

# Information scientists: between editors and data centers

Marianne Brouty

F. Woelfel, S. