# A publication database for optical long baseline interferometry

Fabien Malbet $^{a\dagger}$ , Guillaume Mella $^{a\dagger}$ , Peter Lawson $^b$ , Esther Taillifet $^{a\dagger}$ , Sylvain Lafrasse $^{a\dagger}$ ,  $^a$ Lab. d'Astrophysique de Grenoble (LAOG), UMR 5571 Univ. J. Fourier/CNRS, BP 53, F-38051 Grenoble cedex 9, France;

<sup>b</sup>Jet Propulsion Lab., California Institute of Technology, Pasedena, CA 91109, USA; 

†Jean-Marie Mariotti Center (JMMC), France;

# **ABSTRACT**

Optical long baseline interferometry is a technique that has generated almost 850 refereed papers to date. The targets span a large variety of objects from planetary systems to extragalactic studies and all branches of stellar physics. We have created a database hosted by the JMMC and connected to the Optical Long Baseline Interferometry Newsletter (OLBIN) web site using MySQL and a collection of XML or PHP scripts in order to store and classify these publications. Each entry is defined by its ADS bibcode, includes basic ADS informations and metadata. The metadata are specified by tags sorted in categories: interferometric facilities, instrumentation, wavelength of operation, spectral resolution, type of measurement, target type, and paper category, for example. The whole OLBIN publication list has been processed and we present how the database is organized and can be accessed. We use this tool to generate statistical plots of interest for the community in optical long baseline interferometry.

**Keywords:** Astronomical software, optical, infrared, interferometry, bibliography

#### 1. INTRODUCTION: RATIONALE

Optical interferometry is a technique which requires a high level of critical subsystems illustrated by the fact that one needs to control at the nanometer level optical path difference which can reach several hundred meters, or, to operate several telescopes with some level of adaptive optics. Furthermore, even for the common professional astronomer the link between the measurements and the astrophysical consequences consists in numerous mathematical operations which are not straightforward to understand. Therefore, despite important financial and human investments, it seemed for a while that the astrophysical return was first limited and then restrained to a few specialized areas even though the gain in spatial resolution is recognized as a real breakthrough.

The distance between firstly the efforts and the necessary support from the astronomical community and secondly the results contained in the peer-reviewed literature both in instrumentation but also for the astrophysical advances have led the community to get organized and to publicize its results. This was achieved first by establishing a common point of reference, the web site OLBIN (Optical Long Baseline Interferometry Newsletter edited by P. Lawson, see contribution on this subject in this volume), by forming the IAU commission  $\#54^*$  and by tracking the publication record in the field.

In 2000, the rate of refereed papers published in interferometry of about 30 papers per year was still manageable by hand but ten years later this rate has reached about 100 papers/year and it is still growing. The need to record any new reference in the field is even stronger but it can no longer be done by hand. Therefore we have built a database based on today software capability which allows us to track the evolution of the field by using new information that add extra value for the service to the community.

#### 2. A BIBLIOGRAPHIC DATABASE DIRECTLY LINKED TO ADS

We decided to restrict ourself to publications which have gone through the well-recognized peer-reviewed system in order to guaranty the quality of the papers. As a matter of fact, this type of publication are usually the ones that agencies but also authors want to refer to.

Send correspondence to F. Malbet, E-mail: Fabien.Malbet@obs.ujf-grenoble.fr

<sup>\*</sup>http://olbin.jpl.nasa.gov/iau

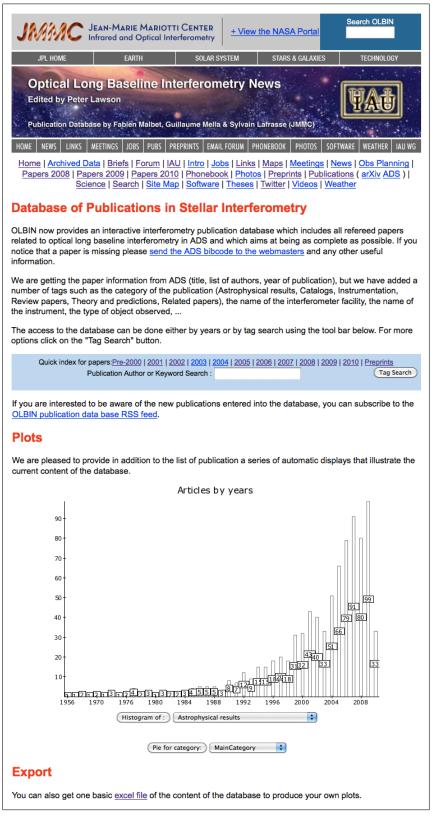


Figure 1. Homepage of the OLBIN publication database.

