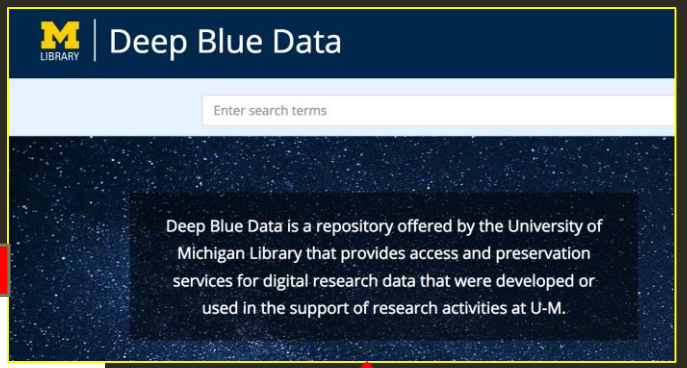
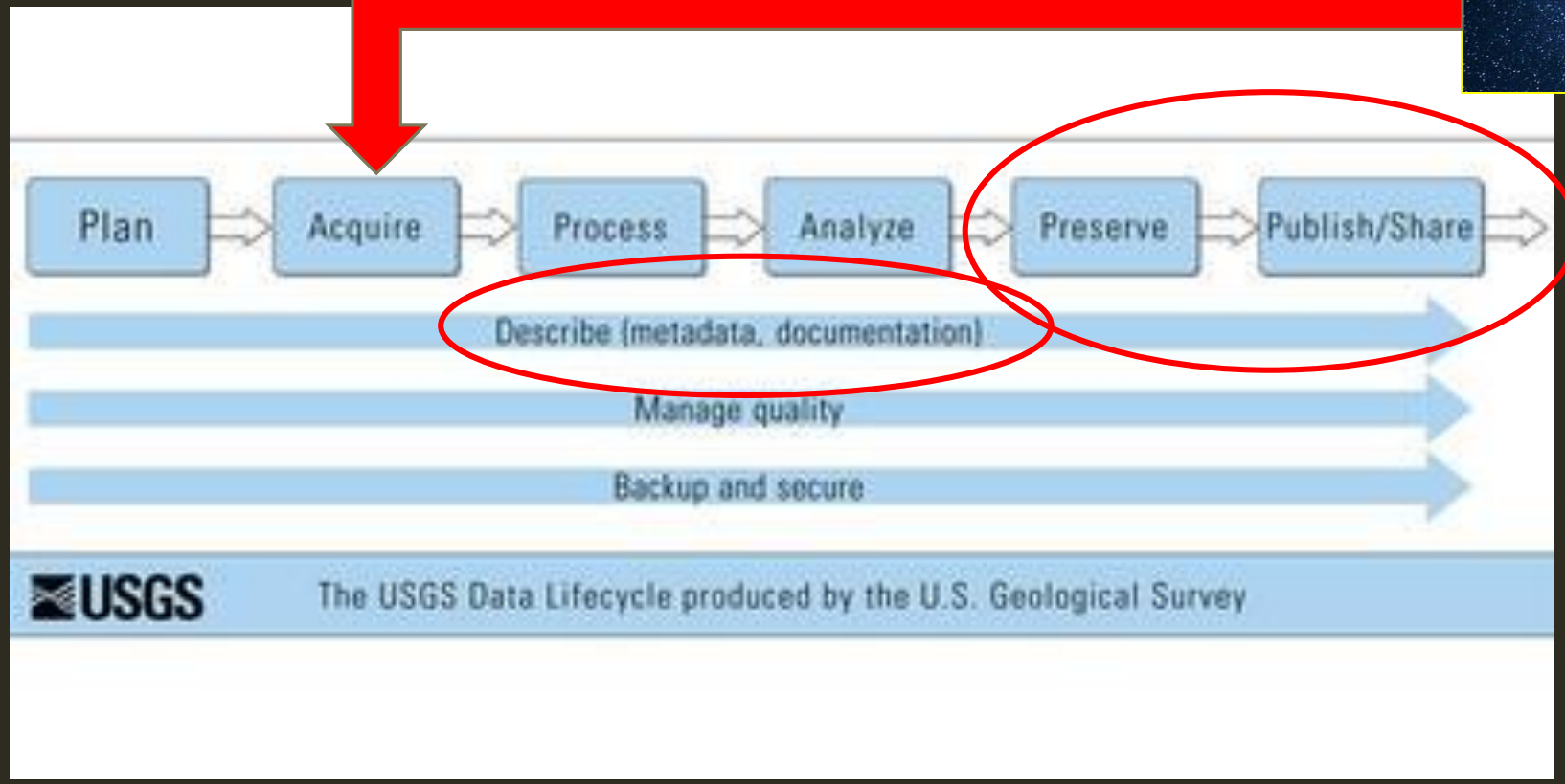
YOU HAVE DATA, WE HAVE QUESTIONS!!

