
References

- Zwier, M. and Chong, L. Reaching biological timescales with all-atom molecular dynamics simulations. *Curr. Opinion in Pharma.* 10: 6, 2010, 745-752
- Zuckerman, DM. and Chong LT. Weighted Ensemble Simulation: Review of Methodology, Applications, and Software. *Annu Rev Biophys. 2017. 46, 43-57.*

Biological Timescales



Biological Timescales



https://i.imgur.com/hniy4My.jpg

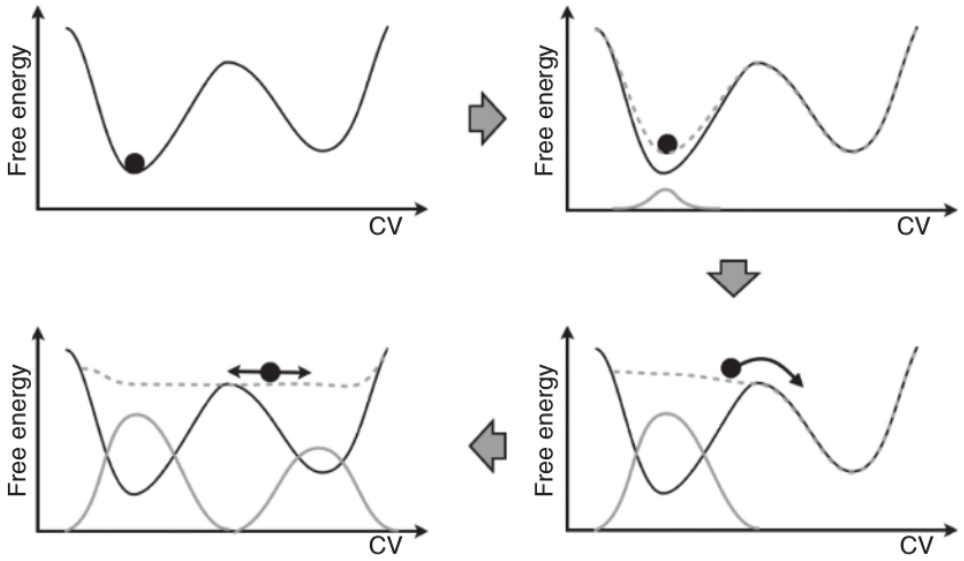
Separation of Timescales

- A central assumption of all ES techniques is that events are rare and waiting times between events are long
- How can the long wait times be eliminated?

http://blog.tutorhub.com/wp-content/uploads/2013/06/shutterstock_68955784.jpg

Altering the Energy Landscape

- Some techniques will alter the free energy landscape of a process to capture a rare event sooner.
- Examples include metadynamics (pictured below) and steered MD



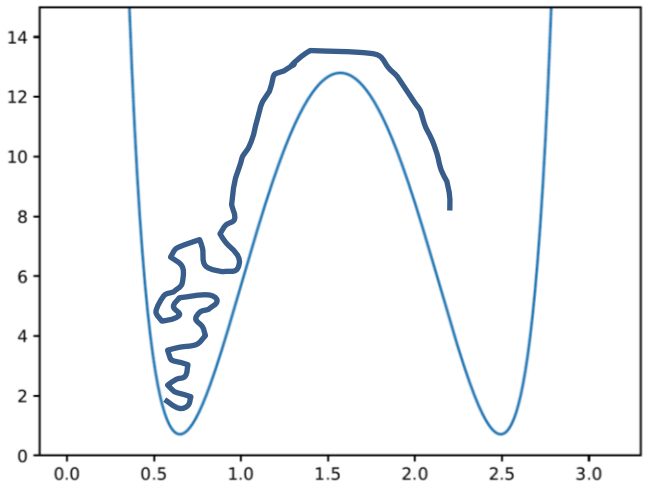
http://www.parrinello.ethz.ch/research/metadynamics/_jcr_content/par/fullwidthimage/image.imageformat.fullwidth.1393180090.png

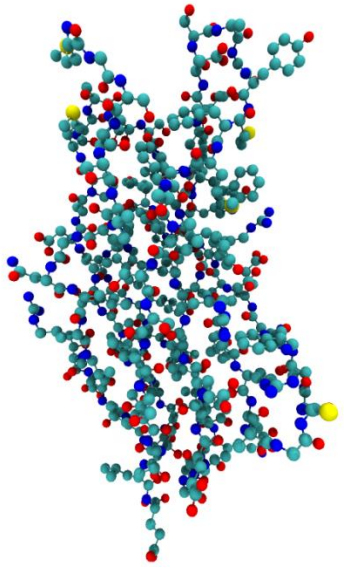
Altering the Energy Landscape

- Some techniques will alter the free energy landscape of a process to capture a rare event sooner.
- Examples include metadynamics (pictured below) and steered MD
- This can be useful for exploring the *thermodynamics* of a process, but what if we are interested in *kinetics*?

Path (Rare Event) Sampling

- Find the detailed dynamics of pathways between metastable states
- Kinetic observables can be determined
- Examples include transition path sampling, forward flux sampling and weighted ensemble.





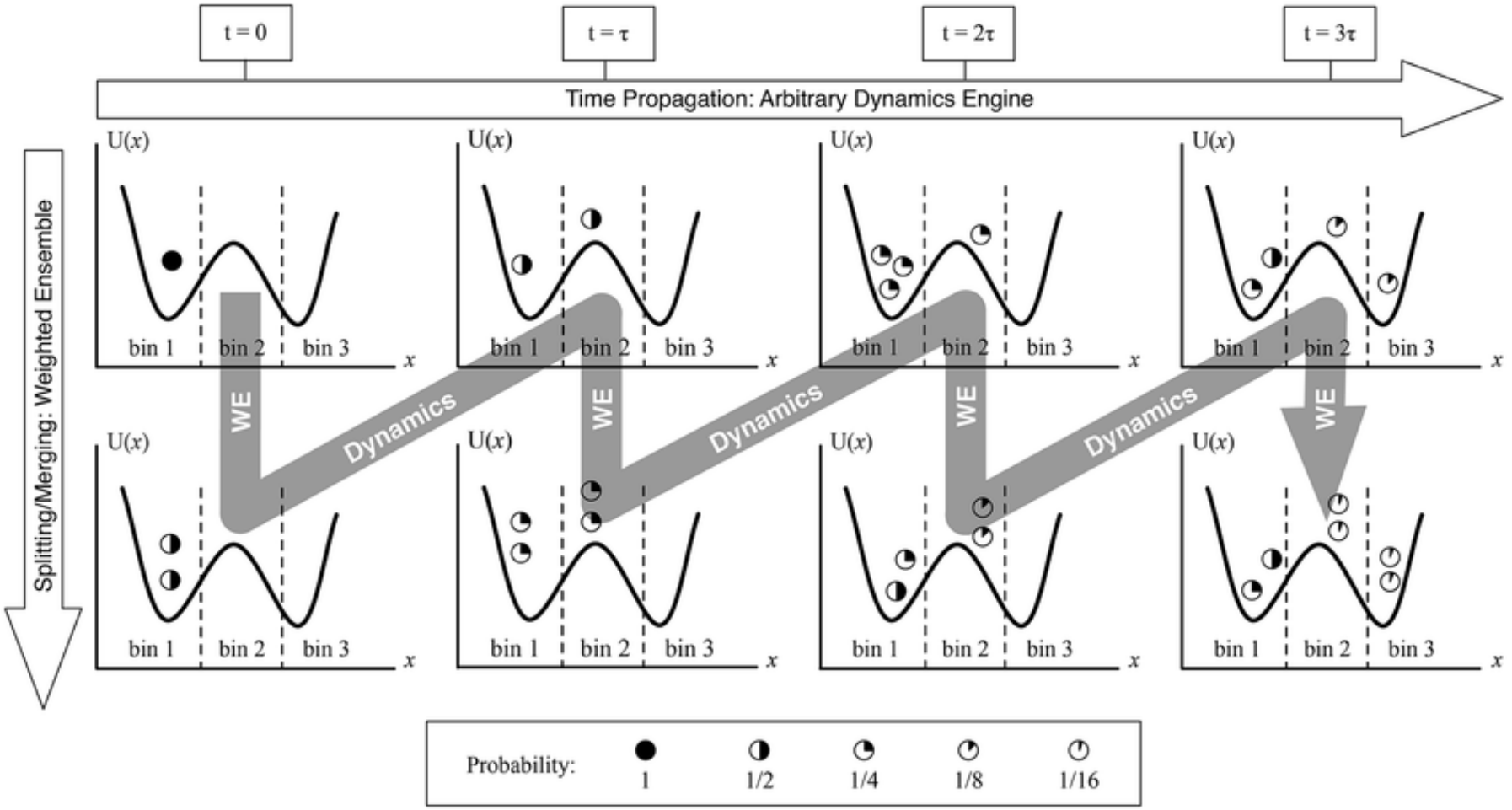
Weighted Ensemble

- Take your configuration space (determined largely by one or more **order parameters**) and divide it up into **bins**
- Start a fixed number of simulations in a bin (called **walkers**), propagate dynamics for each walker
- Assign walkers to bins based off of their calculated order parameter
- Every **iteration**, propagate dynamics and assign to bins