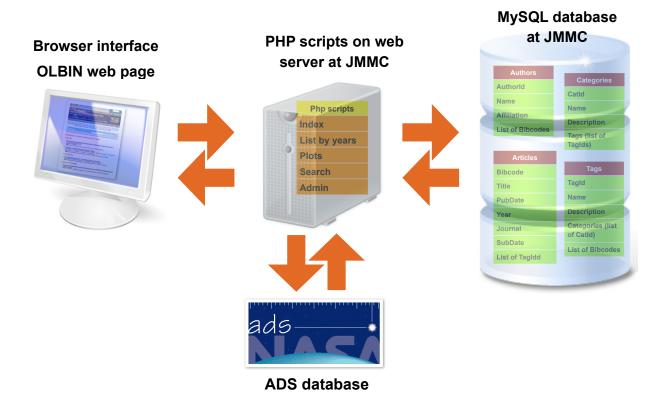


Figure 2. General architecture of the OLBIN publication database

#### 2.1 ADS: the primary source of information

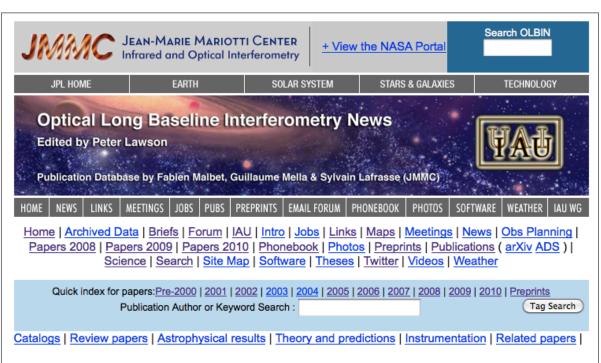
To build this publication database, we want to avoid to enter any already existing information. In astronomy, we can count on the NASA Astrophysics Data System<sup>1</sup> (ADS) which has become the reference in our field covering not only astronomy but also domains like optics and physics. In any case, ADS is very open to inputs from the community through their feedback submission forms on the Internet. ADS provides many information on the publication, including the full bibliographic reference, the list of authors, the year of publications, but also links to the papers on the publisher sites, to the arXiv preprint database, citation rate and other useful information.

Our first goal is to provide an up-to-date list of publication ordered by year of publication and by type of subject (see for example such a listing in Fig. 3 for year 2008). Therefore the output to the user is a web page listing the articles with links to the ADS abstract page of these papers. The user can access to the paper they want by just clicking on the link and are redirected to the ADS pages. There is therefore no need to duplicate the information stored in ADS.

The central entry in the OLBIN publication database is therefore the ADS bibcode. The ADS bibcode is a code made of 19 letters which includes the year of publication, the short name of the publication, the volume, the page and the last name initial of the first author. The ADS abstract page can be accessed through the HTTP protocol using the ADS mirrors address completed with the abs/bibcode link. To enter a new entry in the database, one needs only to enter the ADS bibcode.

However for classification purposes, the title, the year of publication, the list of authors, the affiliations<sup>†</sup> and the bibliographic references (journal, volume and pages) have to to be stored so that the automatic generation of the listings can be performed without having to connect the ADS database.

 $<sup>^{\</sup>dagger}$ Unfortunately, ADS does not provide a way to separate the affiliations of the different authors so that this has to be handle by our database.



# Bibliography in Stellar Interferometry in 2008

# Catalogs (2)

 The angular sizes of dwarf stars and subgiants. Non-linear surface brightness relations in BVR\_clc from interferometry

Kervella P., Fouqué P.

2008 - Astronomy and Astrophysics, Volume 491, Issue 3, 2008, pp.855-858

The Palomar Testbed Interferometer Calibrator Catalog
van Belle G. T., van Belle G., Creech-Eakman M. J., Coyne J., Boden A. F., Akeson R. L., Ciardi D. R.,
Rykoski K. M., Thompson R. R., Lane B. F., PTI Collaboration
 2008 - The Astrophysical Journal Supplement Series, Volume 176, Issue 1, pp. 276-292.

