CENTRE DE DONNÉES ASTRONOMIQUES DE STRASBOURG

EURO VO TECHNOLOGY CENTRE

Automatic recognition of object names in literature

C. Bonnin[1], S. Lesteven[1], A. Oberto[1], S. Derriere[1], F. Genova[1]

[1] CDS – Strasbourg

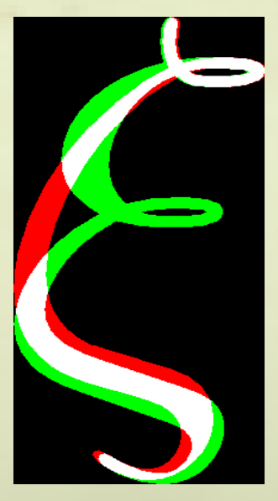
Bibliographic code :

2007MNRAS.374..176L

Connection to the Journal Bibliographical Service to retrieve the PDF document corresponding to the bibliographic code.

Raw PDF document

- Text extraction using PDFBox
- Pictures of special characters using JPedal
- Recognition of graphic symbols
- (here comparison between greek letters
- ζ and ξ)



Extracted text

- Dictionary of nomenclature
- SIMBAD

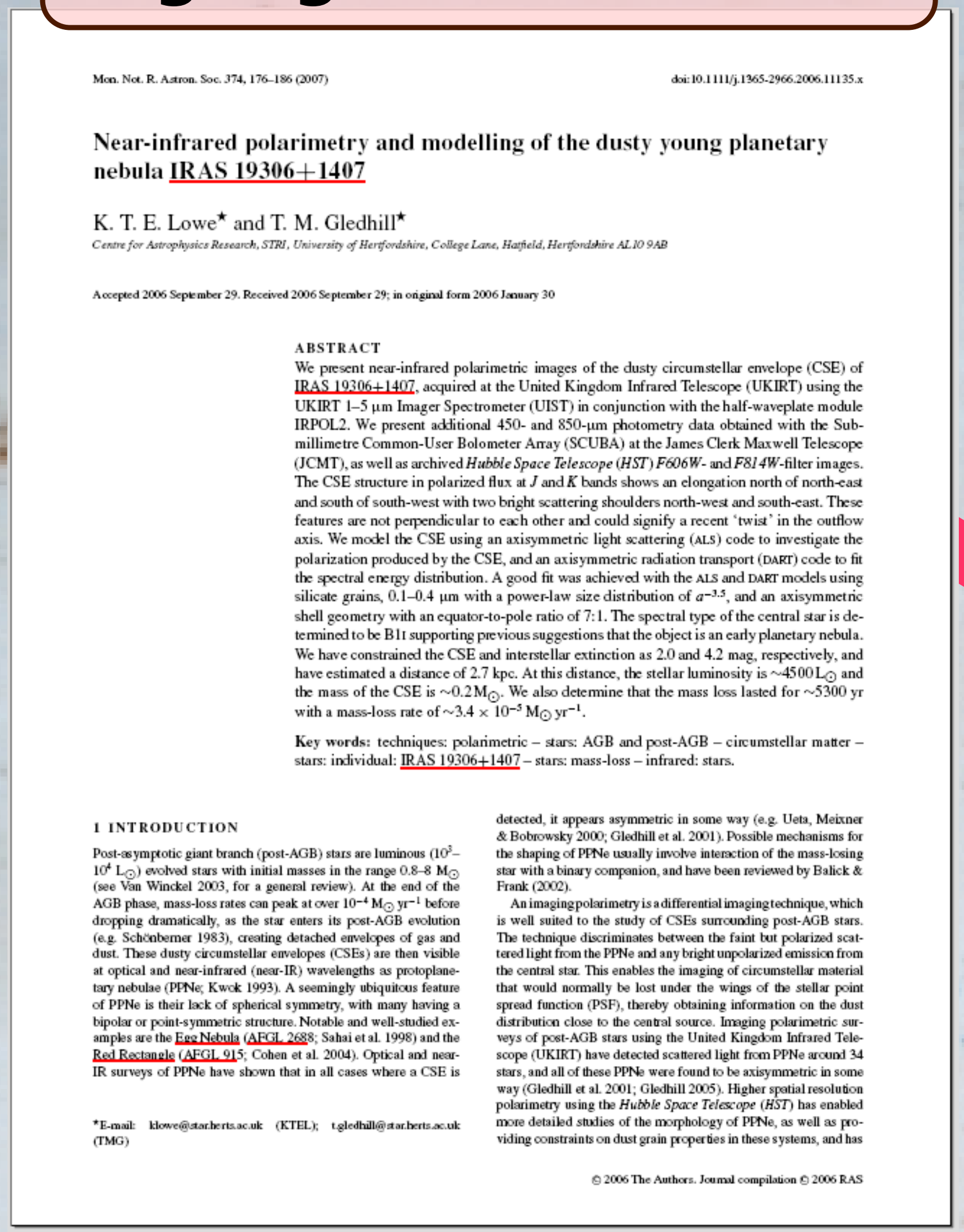
•Searching for object names using :
- the formats from the Dictionary of Nomenclature of Celestial Objects
- usual names such as Aldebaran or The Crab given by SIMBAD
- the variable star names based on constellation names

Use of the Weka Machine Learning engine to help detecting the false positives among the object names.