Weighted Ensemble

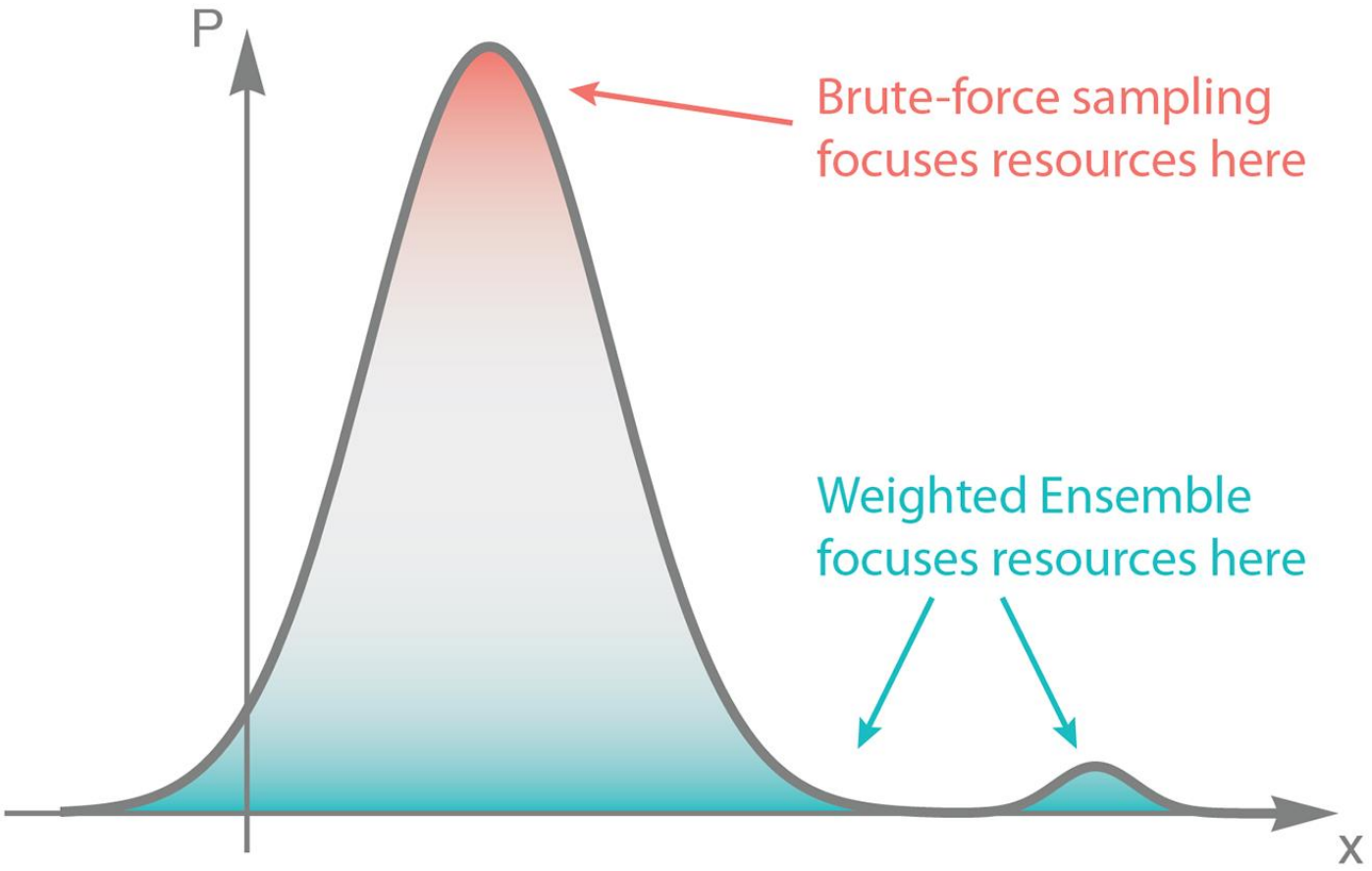
- Each walker (individual simulation) has a certain probability associated with it (starts with a probability of 1)
- There is a limit to the number of walkers that can be in a bin
- If there are more walkers in a bin than the target count, walkers are **merged** (probabilities are combined)
- If there are fewer than the target count, walkers are **split** (probabilities are divided)
- **Splitting and merging is the key to WE**

Weighted Ensemble



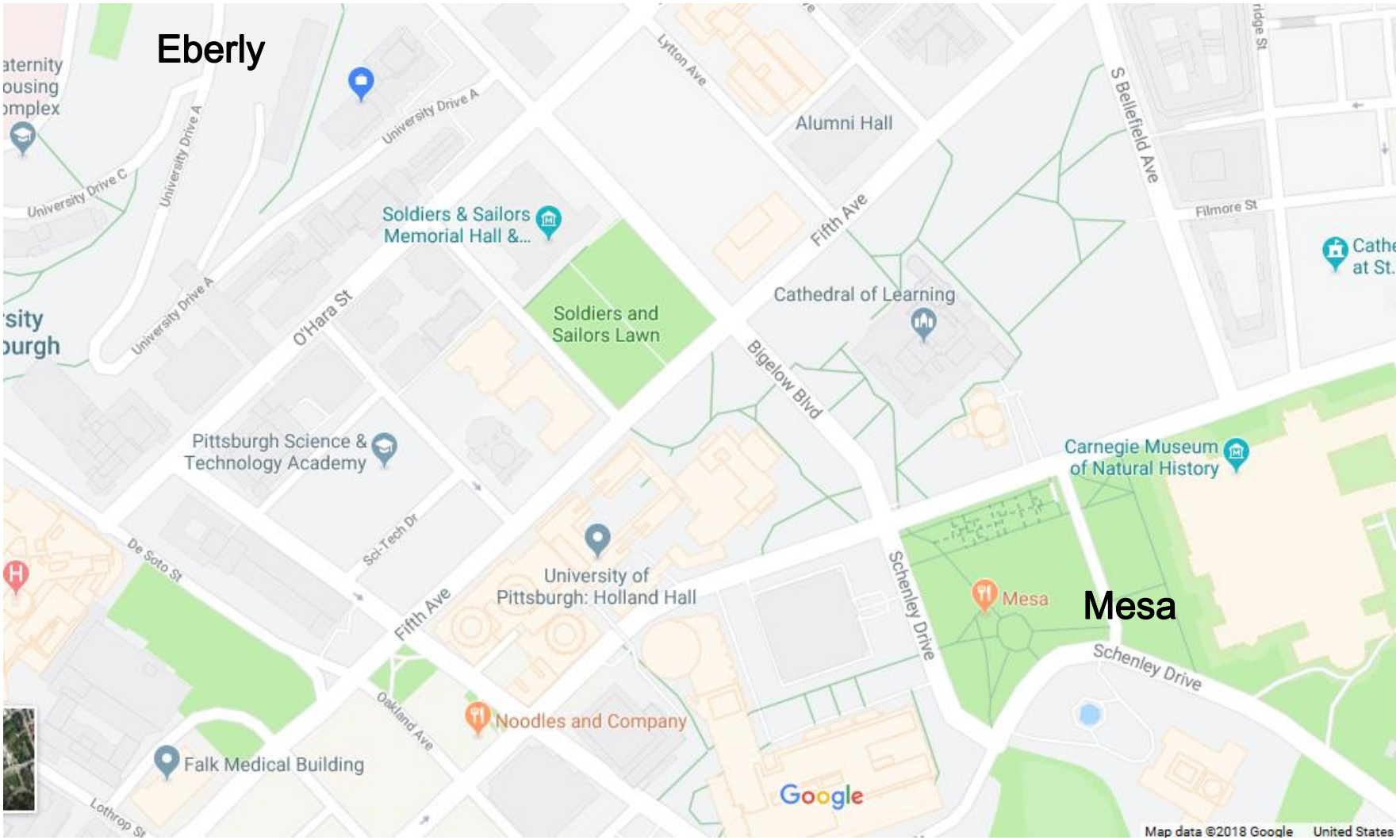
https://www.researchgate.net/profile/Robert_Murphy3/publication/292991222/figure/fig9/AS:341601621692426@1458455581570/Weighted-ensemble-algorithm-One-initial-trajectory-is-iteratively-resampled-until-its.png

Weighted Ensemble

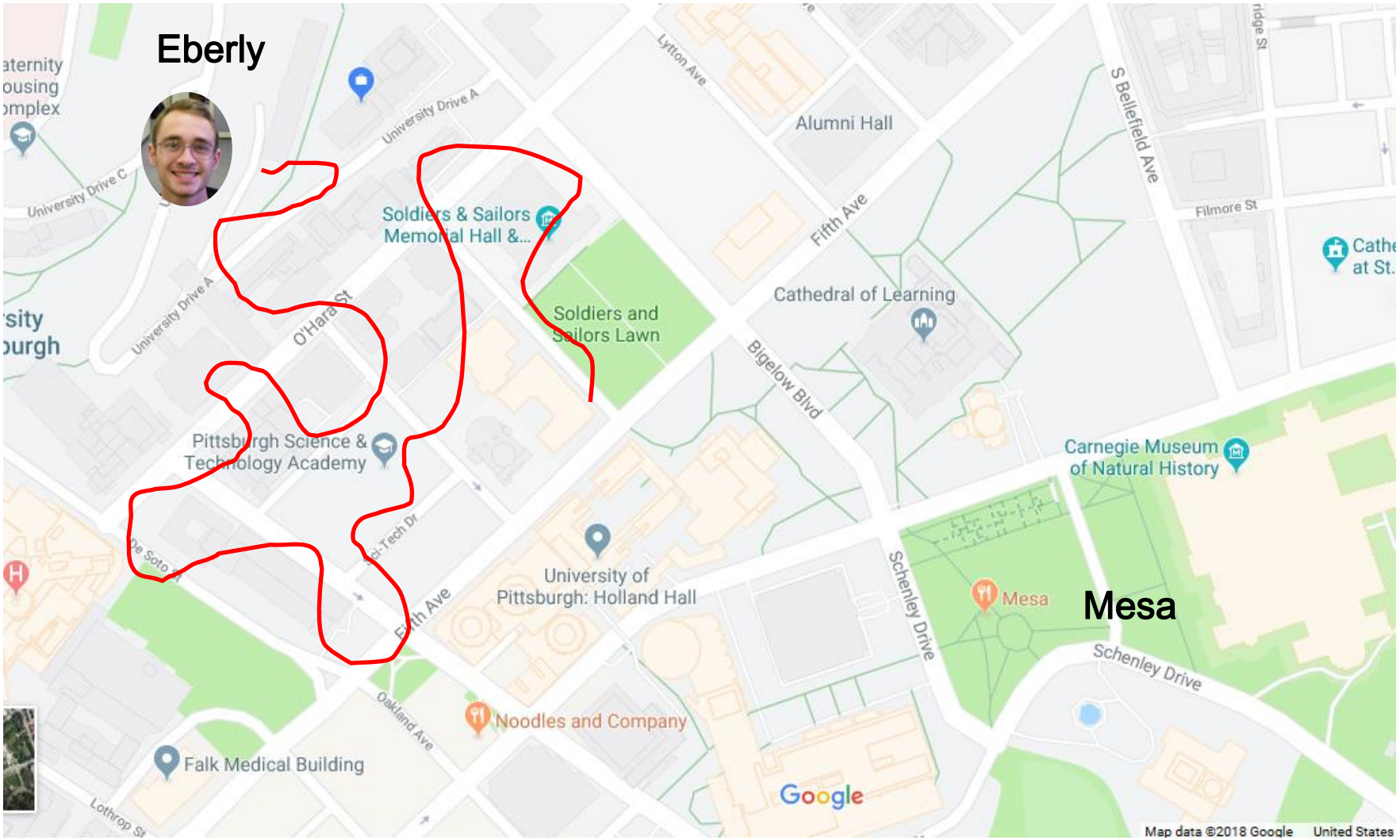


https://journals.plos.org/ploscompbiol/article/figure/image?id=10.1371/journal.pcbi.1004611.g002&size=large