# Review papers (1)

Advanced Imaging Methods for Long-Baseline Optical Interferometry
 Le Besnerais G., Lacour S., Mugnier L. M., Thiebaut E., Perrin G., Meimon S.

 2008 - IEEE Journal of Selected Topics in Signal Processing, Vol. 2, Issue 5, p.767-780

# Astrophysical results (43)

- MWC 297: a young high-mass star rotating at critical velocity
   Acke B., Verhoelst T., van den Ancker M. E., Deroo P., Waelkens C., Chesneau O., Tatulli E., Benisty M., Puga E., Waters L. B. F. M., Verhoeff A., de Koter A.

   2008 Astronomy and Astrophysics, Volume 485, Issue 1, 2008, pp.209-221
- CHARA Array Measurements of the Angular Diameters of Exoplanet Host Stars
   Baines Ellyn K., McAlister Harold A., ten Brummelaar Theo A., Turner Nils H., Sturmann Judit, Sturmann Laszlo, Goldfinger P. J., Ridgway Stephen T.

   2008 The Astrophysical Journal, Volume 680, Issue 1, pp. 728-733.

Figure 3. Lay-out of the OLBIN publication database listing for year 2008.

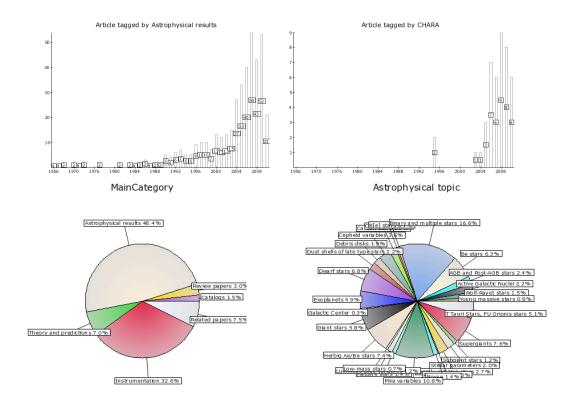


Figure 4. Plots that are automatically from the index OLBIN publication homepage. Top: histograms; bottom pie charts.

# 2.2 Accessing to the database: dynamical pages

The main outputs of the database are dynamical pages that can computed on the fly. This is performed using scripts that can form the HTML code that is then read by the user's Internet browser.

There are several scripts. Some of them are used only for administrative reasons. The main scripts are:

- list of publication for a given year. This is the way the information appeared historically on the OLBIN site. We have kept this functionality, the only change is that these pages are now automatically created upon the user's request. The year list is then ordered by type of publication and by alphabetical order of the first author.
- search page which can retrieve all bibliography entries whose title, author list contain the searched keys. We can also specify in this list the tags to be searched. A combination of tags restricts the list to the entries which contains these tags.
- plots automatically generated with tags. Two types of plots can be produced: histogram of number of papers published per year for a given tag, or, pie chart based on the category of tags (see Fig. 4).
- access to the database content under the form of a comma-separated-values (CSV) file which can be analyzed with usual spreadsheet software.
- administrative pages solely accessible by the OLBIN editors which allow them to add new entries by providing the bibcodes and checking the associated tags, to manage the list of tags and categories of tags, to search the database for entries and manage the checked tags, to delete an entry.

One can check the tags associated with an entry by putting the mouse pointer over the entry link in all dynamical pages generated by the database.

