

Study Questions 1

Where the Rubber meets the Sky: Bridging the Gap between Databases and Science

- 1) The authors propose that there is a Metcalfe's Law for datasets. Explain with a figure what they are talking about. What are some of the assumptions that must be made about the datasets for this Law to hold?
- 2) What are some factors that might mitigate the Data Inflation Paradox of Section 6?
- 3) The term *dataset* is used frequently in the physical sciences (and in this paper), but not so much in commercial contexts. Define dataset.

Section 9 in *The SDSS SkyServer – Public Access to the Sloan Digital Sky Survey Data*

- 4) In Section 9.1.3, the authors describe how they replaced *tag tables* with indices. Are there any circumstances in which a tag table might still be preferred, or is an index always a superior solution?
- 5) Section 9.1.4 in the SkyServer paper describes the Hierarchical Triangular Mesh index. Give two queries (in English) that would benefit from this index.