An Example “Simulation”

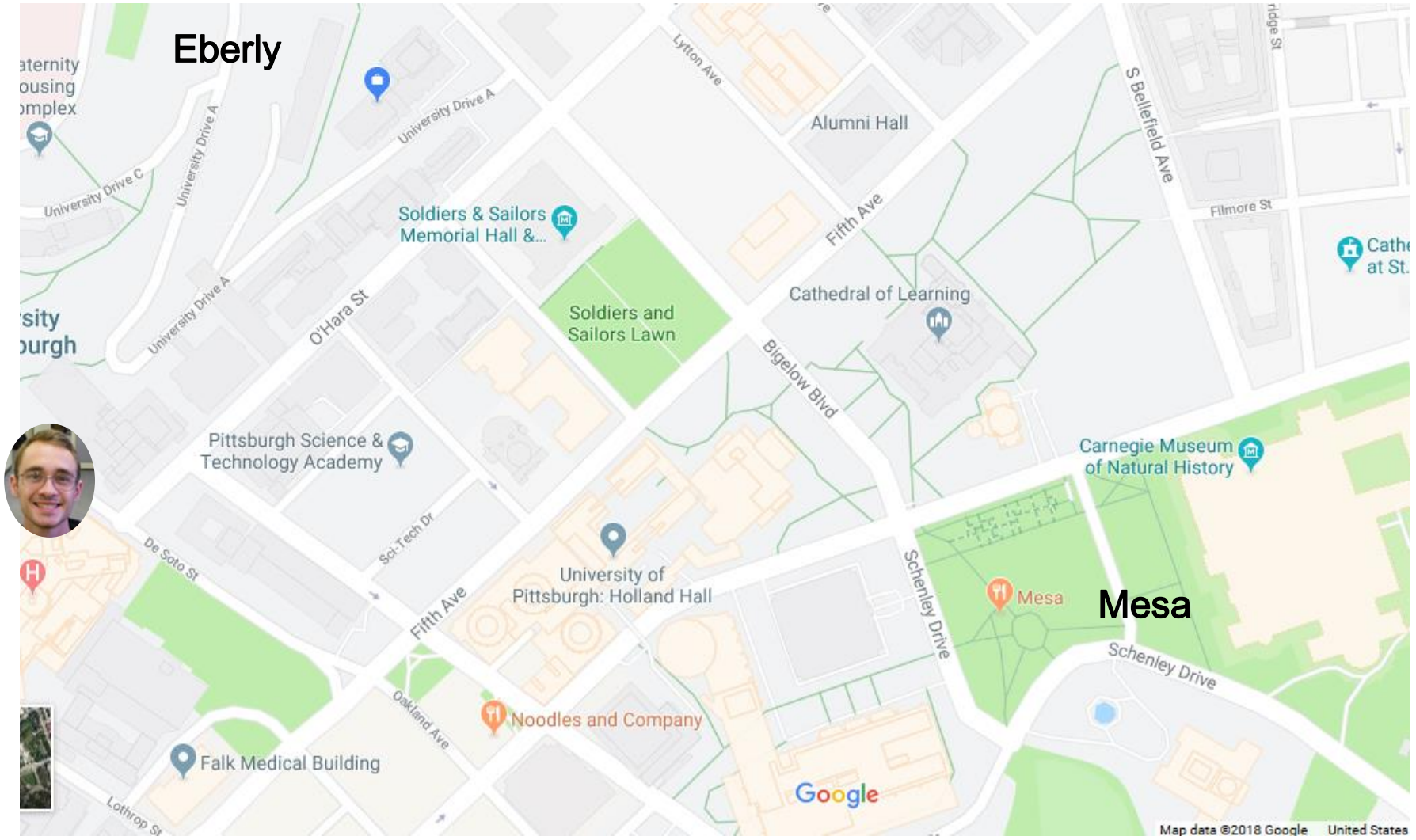


An Example “Simulation”



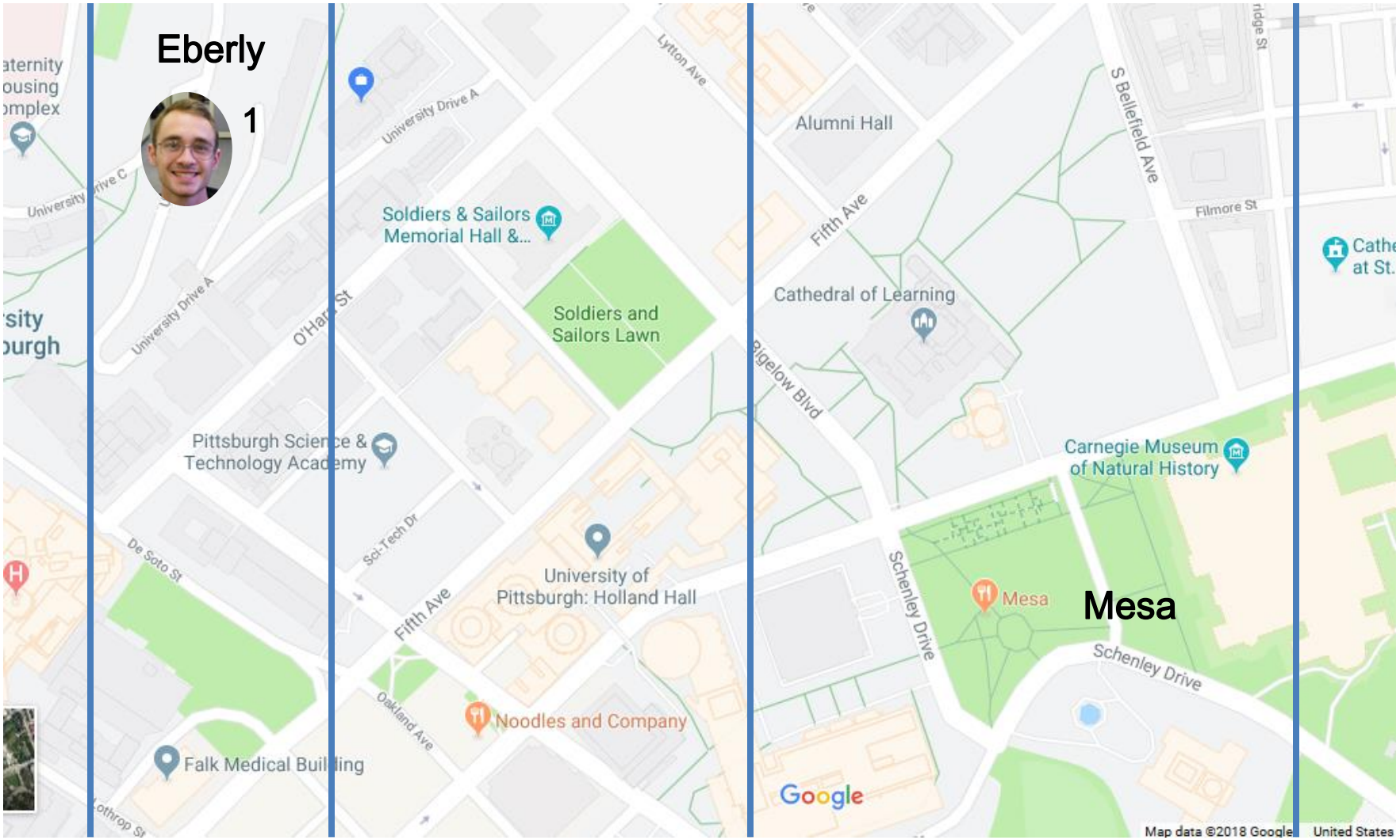
An Example “Simulation”

An Example “Simulation”