WEKA The University of Waikato

The object names are highlighted in the original document and displayed by Acrobat Reader.

Highlighted document



Mon. Not. R. Astron. Soc. 374, 176–186 (2007) doi:10.1111/j.1365-2966.2006.11135.x

Near-infrared polarimetry and modelling of the dusty young planetary nebula IRAS 19306+1407

K. T. E. Lowe★ and T. M. Gledhill★
Centre for Astrophysics Research, STRI, University of Hertfordshire, College Lane, Hatfield, Hertfordshire AL10 9AB

Accepted 2006 September 29. Received 2006 September 29; in original form 2006 January 30

ABSTRACT

We present near-infrared polarimetric images of the dusty circumstellar envelope (CSE) of IRAS 19306+1407, acquired at the United Kingdom Infrared Telescope (UKIRT) using the UKIRT 1–5 μm Imager Spectrometer (UIST) in conjunction with the half-waveplate module IRPOL2. We present additional 450- and 850-μm photometry data obtained with the Sub-millimetre Common-User Bolometer Array (SCUBA) at the James Clerk Maxwell Telescope (JCMT), as well as archived *Hubble Space Telescope* (*HST*) *F606W*- and *F814W*-filter images. The CSE structure in polarized flux at *J* and *K* bands shows an elongation north of north-east and south of south-west with two bright scattering shoulders north-west and south-east. These features are not perpendicular to each other and could signify a recent 'twist' in the outflow axis. We model the CSE using an axisymmetric light scattering (ALS) code to investigate the polarization produced by the CSE, and an axisymmetric radiation transport (DART) code to fit the spectral energy distribution. A good fit was achieved with the ALS and DART models using silicate grains, 0.1–0.4 μm with a power-law size distribution of $a^{-3.5}$, and an axisymmetric shell geometry with an equator-to-pole ratio of 7:1. The spectral type of the central star is determined to be B1I supporting previous suggestions that the object is an early planetary nebula. We have constrained the CSE and interstellar extinction as 2.0 and 4.2 mag, respectively, and have estimated a distance of 2.7 kpc. At this distance, the stellar luminosity is $\sim 4500\,L_\odot$ and the mass of the CSE is $\sim 0.2\,M_\odot$. We also determine that the mass loss lasted for ~5300 yr with a mass-loss rate of $\sim 3.4 \times 10^{-5}\,M_\odot\,yr^{-1}$.

Key words: techniques: polarimetric – stars: AGB and post-AGB – circumstellar matter – stars: individual: IRAS 19306+1407 – stars: mass-loss – infrared: stars.

1 INTRODUCTION

Post-asymptotic giant branch (post-AGB) stars are luminous (10^3–$10^4\,L_\odot$) evolved stars with initial masses in the range 0.8–8 $M_\odot$ (see Van Winckel 2003, for a general review). At the end of the AGB phase, mass-loss rates can peak at over $10^{-4}\,M_\odot\,yr^{-1}$ before dropping dramatically, as the star enters its post-AGB evolution (e.g. Schönberner 1983), creating detached envelopes of gas and dust. These dusty circumstellar envelopes (CSEs) are then visible at optical and near-infrared (near-IR) wavelengths as protoplanetary nebulae (PPNe; Kwok 1993). A seemingly ubiquitous feature of PPNe is their lack of spherical symmetry, with many having a bipolar or point-symmetric structure. Notable and well-studied examples are the Egg Nebula (AFGL 2688; Sahai et al. 1998) and the Red Rectangle (AFGL 915; Cohen et al. 2004). Optical and near-IR surveys of PPNe have shown that in all cases where a CSE is detected, it appears asymmetric in some way (e.g. Ueta, Meixner & Bobrowsky 2000; Gledhill et al. 2001). Possible mechanisms for the shaping of PPNe usually involve interaction of the mass-losing star with a binary companion, and have been reviewed by Balick & Frank (2002).

