Of Programs and Papers: Making the *Chandra* Connection

Sherry Winkelman and Arnold Rots

Proposal/Grant/Paper Linking Process

The *Chandra* bibliography is a complicated database storing metadata about publications closely connected in some way to the *Chandra* mission. It includes all publications related to the *Chandra* mission, from *Chandra* science papers to publications about the instruments, software, and operations of the CXC. In the last year we have expanded the bibliography to include metadata describing, among other things: the general content of papers; justifications for inclusion in the bibliography; where/how *Chandra* data links are made to papers; and *Chandra* related details about acknowledgments and grants included in publications. These metadata are incomplete and will take time to be fully populated, but when finished, the *Chandra* bibliography will be a much richer resource for scientific, bibliographic, and historical investigations.

Our initial focus is to populate metadata related to *Chandra* theory and archive grant acknowledgments, since that is the only mechanism that allows us to track such publications related to those programs except perhaps for dataset IDs in the archival case. We have completed an ADS full-text search for all grants from *Chandra* cycles 1-14 against all searchable papers in ADS as of June 27, 2013. For each archive and theory grant reference, we have determined whether the paper is related to the abstract of the proposal linked to the grant, and to which institution the grant was attributed.

Maintaining the *Chandra* Bibliography

The CDA relies heavily on ADS for locating potential *Chandra*-related publications and we have automated as much of the process as possible to reduce the number of papers we examine that are unrelated to *Chandra*. We query ADS once a week for all papers which have "chandra," "axaf" or "x-ray" in the title, abstract or keywords, adding new bibcodes to our working database. We average 100 new bibcodes per week. We then download each of those papers, convert them to text, and flag for manual inspection those papers containing the word 'chandra'. At this point we have 40 new bibcodes a week to classify, 45% of which are *Chandra* science papers as defined at http://cxc.harvard.edu/cda/bibstats/bibstats.html.

In the next few months we will collect new bibcodes weekly by using the full-text search interface from ADS to scan *Chandra* grants acknowledged in publications. We expect this to yield about 5 bibcodes more per week to classify. We would like to eventually expand our initial search against ADS to include all astronomy papers. Until we have some semantic search capabilities able to focus on *Chandra* science papers, this may increase our classification load significantly while only adding a few percent to the *Chandra* science paper category, so is therefore not currently cost-effective.

Recommendations to Authors

There are several things that you as PIs and authors can do to make grant acknowledgments meaningful and a useful resource to the *Chandra* bibliography: (1) acknowledge your grant as coming from *Chandra*; (2) check that the grant number is correct; and (3) only acknowledge your *Chandra* grant if the work presented in the paper is related to the proposal that the grant comes from. We also encourage authors to use dataset identifiers in their AAS manuscripts as a means of creating direct links to the data used in their publications. You can find details on using and/or creating dataset identifiers at the CDA website, http://cxc.harvard.edu/cda/datasetid.html.