Susan Borda – Data Workflows Specialist UMich Library
Scott Witmer – Digital Preservation Specialist UMich Library

https://deepblue.lib.umich.edu/data/

RESEARCH DATA LIFECYCLE





https://www.usgs.gov/products/data-and-tools/data-management/data-lifecycle

REPRODUCIBILITY GOALS

With publisher or funder requirements for sharing is there a further requirement for reproducibility or replicability*?

If so, is the dataset complete, are all the scripts, files, and documentation included?

Is the data shared with the intent of reuse?

* Reproducibility: different team, different experimental setup
Replicability: different team, same experimental setup (Plesser, 2018; see *ACM terminology*)

"Is there a reproducibility crisis?" (Baker, 2016)

LIBRARY

REUSE GOALS

Does the data have a life beyond the paper?

--- data deposited directly supports figures – do we need the raw data as well?

Is data reuse important to you/your discipline?

Do you think about other disciplines using your data?

Do institutional repositories play a role here? Or do you prefer, data portals, disciplinary repositories, lab websites managed by other researchers?

"Changes in Data Sharing and Data Reuse Practices and Perceptions among Scientists Worldwide" (Tenopir et al, 2015)

M LIBRARY

REUSE GOALS

Do you re-use data created by other researchers? What are you looking for in a dataset, what's important?

- Data source (peers or portals)?
- Documentation?

Have you had issues accessing other people's research data and/or with format obsolescence, integrity or something else?

PRESERVATION GOALS

Does all data need to be preserved?

Simulation data – as technologies develop over time is older simulation data redundant? If not, what should we be preserving, the input parameters and model information?

Observational data – in most cases cannot be recollected, what needs to be preserved?