0	A	B	C	D	E	F	G	H		J	K	L	M	N	0	P Q	R
	bibcode	title	journal	pubdate	authors	year	tag VLTI	tag IOTA	tag PTI	tag CHARA	tag AMBER	tag MIDI	tag PRIMA	tag MIRC	tag VEGA	tag JMMC_gri tag NPOI	tag Review p
		140 Ruling Out Pt					10										
		S18 SIO Masers is					10										
4	2010ApJ	.716 Is Eta Carina	The Astrophy	2010-08-01	0( Groh, J. H.:W	2	010	1			1						
5	2010ApJ.	.715 Luminosity-va	The Astrophy	2010-06-01	Of Ghez, Andrea	2	10										
6	2010A.I	139 Separated Fr	The Astronor	2010-08-01	Of Farrington, C.	2	010										
		139 The Spatially					10										1
		.51 The Hit line 1					10										-
		51 First step to c					010										
		51 Filipsoidal pri					010										
								1								1	
		51 Angular diam					010	1								1	
		6A Radiative hyd					010										
		P12 Simultaneous					10										
		715 The Algol Trip					010										1
15	2010Natur	r.46 Infrared imag	Nature, Volum	2010-04-01	Of Kloppenborg.	2	10								1		
16	2010An.I	.712 The Dust Sur	The Astrophy	2010-04-01	Of Wishnow F. I	2	110										
		26S Target star or				2	10										
18	20101894	J4 Flight-Like Gr	IFFF System	2010-03-01	Of Scharf Danie	2	10										
19	20101918	J4 Flight-Like Gr	IEEE System	2010-03-01	Of Schorf, Donie	2	110										
		27 A new concer					110	4									
							110	1									
21	2010ApJS	318 SiO Masers is	The Astrophy	4 2010-03-01	Ut Cotton, W. D.												
		51 Resolving the					10	1					1				
		6A The radius ar					010										
24	2010ApJ	71(Angular Diam	The Astrophy	₹2010-02-01	O(Baines, Ellyn		110										
		.71( Ruling Out St					10										
26	2010A&A	51 VLTVAMBER	Astronomy a	2010-02-01	Of Chiavassa, A.	2	010	1			1					1	
27	2010A%2	6A., Strong starlig	Astronomy a	2010-02-01	Of Martin, S. R.:	2	10										
		6A The complex					110	1								1	
		6A. Strong near-					110	1								- 1	
30	20108488	P., 12 Mid-Infrared A	The Publicati	2010-01-01	Of Peters Rober	2	10										
		12 Image recons					010										_
		70! A New View of					010										
		50 Phase closur					110										1
								1			1					1	
		6A Nulling interfe					010										
		51 The origin of					009	1					1				
		AA., Two-wavelen					009										
37	2009JOpt	A1 Multi-spectral	Journal of Op	2009-12-01	0(HA©nault, Fri	2	009										
38	2009ApJS	318 SIO Masers II	The Astrophy	2009-12-01	Of Cotton, W. D.	2	009		1								
39	2009Ap.I	701 Starspot Jitte	The Astrophy	2009-12-01	Of Makarov V V	2	009										
40	2009An.I	.707 The Pulsation	The Astrophy	2009-12-01	Of Lacour, S. Th	2	009		1								
		.70f Parasitic Inte					009										
		Rytr Formation an					009										
		50 VEGA: Visible					009										
		50 VEGA: VISION					009										
		50 Imaging the s					109										
								1									
		50 Exploring the					009										
		50 The PRIMA fr					009	1									
		6A Intensity inter					009										
		pt4 Closure phas					009										
50	2009ApJ	.70! Dust Emissio	The Astrophy	2009-11-01	00 Burtscher, L.:	. 2	009	1					1				
51	2009ApJ	.70! Interferometri	The Astrophy	2009-11-01	Ol Csizmadia, S.	2	009										
		50 A binary engi					009	1			1		1			1	
		50 Resolving the					009	1					1			- 1	
54	20000080	50 VLTI/AMBER	Astronomy a	2000 11-01	Of Millour, E. Drie	2	109	4								- 1	
		50 Radiative hyd					009		4							- '	
									1								
		AS. A list of bright					009	1									
57	2009ApO;	pt4 Polarimetric s	Applied Option	2009-10-01	Ol Murakami, Na		009										
		MALLEN AND A	which the beautiful	70000 10 01	Articles, a pr	^	AAA .										

Figure 5. Comma-separated-values (CSV) file that is automatically from the index OLBIN publication homepage which can be processed by any spreadsheet software.

Since we are using ADS, the first thing to do when entering a paper is to check that it is in ADS and, if not, we have to request it to be registered. Similarly, if mistakes are found, then they should be corrected in ADS since it is the most used publication database in astronomy.

#### 2.3 Metadata: publication tags

In order to add extra-values to the database and to provide other way of classifying the resulting lists, we added the notion of *tags* to each entry to better define the different entries. Any paper can be labeled by any number of tags. In order to sort out the database, we also created *categories* of tags which are lists of tags of same nature: type of papers, facilities, instruments, astrophysical topics, technique,...

