Two years of ALMA bibliography -- lessons learned

Silvia Meakins[1], Uta Grothkopf[1], Marsha Bishop[2], Felix Stoehr[1] & Ken Tatematsu[3]

[1] ESO, Garching, Germany [2] NRAO, Charlottesville, VA, USA [3] NAOJ, Tokyo, Japan

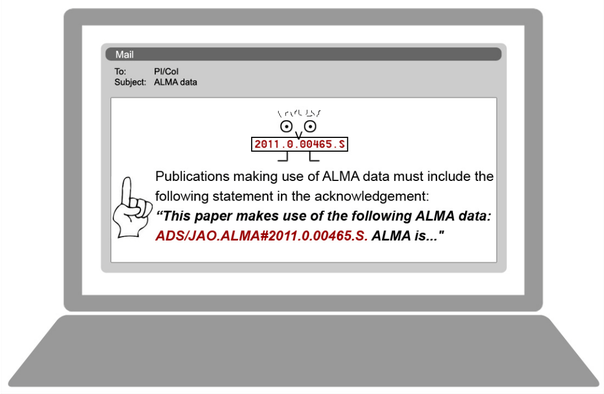
Telescope bibliographies are integral parts of observing facilities. They are used to associate the published literature with archived observational data, to measure an observatory's scientific output through publication and citation statistics, and to define guidelines for future observing strategies.

The ESO and NRAO librarians as well as NAOJ jointly maintain the ALMA (Atacama Large Millimeter/submillimeter Array) bibliography, a database of refereed papers that use ALMA data.

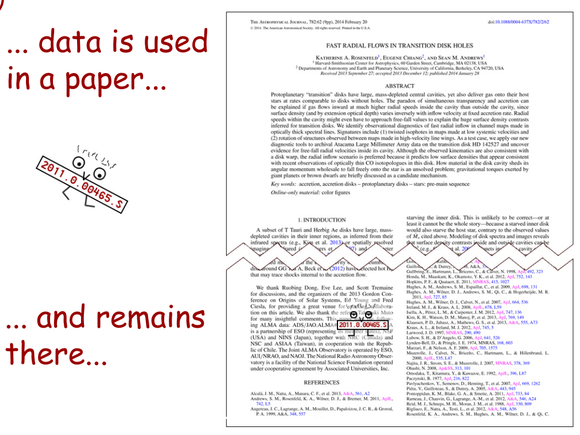
This poster illustrates how relevant articles are identified, which procedures are used to tag entries in the database and link them to the correct observations, and how results are communicated to ALMA stakeholders and the wider community. Efforts made to streamline the process will be explained and evaluated, and a first analysis of ALMA papers published after two years of observations will be given.

WORKFLOW – from data to statistics

1 PI receives data…



2 … data is used in a paper…



… and remains there…

3 The Fulltext Search tool FUSE (©ESO Library) is used to identify papers that use ALMA data. FUSE semi-automatically screens articles for ALMA-related keywords and highlights results in context. Each paper is then visually inspected by the librarians to decide whether or not it shall be included in the bibliography.

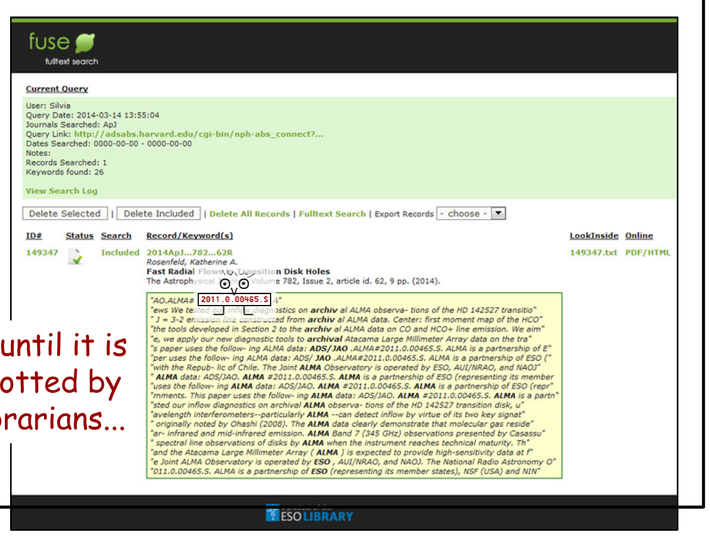
Policy for paper inclusion:

Papers must use partly or exclusively ALMA data.

