

Abstract:

Wolfgang Voges

The Data behind a Publication: An Astronomer's View

The amount of data being produced yearly in astronomy will soon reach the Peta-byte limit. Observational data covers the whole frequency or wavelength regime (radio, infrared, optical, ultraviolet, X-ray, and gamma-ray) and is collected from both ground-based telescopes and instruments on satellites. At the same time, simulations on supercomputers produce equally daunting quantities of theoretical data. Innovative methods and tools are needed to ingest, digest and concentrate this data before publishing the results and triggering a new cycle of knowledge discovery.

This talk addresses a new conception of the relationship between data and publications. It is increasingly important to make not only the final, condensed results available, but also the raw data, detailed methods, and complete output. This supplementary information, which may consist of lists, catalogues, image collections, or computer programs, has to be in electronic form to be used effectively.

It is now common policy in the field of astronomy that data are made public soon after they have been recorded, processed, or calculated. This accords with the spirit of publicly funded research, enhances the quality of the science, discourages sloppy or unethical conduct, and encourages rapid progress.