At the time being, there are about 70 tags which label the papers in the OLBIN database. We have classified them into 7 categories:

- Type of papers: Astrophysical results, Catalogs, Instrumentation, Related papers, Review papers, Theory and predictions;
- Facility: CHARA, COAST, GI2T, I2T, IOTA, IRMA, ISI, Keck, LBTI, Mark III, Narrabri Stellar Intensity Interferometer (NSII), NPOI, PTI, SIM, SUSI, VLTI;
- Instrument: AMBER, CHARA Classic, FLUOR, IONIC, MIDI, MIRC, PRIMA, VEGA, VINCI;
- Astrophysical topic: AGB and Post-AGB stars, Active Galactic Nuclei, B[e] stars, Be stars, Binary and multiple stars, Calibrators, Carbon stars, Cepheid variables, Debris disks, Dust shells of late type stars, Dwarf stars, Exoplanets, Galactic Center, Giant stars, Herbig Ae/Be stars, Low-mass stars, Luminous Blue Variables, Massive stars, Mira variables, Novae, R CrB stars, Rapidly rotating stars, Stellar parameters, Subgiant stars, Supergiants, T Tauri Stars, FU Orionis stars, Wolf-Rayet stars, Young massive stars;
- Wavelength of operation: visible, near-infrared (NIR), mid-infrared (MIR);
- Spectral resolution mode: broad-band, narrow-band, low resolution ( $\mathcal{R} \leq 100$ ), medium resolution ( $100 \leq \mathcal{R} \leq 3000$ ), high resolution ( $3000 \leq \mathcal{R}$ );
- **Technique**: Astrometry, Closure phases, Differential phase, Fringe tracking, Images, Intensity interferometry, Nulling, Phase reference, Squared visibilities,...

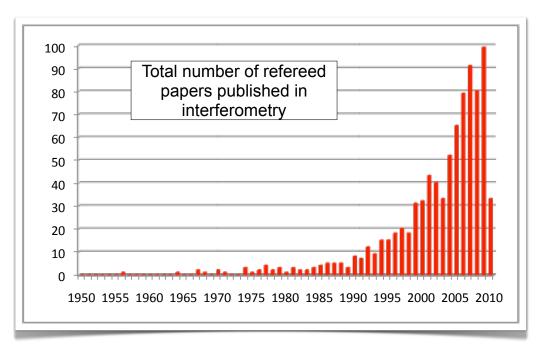


Figure 6. Total number of refereed publications published since 1950.

Adding tags must be easy as well as reorganizing the categories of tags. However, one should keep in mind that when adding tags, the whole database should be scanned again in order to make sure that no entries have been forgotten. This is the main weakness of this architecture.

#### 2.4 Technical solution

Figure 2 represents the technical architecture which has be chosen to build the OPLBIN publication database. This database is based on MySQL and it is located in the  $JMMC^{\ddagger}$  server in Grenoble. This database contains different tables with the following information attached:

- Articles: Bibcode, Title, PubDate corresponding to the publication date, Year, Journal, SubDate corresponding to the date of submission in the database, list of TagId which identify to the tags.
- Authors: AuthorId, Name of the author, Affiliation of the author, list of attached Bibcode.
- Tags: TagId which uniquely identify the tag, Name of the tag, Description of the tag (not always filled in), list of attached CatId, list of attached Bibcode.
- Categories: CatId which uniquely identifies the category, Name of the category, Description of the category, list of attached TagId which are gathered in this category.

The MySQL database can be accessed through XML or PHP scripts which construct dynamically the pages that can been read using a browser. Input data can also be introduced through forms in order to store new information mainly in administrative mode. XML scripts are used to interact with ADS when a new entry is given.

#### 3. CONTENT OF THE OLBIN PUBLICATION DATABASE

In this Section, we would like to give examples of applications concerning the publication in interferometry. They originate directly from the data base and the use of the CSV export file.

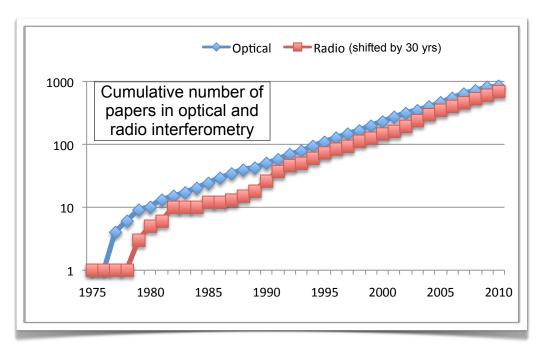


Figure 7. Evolution of the cumulative number of refereed publications published since 1950 between the optical (blue losanges) and the radio (red squares) domains.