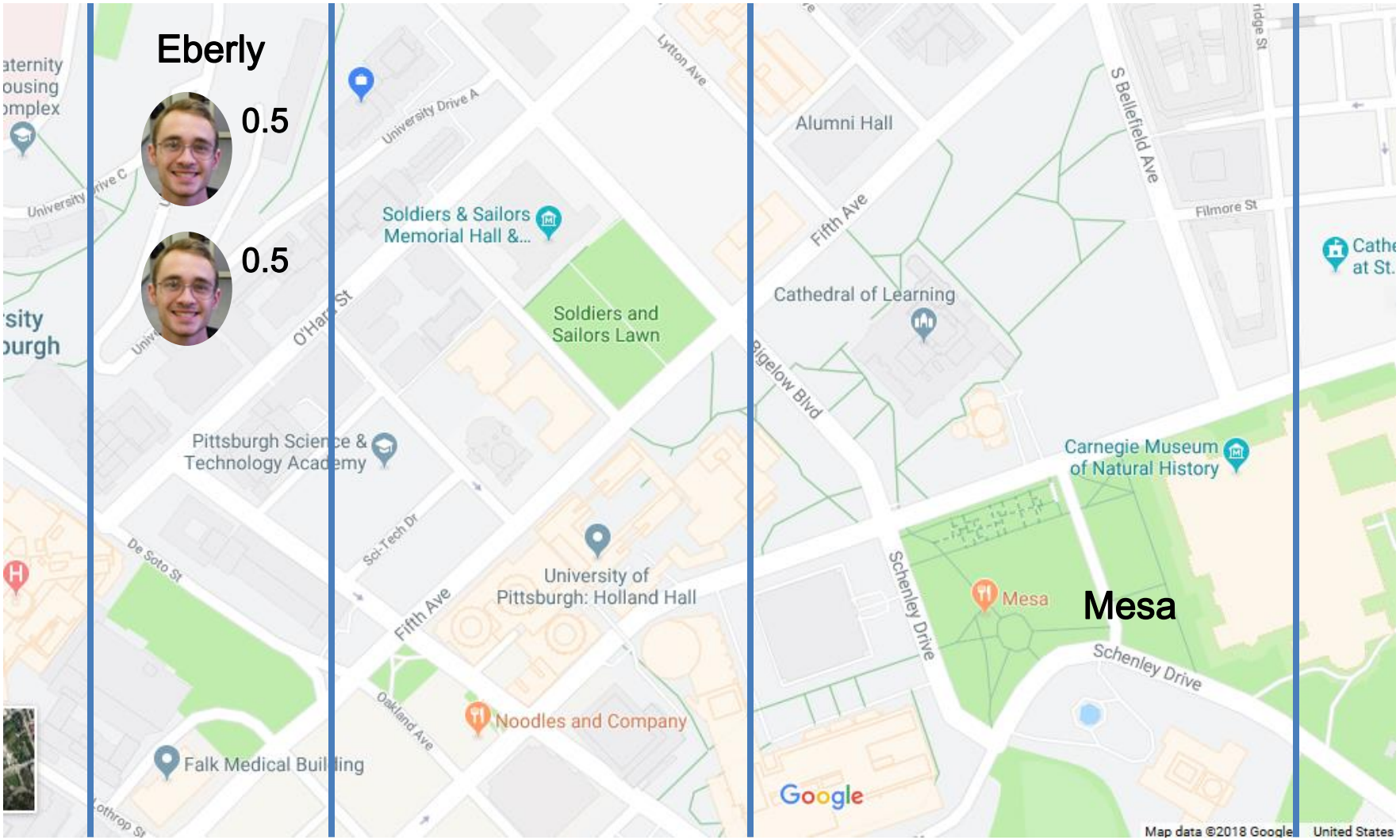


An Example “Simulation”

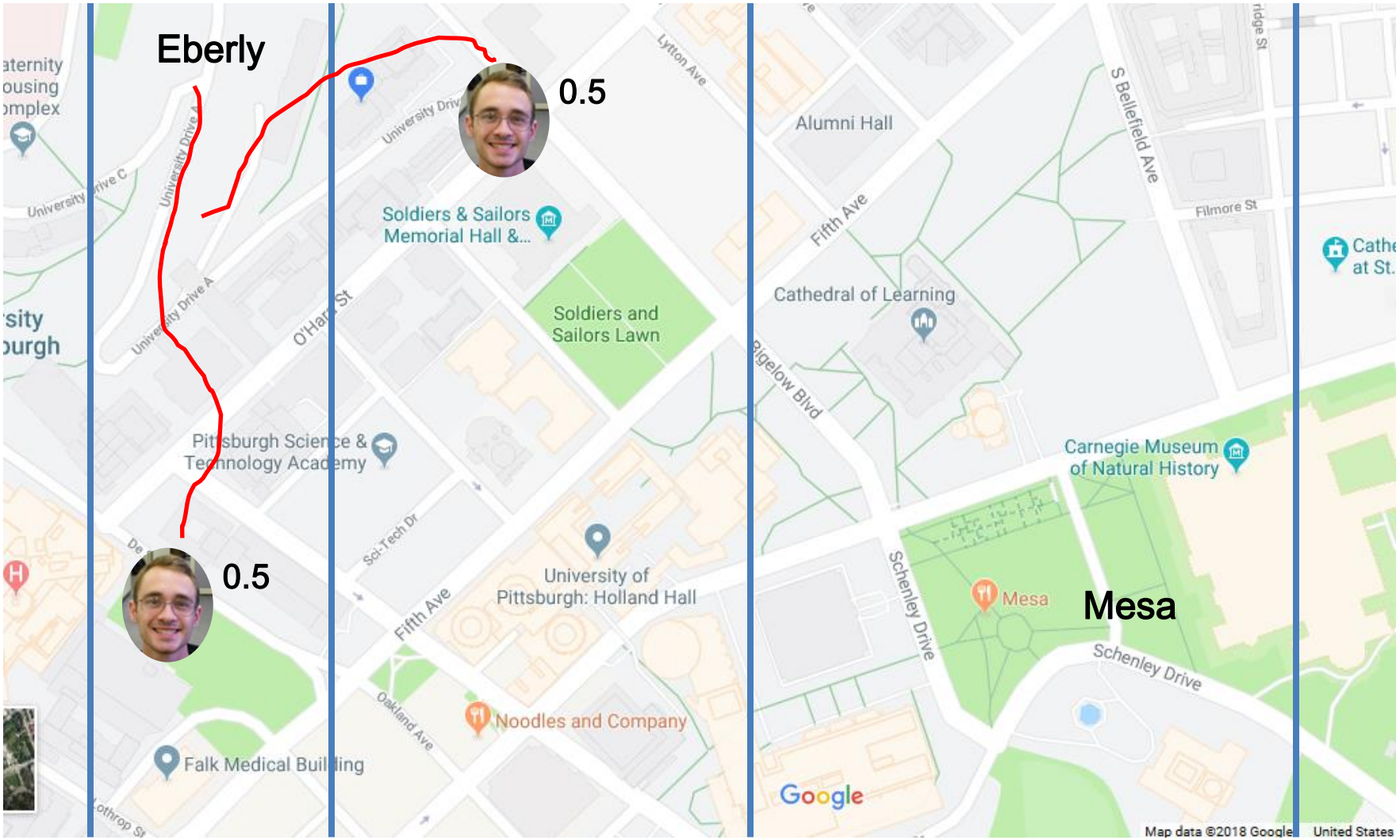
An Example “Simulation”



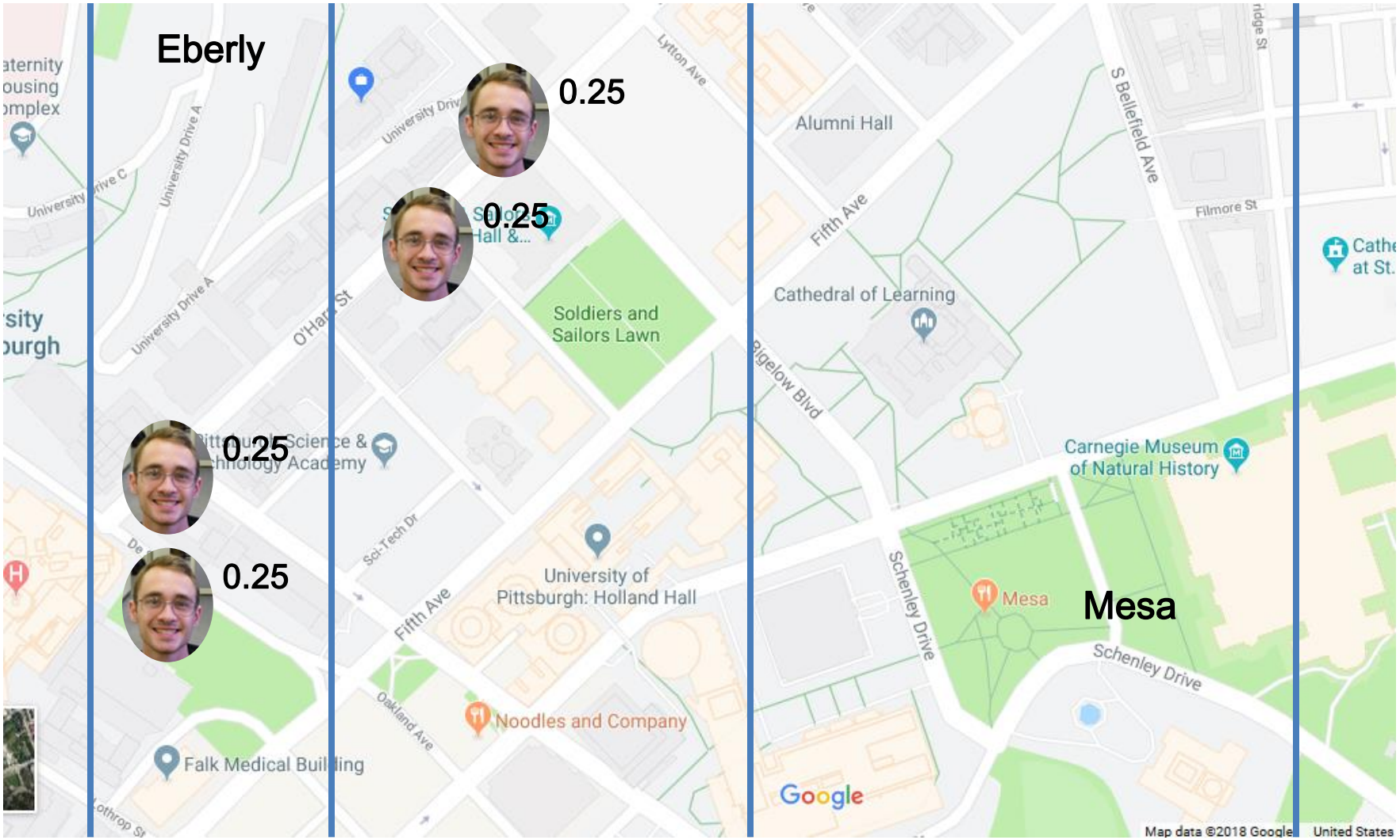
An Example “Simulation”