An imaging polarimetry is a differential imaging technique, which is well suited to the study of CSEs surrounding post-AGB stars. The technique discriminates between the faint but polarized scattered light from the PPNe and any bright unpolarized emission from the central star. This enables the imaging of circumstellar material that would normally be lost under the wings of the stellar point spread function (PSF), thereby obtaining information on the dust distribution close to the central source. Imaging polarimetric surveys of post-AGB stars using the United Kingdom Infrared Telescope (UKIRT) have detected scattered light from PPNe around 34 stars, and all of these PPNe were found to be axisymmetric in some way (Gledhill et al. 2001; Gledhill 2005). Higher spatial resolution polarimetry using the *Hubble Space Telescope* (*HST*) has enabled more detailed studies of the morphology of PPNe, as well as providing constraints on dust grain properties in these systems, and has

★E-mail: klowe@star.herts.ac.uk (KTEL); t.gledhill@star.herts.ac.uk (TMG)

© 2006 The Authors. Journal compilation © 2006 RAS

- **Results :**

- Improvement and automatic verification of the Dictionary of Nomenclature of Celestial Objects
- More verification of data entered in SIMBAD
- Documents treated more quickly and more exhaustively
- Help for the librarians to concentrate on the added value of their work
- Beta version currently in use

- More information will be made available to find the most relevant papers in the object reference lists (number and position of occurrences).

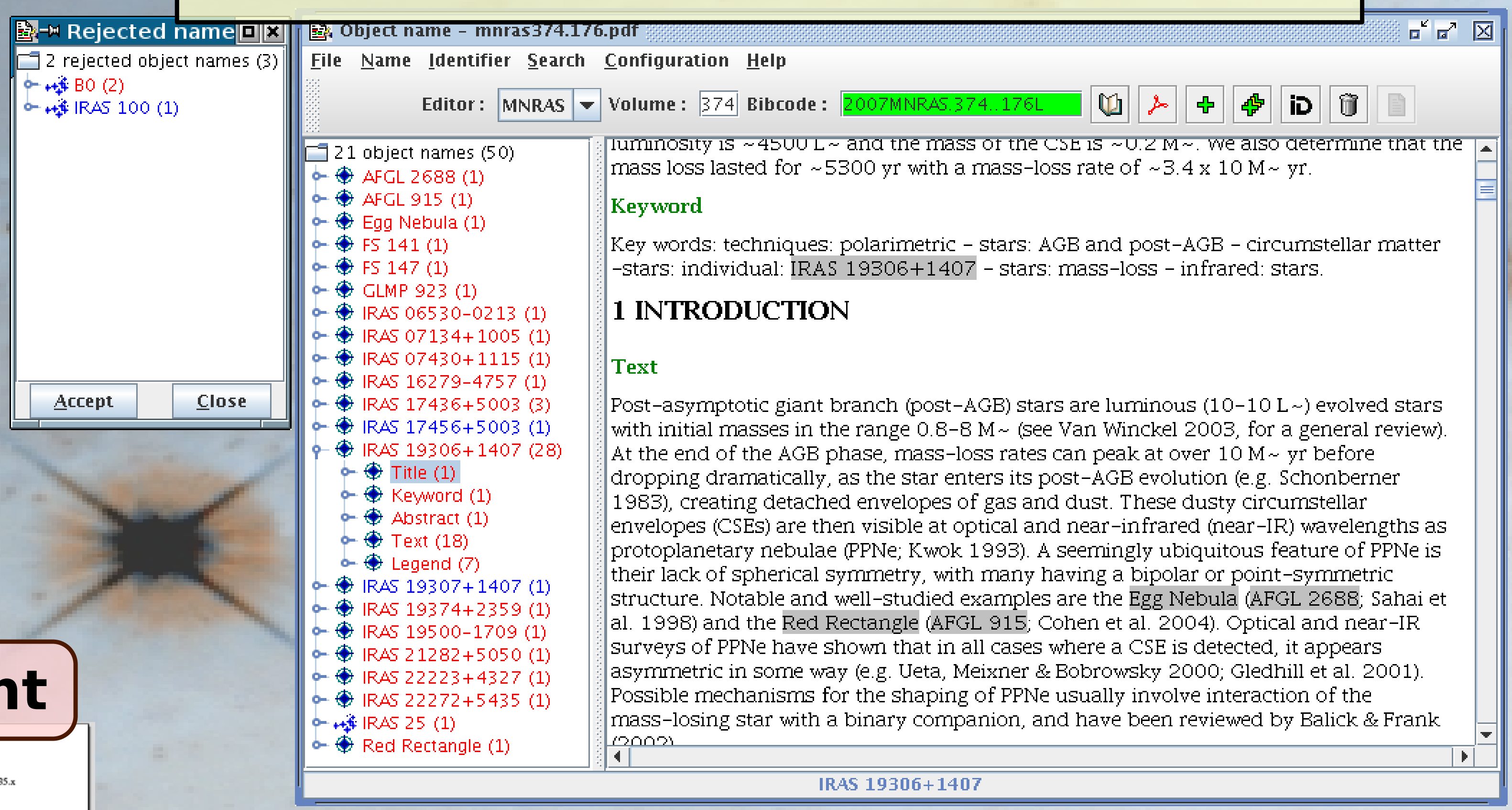
Today :

SIMBAD is a database of astronomical objects that provides (among other things) their bibliographic references in a large number of journals. Currently, these references have to be entered manually by librarians who read each paper.