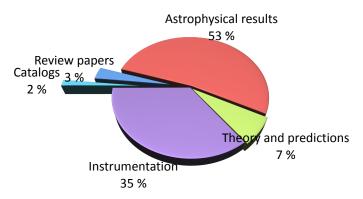


Figure 8. Repartition of the different types of refereed articles.

# 3.1 Growth of interferometry results

At the date of presentation of this work, there is a total of more than 850 refereed publications in the database. There is also about 70 tags and 7 categories. The evolution of the number of publications is displayed in Fig. 6. The growth is exponential with a rate of publication less than 5 papers per year in the 1980's that reached 30 papers per year in 2000 and almost 100 papers per year in 2010.

An interesting way to analyze this growth is to compare it to the growth of papers produced by radio-interferometers at the beginning of the domain. Figure 7 shows the cumulative number of papers in optical and infrared interferometry (blue losanges) compared to radio-interferometry (red squares). If we shift the radio curve by 30 years then it becomes obvious that optical interferometry follows the same type of growth of radio-interferometry.

The type of papers is presented in Fig. 8. More than half the papers correspond to actual astrophysical results, 35% of them are articles on instrumentation and concept and in the remaining 10% we have theory and

<sup>&</sup>lt;sup>‡</sup>Jean-Marie Mariotti Center, http://www.jmmc.fr

Table 1. Tags of the main category astrophysical topics sorted in intermediate astrophysical sub-categories.

Stellar parameters	Evolv	ed stars	Hot active stars
Supergiants	R CrB stars	Mira variables	Massive stars
Giant stars	Wolf-Rayet stars	Cepheid variables	Luminous Blue Variables
Subgiant stars	Novae	Carbon stars	B[e] stars
Dwarf stars	Dust shells of late	AGB and Post-AGB	Be stars
Rapidly rotating stars	-type stars		

Young stellar objects	Low-mass objects	Galaxies	Multiple stars
T Tauri Stars, FU Orionis	Low-mass stars	Galactic Center	Binary and multiple systems
Herbig Ae/Be stars	Exoplanets	Active Galactic Nuclei	
Young massive stars	Debris disks		

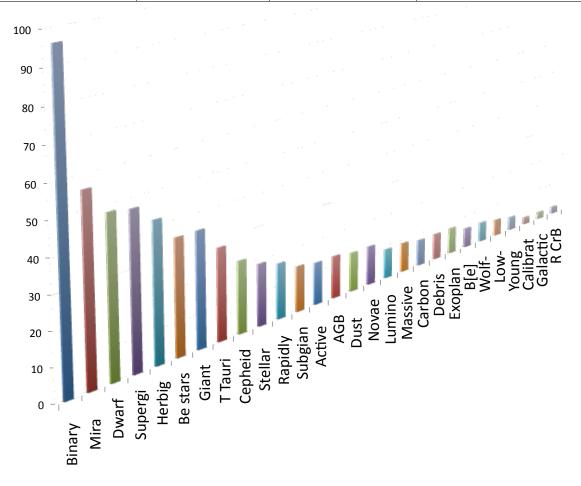


Figure 9. Repartition of the number of publications per topics.

predictions, catalogs and reviews.

# 3.2 Astrophysical topics

Figure 9 shows the repartition of papers published per topics in decreasing order. The number of topics are too detailed to distinguish a particular trend. We have therefore defined general astrophysical categories which gather several topics (see Table 1). Figure 10 summarizes the share of the different astrophysical main categories. The largest category concerns the stellar parameters (31%) which is not surprising since it is the first science

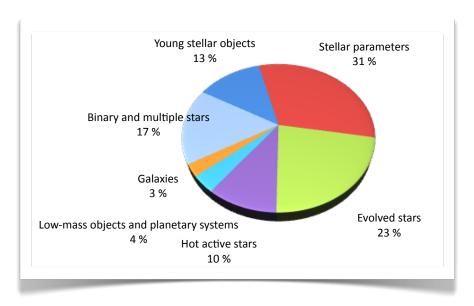


Figure 10. Pie chart of the main science categories of the OLBIN publication database.