Papers that
- merely quote results from the literature
- describe instrumentation or software
- simply mention ongoing projects
- suggest future observations
- develop models or run simulations

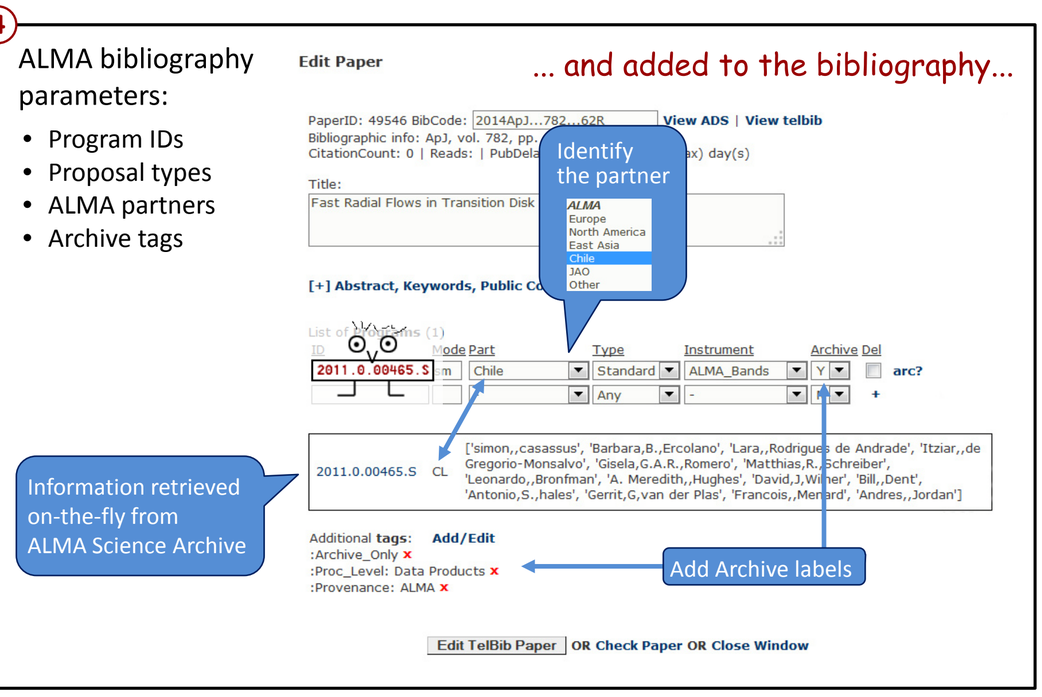
are not included.

… until it is spotted by librarians…

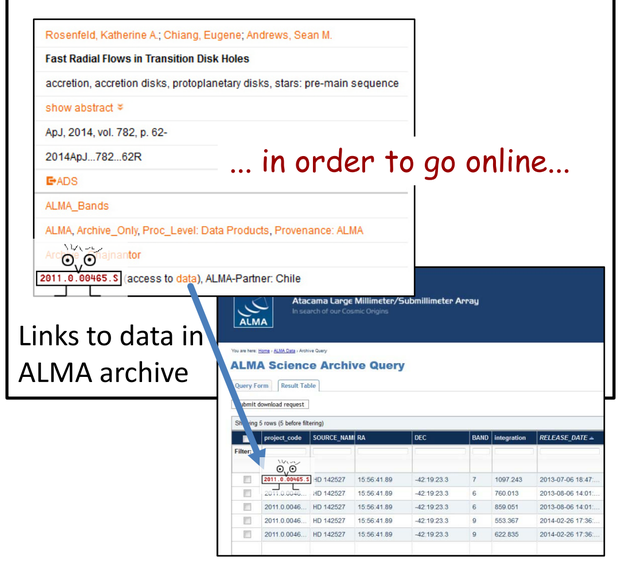


4 ALMA bibliography parameters:
- Program IDs
- Proposal types
- ALMA partners
- Archive tags

… and added to the bibliography…



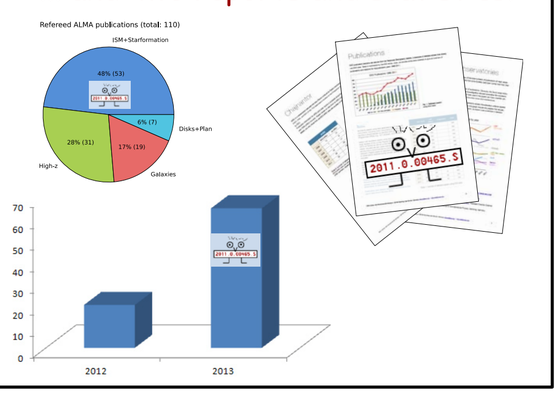
5 Public interface



… in order to go online…

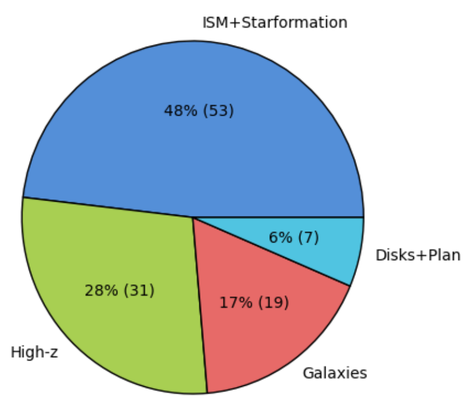
Links to data in ALMA archive

6 … and into reports and statistics.