- Objectives :
-

Detect object names in published articles and propose them for validation by the experts.
This includes :
- Search for possible object names in PDF documents
- Highlight the names in the documents without changing the presentation
- Retrieve information about the objects in SIMBAD
- Display the possibilities and let the librarian make the final choice
- Gather some additional information such as the number of occurrences and the positions in the document
- Enter the references and additional information in SIMBAD

A graphic user interface written in Java Swing can perform all these operations.



2 rejected object names (3)
- B0 (2)
- IRAS 100 (1)

Editor : MNRAS Volume : 374 Bibcode : 2007MNRAS.374..176L

21 object names (50)
- AFGL 2688 (1)
- AFGL 915 (1)
- Egg Nebula (1)
- FS 141 (1)
- FS 147 (1)
- GLMP 923 (1)
- IRAS 06530-0213 (1)
- IRAS 07134+1005 (1)
- IRAS 07430+1115 (1)
- IRAS 16279-4757 (1)
- IRAS 17436+5003 (3)
- IRAS 17456+5003 (1)
- IRAS 19306+1407 (28)
  - Title (1)
  - Keyword (1)
  - Abstract (1)
  - Text (18)
  - Legend (7)
- IRAS 19307+1407 (1)
- IRAS 19374+2359 (1)
- IRAS 19500-1709 (1)
- IRAS 21282+5050 (1)
- IRAS 22223+4327 (1)
- IRAS 22272+5435 (1)
- IRAS 25 (1)
- Red Rectangle (1)

luminosity is ~4500 L~ and the mass of the CSE is ~0.2 M~. We also determine that the mass loss lasted for ~5300 yr with a mass-loss rate of ~3.4 x 10 M~ yr.

Keyword

Key words: techniques: polarimetric – stars: AGB and post-AGB – circumstellar matter –stars: individual: IRAS 19306+1407 – stars: mass-loss – infrared: stars.

1 INTRODUCTION

Text

Post-asymptotic giant branch (post-AGB) stars are luminous (10–10 L~) evolved stars with initial masses in the range 0.8–8 M~ (see Van Winckel 2003, for a general review). At the end of the AGB phase, mass-loss rates can peak at over 10 M~ yr before dropping dramatically, as the star enters its post-AGB evolution (e.g. Schonberner 1983), creating detached envelopes of gas and dust. These dusty circumstellar envelopes (CSEs) are then visible at optical and near-infrared (near-IR) wavelengths as protoplanetary nebulae (PPNe; Kwok 1993). A seemingly ubiquitous feature of PPNe is their lack of spherical symmetry, with many having a bipolar or point-symmetric structure. Notable and well-studied examples are the Egg Nebula (AFGL 2688; Sahai et al. 1998) and the Red Rectangle (AFGL 915; Cohen et al. 2004). Optical and near-IR surveys of PPNe have shown that in all cases where a CSE is detected, it appears asymmetric in some way (e.g. Ueta, Meixner & Bobrowsky 2000; Gledhill et al. 2001). Possible mechanisms for the shaping of PPNe usually involve interaction of the mass-losing star with a binary companion, and have been reviewed by Balick & Frank (2002).

IRAS 19306+1407

Each object name is checked for existence in SIMBAD and the object description is retrieved. A librarian makes the final validation.

Identifiers list

Name	Position	Nb
AFGL 2688	text	2
AFGL 915	text	2
FS 141	text	1
FS 147	text	1
IRAS 06530-0213	text	1
IRAS 07134+1005	text	1
IRAS 07430+1115	text	1
IRAS 16279-4757	text	1
IRAS 17436+5003	text	4
IRAS 19306+1407	text, title, keyword, abstract, legend	30
IRAS 19374+2359	text	1
IRAS 19500-1709	text	1
IRAS 21282+5050	text	1
IRAS 22223+4327	text	1

- The information is entered directly into SIMBAD :
- Object identifier
- Name as it is written by the author
- Position of the object citation in the text (title, abstract, keyword, table, ...)
- Number of occurences

Update

SIMBAD

R&D in the framework of the VO-TECH Design Study

ADASS September 2007 London, UK