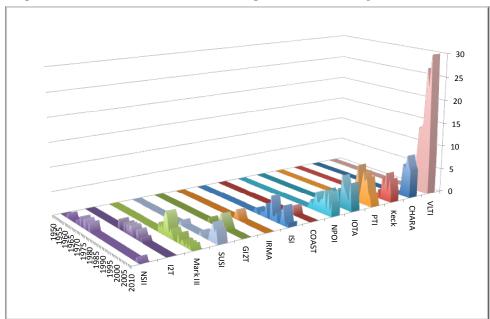


Figure 11. Evolution of the number of publication which have been tagged to a past or operating facility.

that has been achieved by optical interferometry. The second largest category consists of evolved stars (23%) which have benefited from the high brightness of some of them. Multiple systems (17%) are playing an important role as well, whereas young stellar objects constitute the fourth category (13%) which has been growing a lot during this last decade. The remaining stellar categories are hot active stars and low-mass objects and planetary systems. The last objects are more difficult to observe. Finally extragalactic observations already amount to 3% of the total number of astrophysical results.

# 3.3 Facilities

Figure 11 shows the evolution of astrophysical results published per facility. One can notice the historical facilities (NSII, I2T, Mark III, IRMA, COAST, GI2T), the more modern installations (ISI, SUSI, NPOI, IOTA, PTI) to

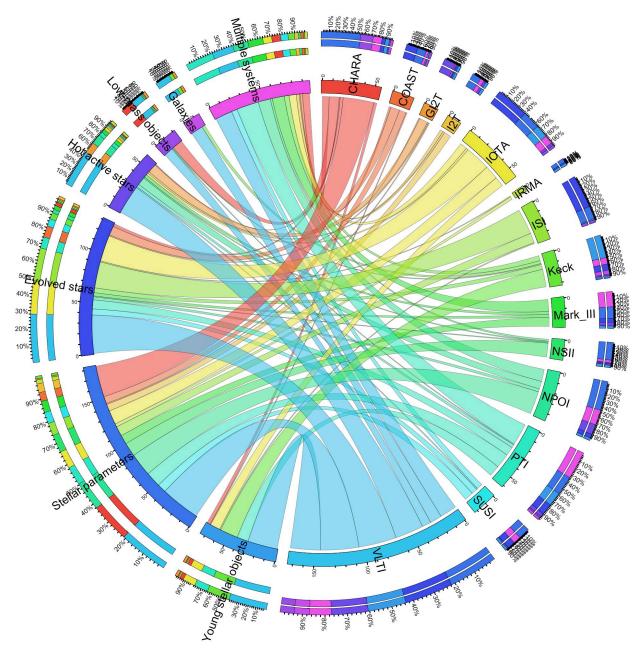


Figure 12. Table of the astrophysical topic vs the facility represented with the Circos Table viewer (see http://mkweb.bcgsc.ca/circos).

the latest large facilities (Keck, CHARA and VLTI). The fact that the VLTI is offered to a large community boosts its number of publication.

The diagram displayed in Fig. 12 is particularly interesting. It displays the correspondence between two categories: the astrophysical topics (left) and the facilities (right). For example, VLTI has a rather broad range of topics compared to ISI. On the other side, low-mass objects have been tackled mainly by CHARA and the VLTI.

#### 4. CONCLUSION AND PERSPECTIVES

We have described in this paper the implementation of a publication database to list all refereed publication in optical long baseline interferometry. We have included some statistical plots which illustrate the possible applications of this database. We need the feedback of the users to correct the citations, the tags and be aware of all papers.

This tool could be the reference for the different groups in order to list their publications. We may develop then specific pages (instruments, interferometers, science,..). These pages might also be used by our agencies to evaluate the outcome of interferometry. We could also contemplate to get the citations rate from ADS, but since it changes everyday basically, it would require to update the database on a daily basis. Another important perspective is to create a direct link between the publications and the published data.

Finally do not forget to use this database in your work:

http://olbin.jpl.nasa.gov in the *Publications* menu or http://www.jmmc.fr/bibdb/

# ACKNOWLEDGMENTS

We would like to thanks all our colleagues who have discussed with us the opportunity of such a tool. We are especially grateful to S. Ridgway who compiled the number of papers published in radio-interferometry.

#### REFERENCES

[1] Kurtz, M. J., Eichhorn, G., Accomazzi, A., Grant, C. S., Murray, S. S., and Watson, J. M., "The NASA Astrophysics Data System: Overview," A&AS 143, 41–59 (Apr. 2000).