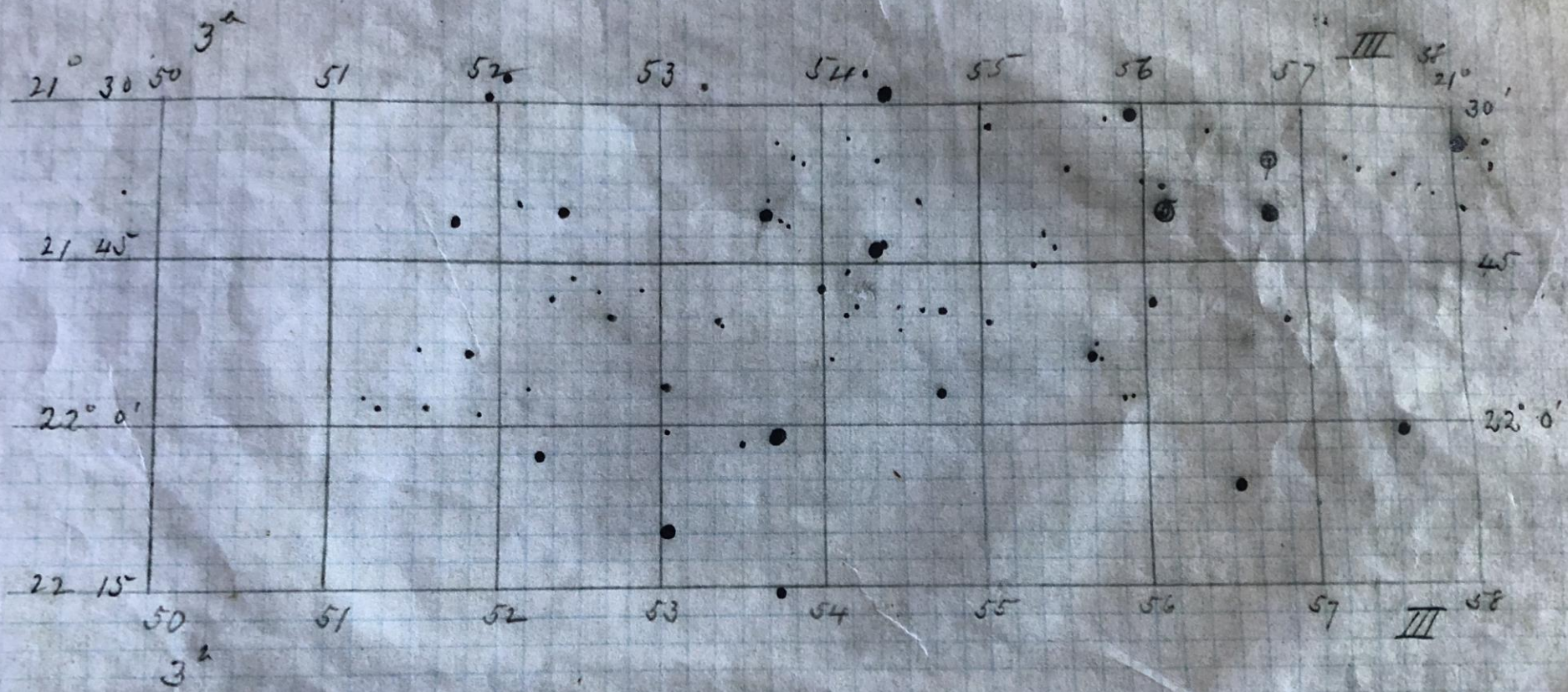
Experimental data – what needs to be preserved?

Is the dataset complete, are all the scripts, files, and documentation included?

"Research Data Preservation" (Currier et al, 2018)

M LIBRARY

Annular density for Dec. 11th 9h 3h 53m 15s 21° 26'



Places of stars as above 1877 Dec 11th at 8h 30m

Tuesday Dec 11th 1877
8h – 9h P.M. (meteor data)

CONTACT US!

Susan Borda – sborda@umich.edu

Scott Witmer – switmer@umich.edu

Deep Blue Data - https://deepblue.lib.umich.edu/data/

https://www.lib.umich.edu/research-data-services

https://www.lib.umich.edu/preservation-and-conservation/digital-preservation-unit

M LIBRARY Research Data Services

References:

Baker, M (2016) 1,500 scientists lift the lid on reproducibility. Nature 533, 452–454 (26 May 2016) doi:10.1038/533452a

Currier, B. D., Kim, B., Edwards, C., & Butler, C. R. (2017). Research Data Preservation. Presentation at the NDSA and Digital Preservation 2017, Pittsburgh, PA. Accessed through LIS Scholarship Archive. *doi:* 10.31229/osf.io/j7meu

Plesser HE (2018) Reproducibility vs. Replicability: A Brief History of a Confused Terminology. Front. Neuroinform. 11:76. doi: 10.3389/fninf.2017.00076

Tenopir C, Dalton ED, Allard S, Frame M, Pjesivac I, et al. (2015) Changes in Data Sharing and Data Reuse Practices and Perceptions among Scientists Worldwide. PLOS ONE 10(8): e0134826. https://doi.org/10.1371/journal.pone.0134826

M LIBRARY