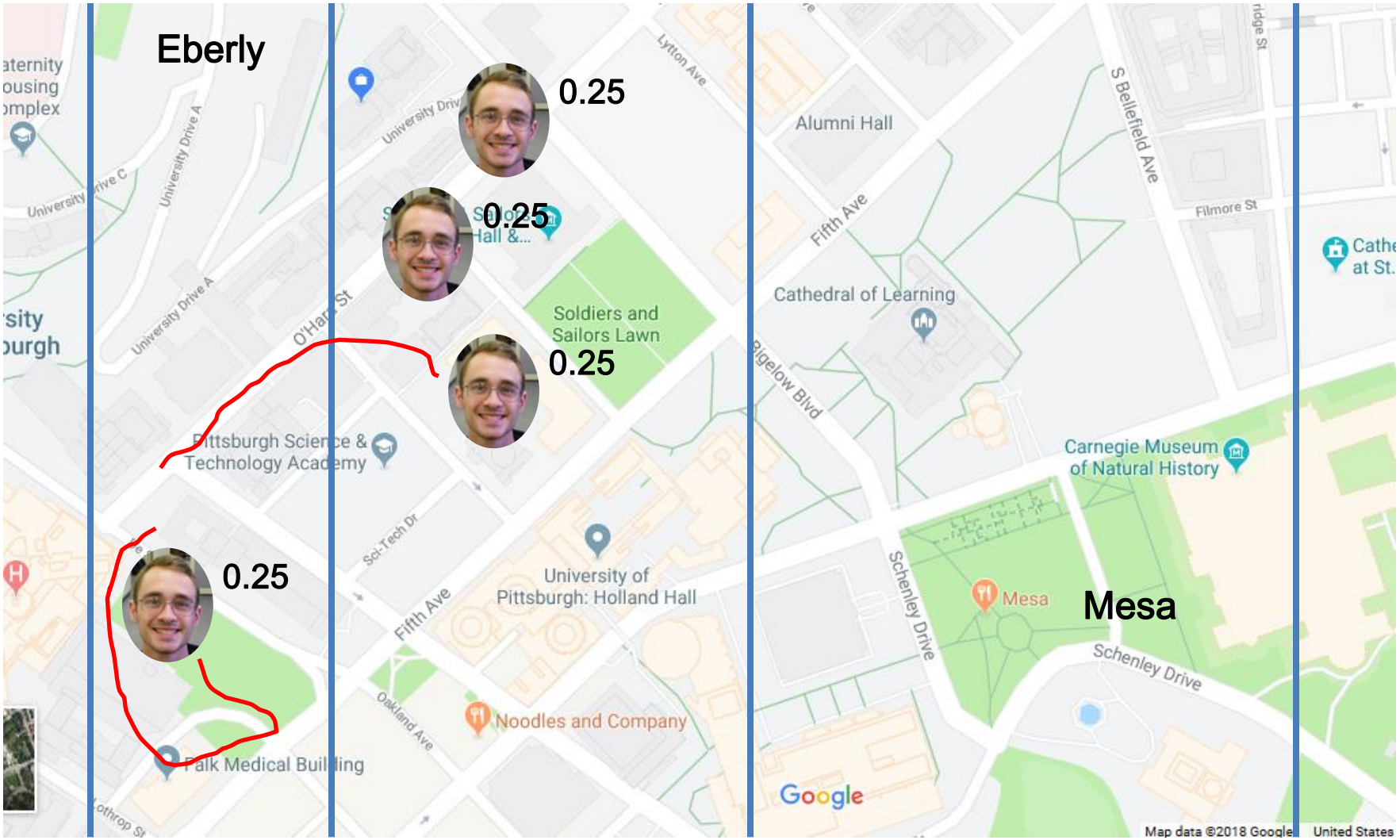
An Example “Simulation”



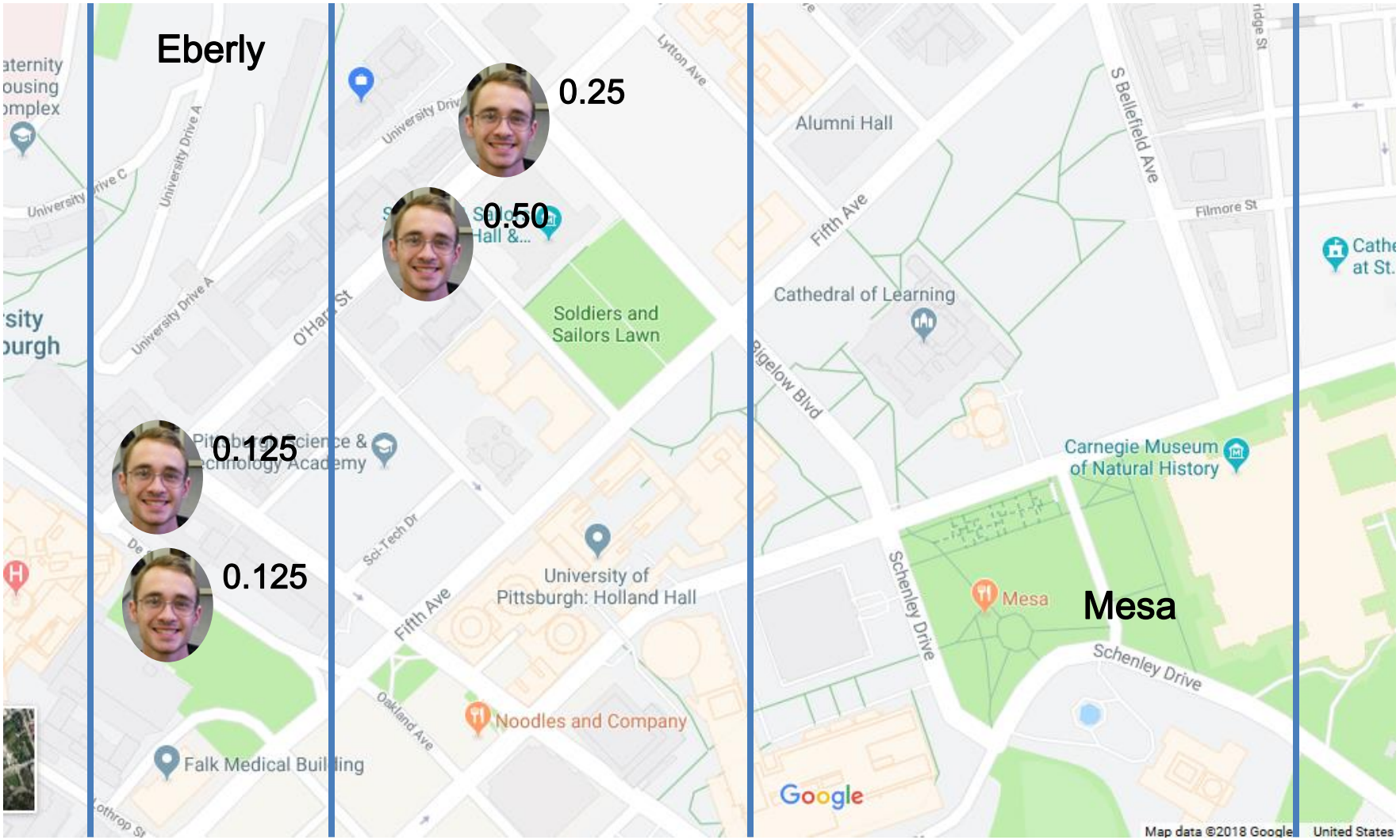
An Example “Simulation”



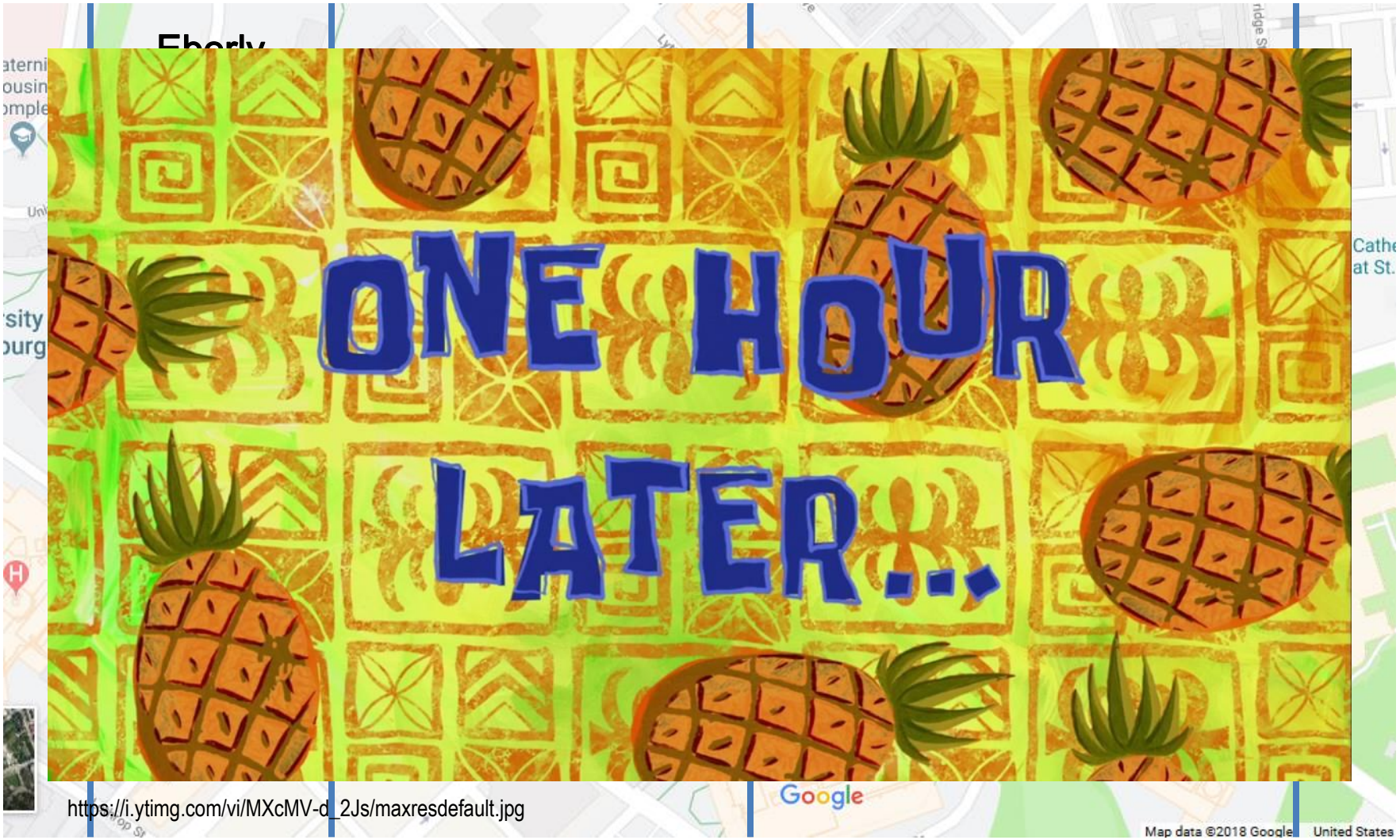
An Example “Simulation”