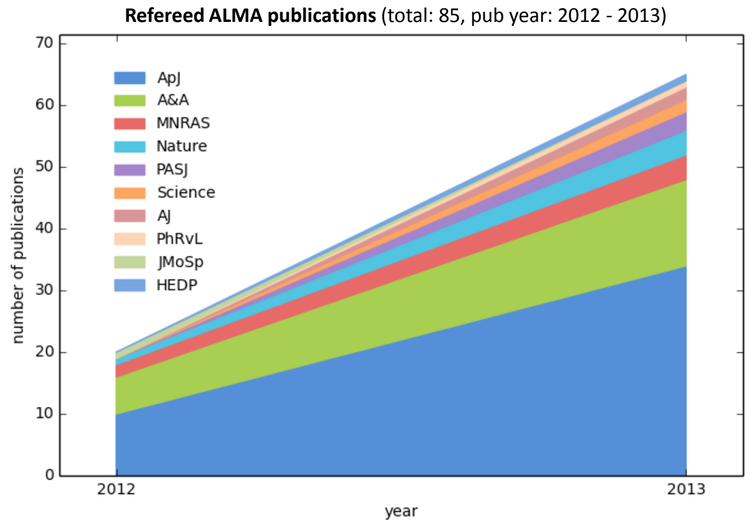
STATISTICS

Refereed ALMA publications (total: 110, pub year: 2012 - Apr 2014)



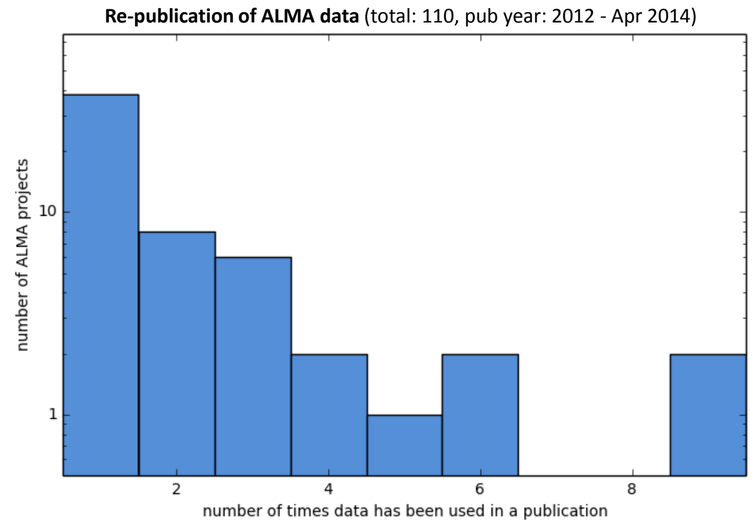
Subject Category: Currently, approx. half of the publications are done with data from the "ISM+Starformation" category, which was also the category that received the largest request for time.

Refereed ALMA publications (total: 85, pub year: 2012 - 2013)



Journal distribution: A large fraction (8%) of publications has been published in high-impact journals (Nature and Science).

Re-publication of ALMA data (total: 110, pub year: 2012 - Apr 2014)



Re-publications: Already now, there are several projects that have been published several times. The largest number of republications stem from one science verification and one standard project which have produced 9 papers.
Note the log y-axis.

LESSONS LEARNED

Regular discussion about policies and procedures among all librarians, ALMA archive specialists and ALMA Management is essential in order to provide a bibliography that is as **complete, consistent, and multi-faceted** as possible

Using a **semi-automatic tool** such as the FUSE Fulltext Search system **facilitates the identificaton** of relevant papers enormously.

Information on how data should be acknowledged in papers has to be **communicated to PIs/CoIs clearly and repeatedly**. It is also helpful to inspect preprints and to "educate" authors by contacting them if the programID is missing

Visual inspection of papers by the bibliography curators is necessary in order to spot cases where program IDs are missing, incomplete, or wrong.

Publication statistics are **key performance indicators**; it is crucial to design, develop, and implement a **bibliography from the very beginning of the mission**.

SPIE Paper No. 9149-81, Montréal, Canada, June 2014