

MASSIVE RADIO ASTRONOMY SURVEYS AT ARECIBO USING  
ALFA: DATA MINING AND MANAGEMENT

James M. Cordes

Astronomy Department, Cornell University, Ithaca, NY 14853 USA

The Arecibo L-band Feed Array (ALFA) provides a seven-fold increase in survey throughput for extragalactic, galactic and pulsar surveys using the Arecibo telescope in Puerto Rico. These surveys are being conducted through large scientific consortia that are international in membership and which include substantial student involvement. Surveys have begun recently and will extend at least five years into the future. Both the astrophysical return and the day-to-day management of the large data volume demand new ways for conducting the surveys that involve students' participation in many different ways. In addition to traditional astronomical activities, these also include data mining and data archival and, thus, closer interaction with colleagues in computer science. In particular, the surveys require construction of data bases for all stages, from the data acquisition to data archival. For the pulsar surveys, which will yield the largest data volume of any of the ALFA surveys (e.g. one petabyte in about five years), we plan to archive all the raw data as well as the data products from the processing. Processing of the raw data and exploration of data products will be done through a web-based system sited at the Cornell Theory Center (CTC). In addition, a database is sited at the Arecibo Observatory that addresses the immediate needs for observations of a pre-defined grid of positions, data quality, and data transport. Data are moved to remote processing sites using inexpensive disks and then sent to the CTC for archival as well as processing. At Arecibo, the database system uses MySQL while at CTC, Microsoft Server SQL is used. Tools are being developed that allow two-way interchange between the two database systems using web services. ALFA surveys will contribute catalogs and refined data products to the National Virtual Observatory (NVO). My talk will discuss the surveys and consortia, data management and archival, some early science results, and student participation in these activities, including NVO participation.

Abstract Submission Form

2006 National Radio Science Meeting

Abstract: cordes8216

Date Received: September 17, 2005

1. (a)  
James Cordes  
Astronomy Department  
520 Space Sciences Building  
Ithaca, NY  
14853 USA  
jmc33@cornell.edu  
(b) 607 255 0608  
(c) 607 255 8803
2. J - Radio Astronomy
3. (a)
4. I - Invited Paper, Program chair:  
Robert Mutel
5.   
ROWS=3;COLS=50;This  
submission is for the session  
entitled "Radio Astronomy  
Teaching Laboratories and  
Instruments";