An Example “Simulation”

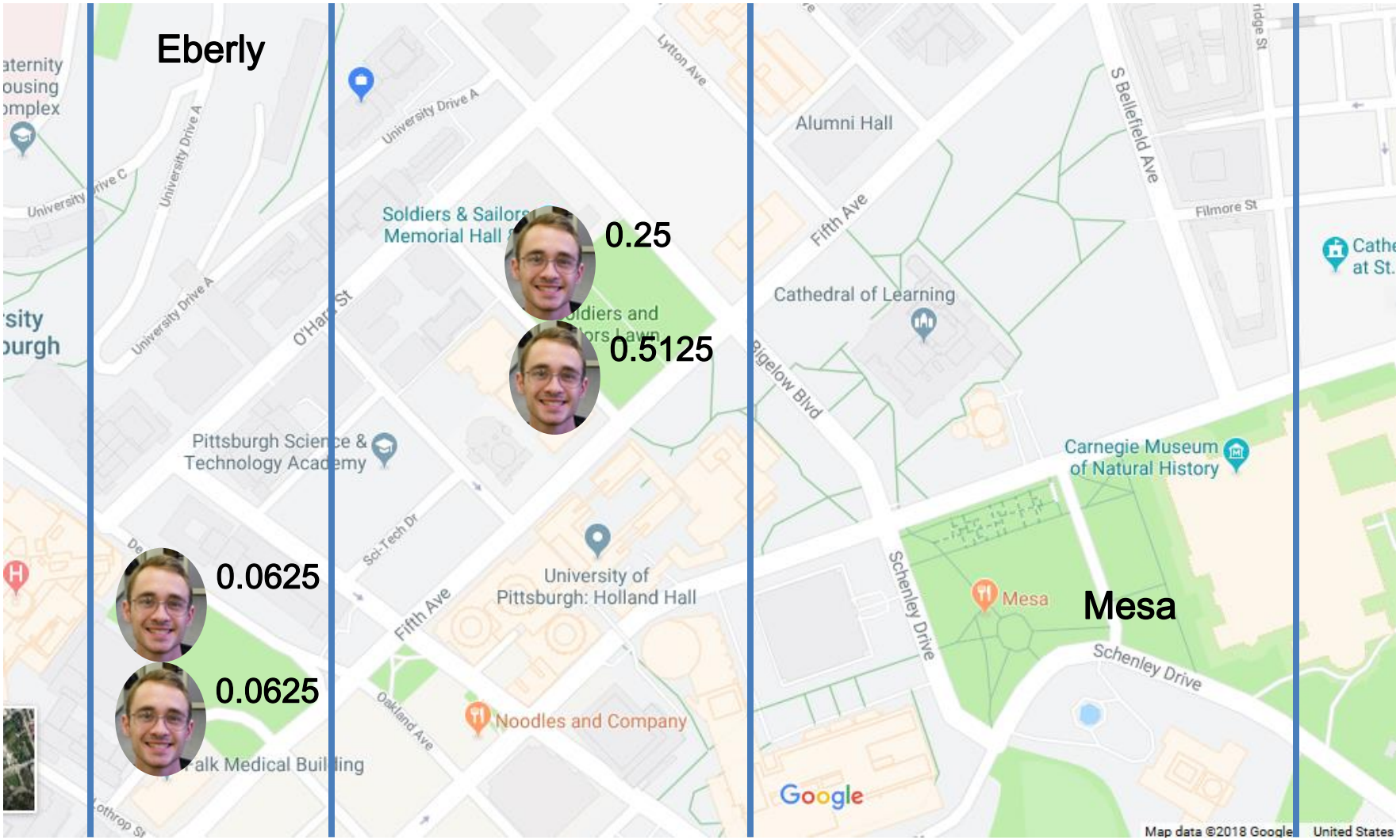


An Example “Simulation”

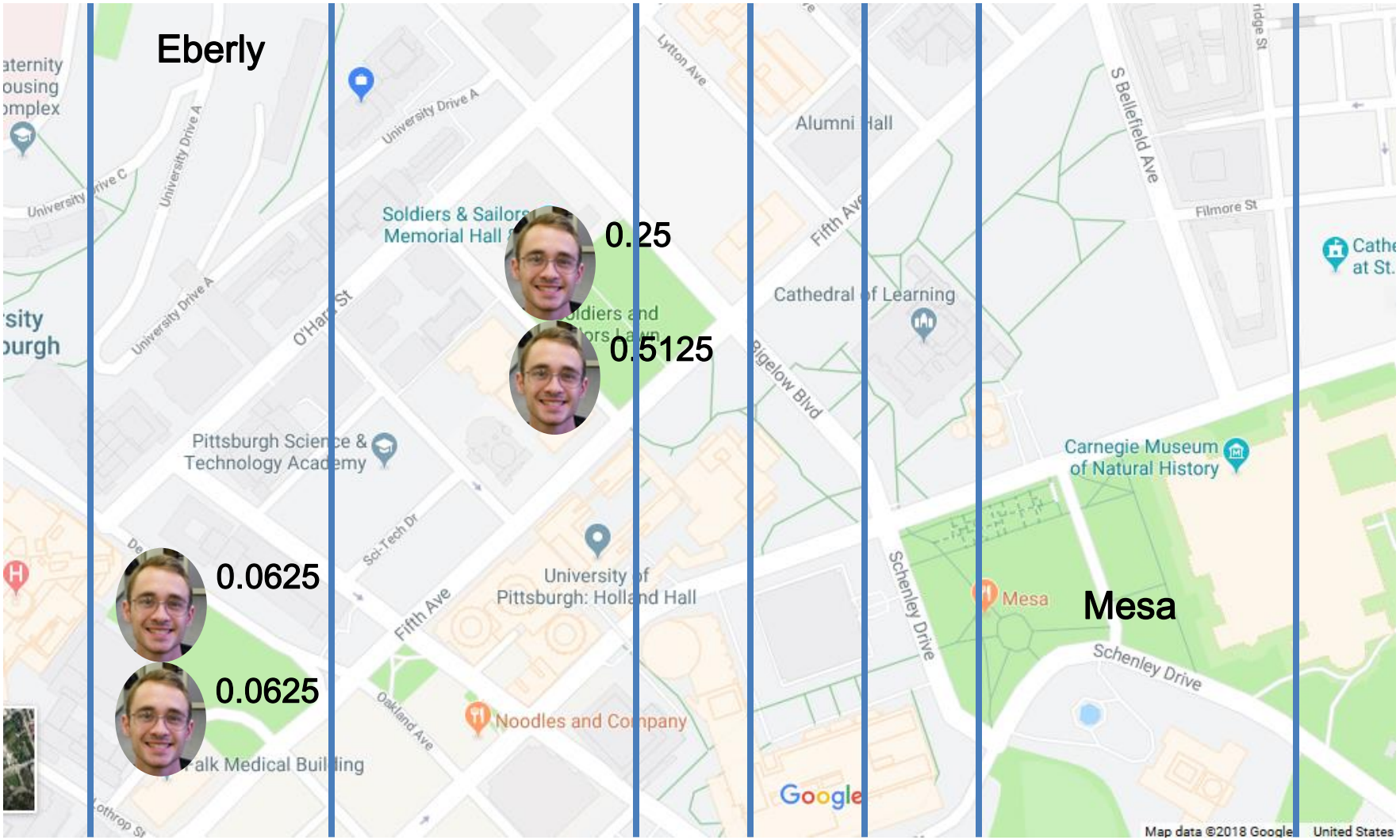


https://i.ytimg.com/vi/MXcMV-d_2Js/maxresdefault.jpg

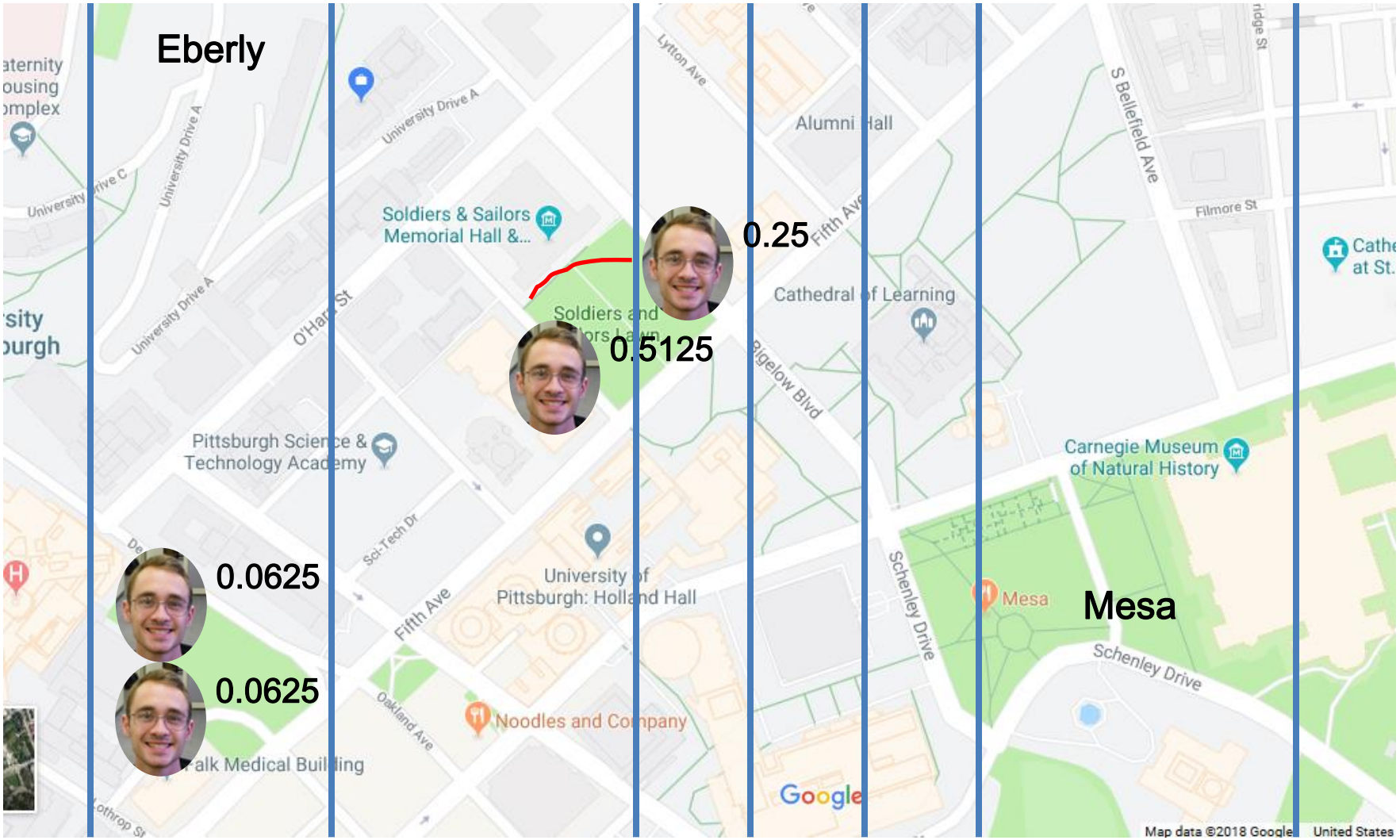
An Example “Simulation”



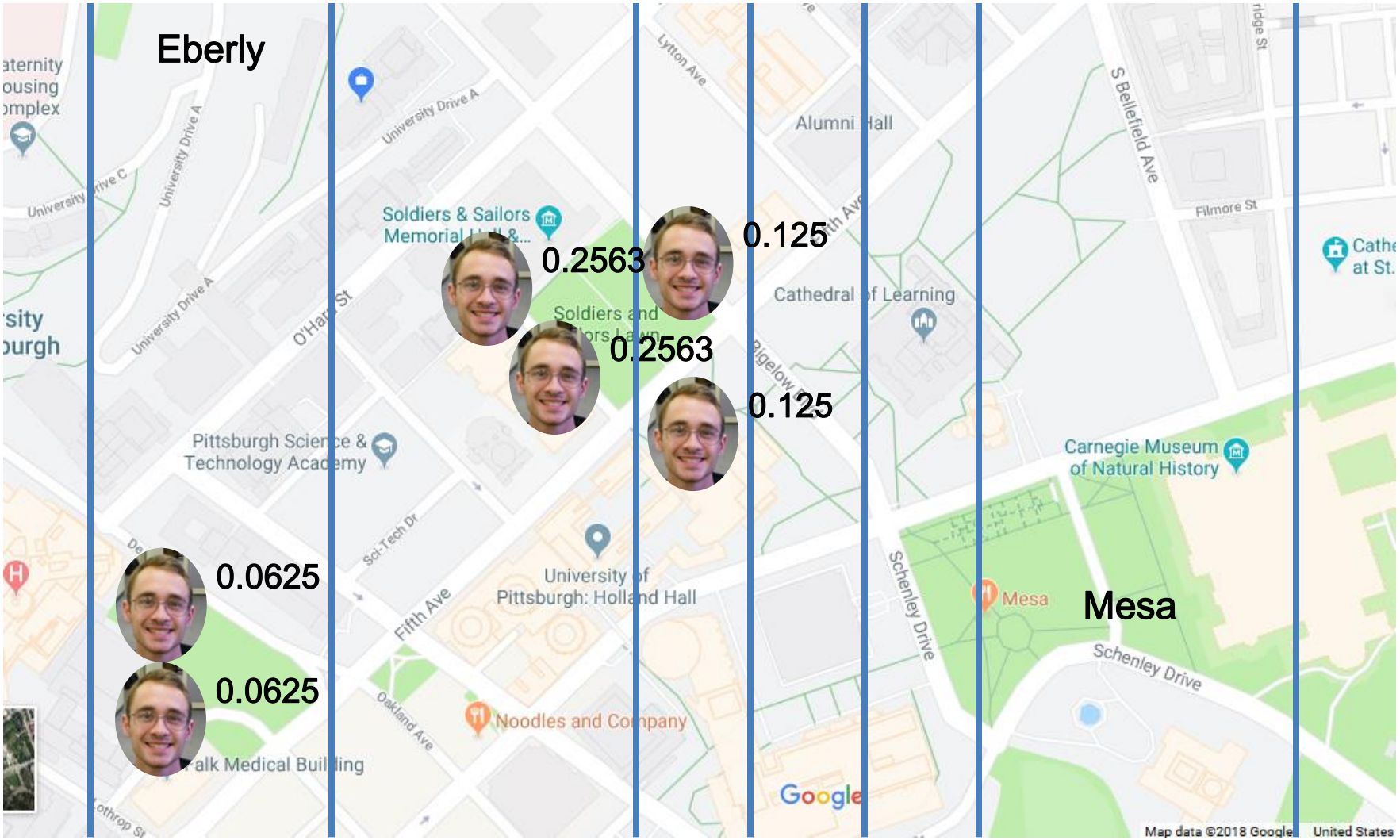
An Example “Simulation”



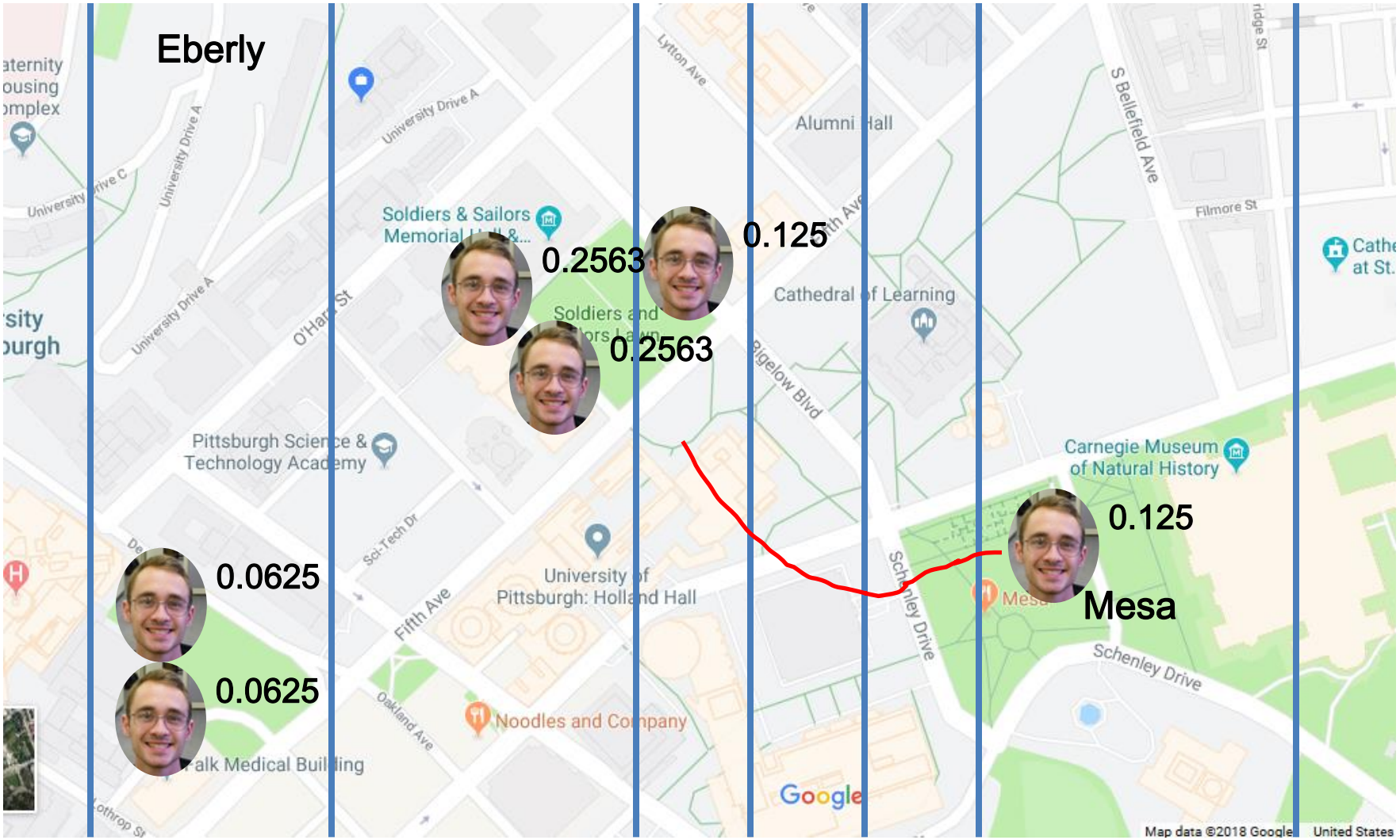
An Example “Simulation”



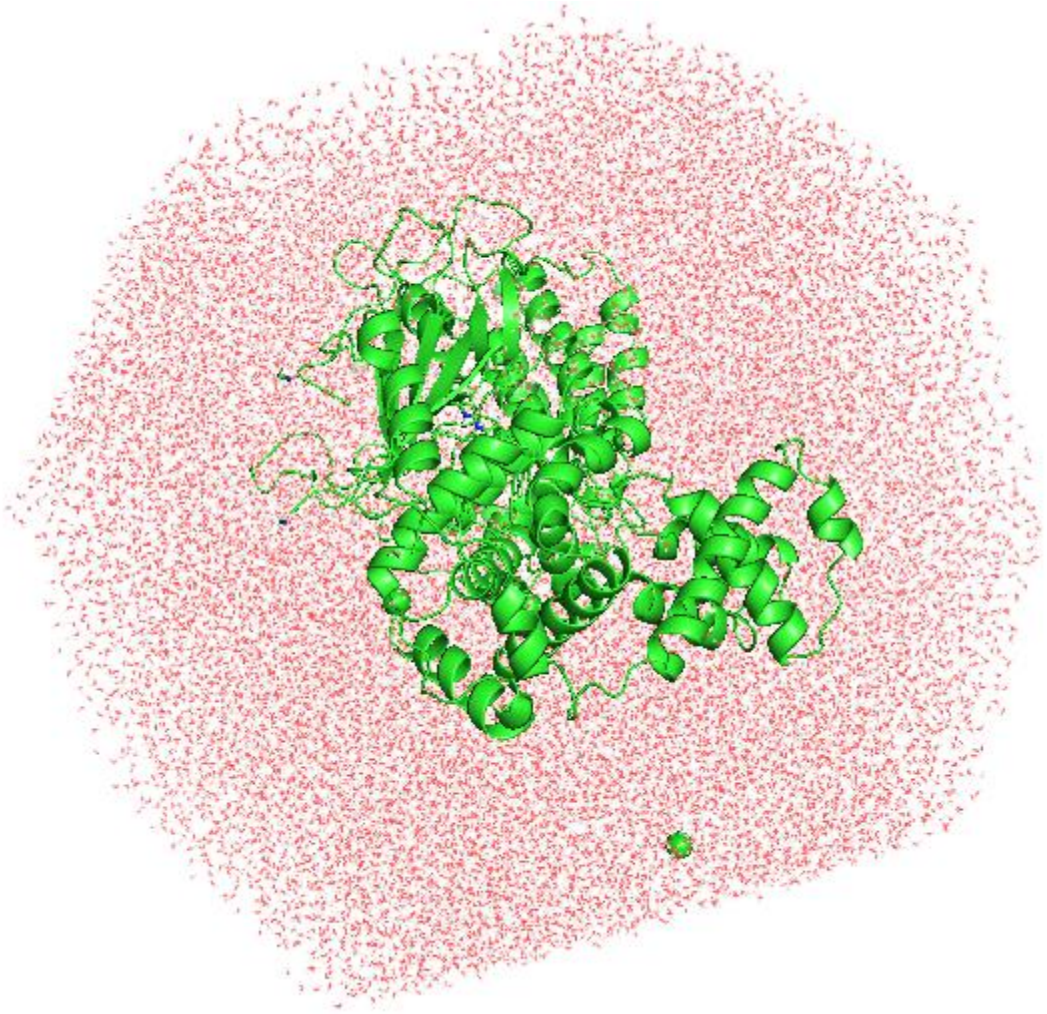
An Example “Simulation”



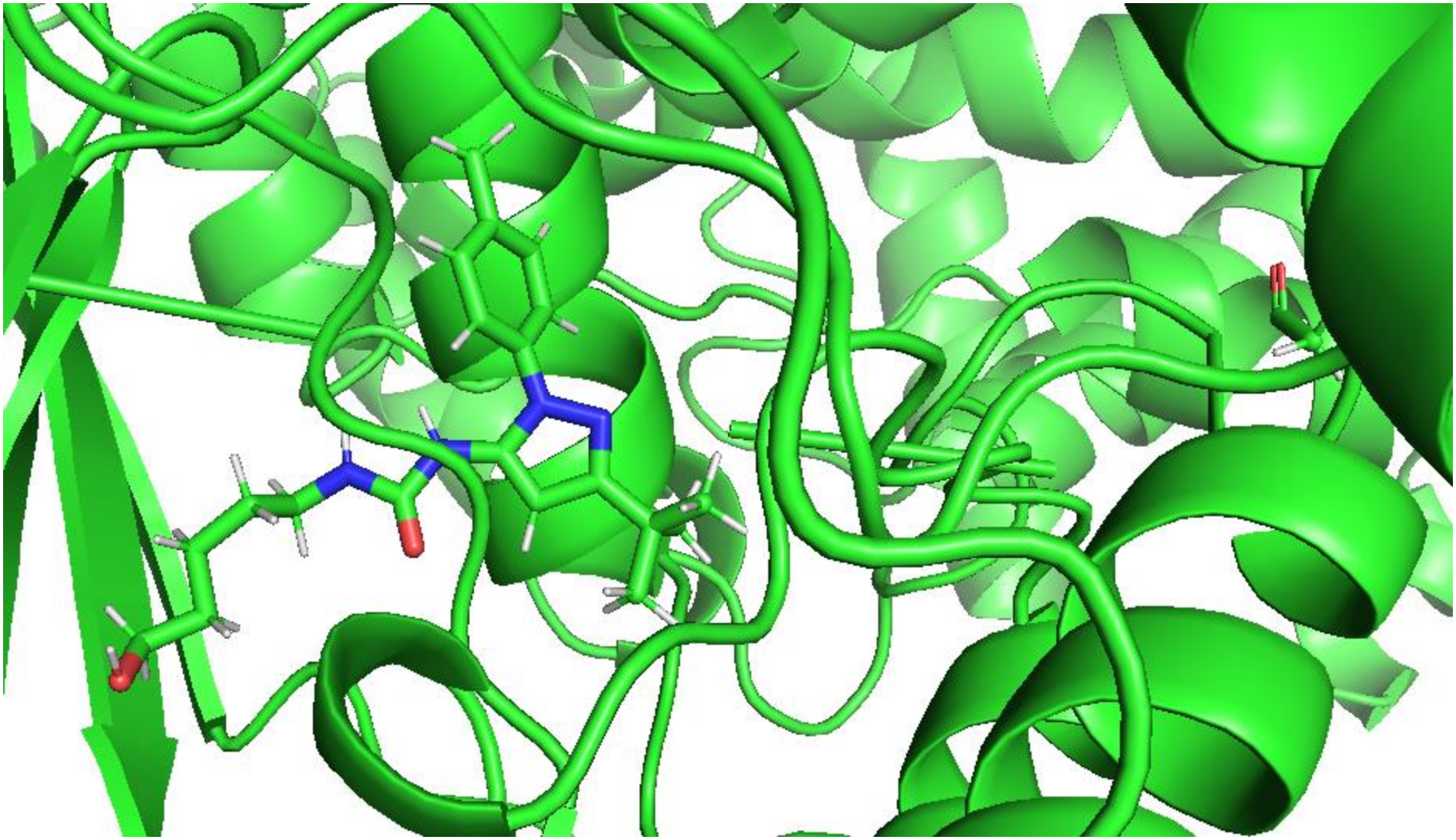
An Example “Simulation”



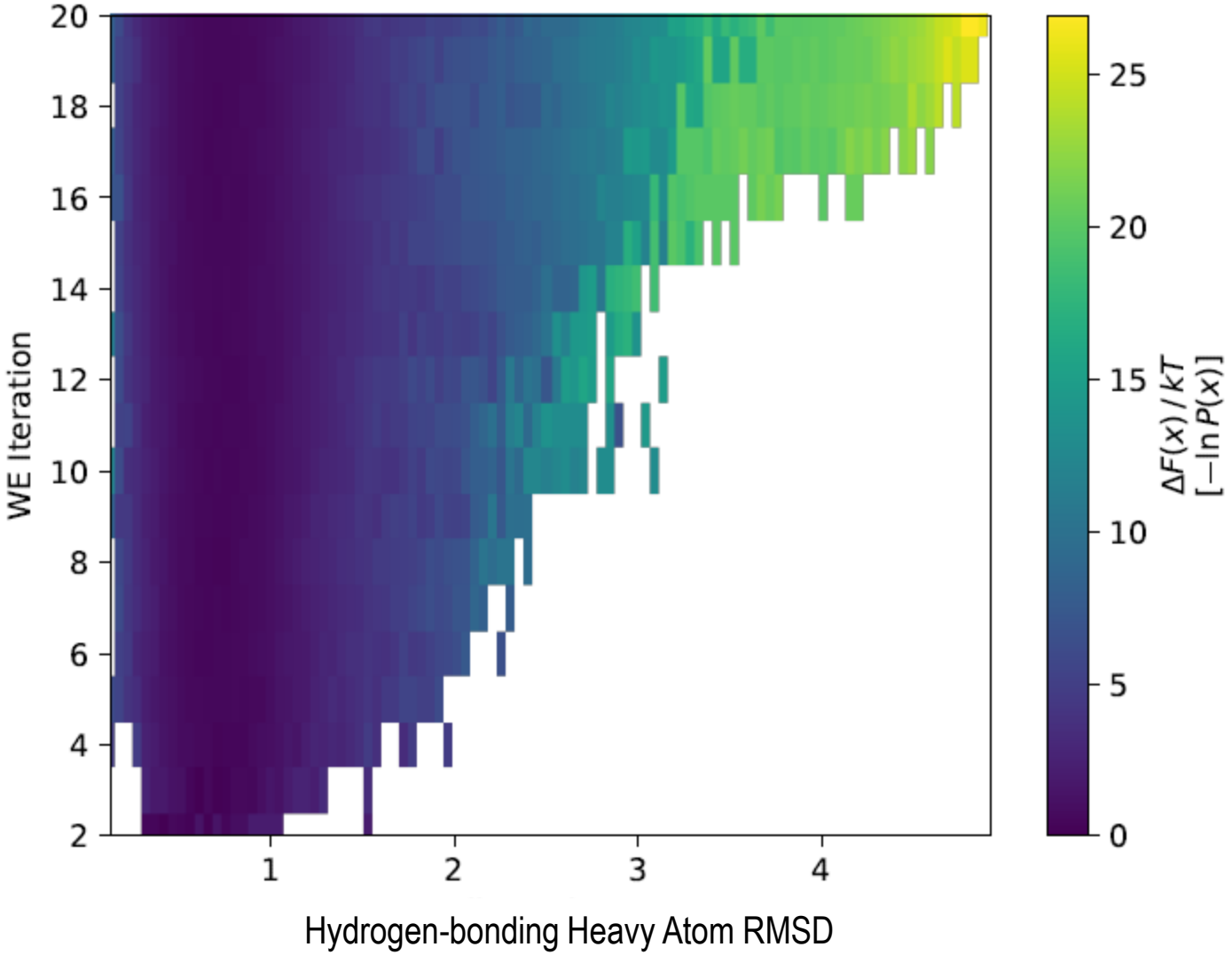
A “Real” Example



A “Real” Example



A “Real” Example



Complications

- It's really difficult to choose *good* parameters for WE (timestep of dynamics, binning scheme, order parameter, etc.)
- **One** event may not be representative of the **ensemble** of events
- Difficult to deal with error (bootstrapping)
- Target state may not be known
- Massive amounts of data are generated, it's very difficult to go back through and re-calculate something if forgotten

References

- Zwier, M. and Chong, L. Reaching biological timescales with all-atom molecular dynamics simulations. *Curr. Opinion in Pharma.* 10: 6, 2010, 745-752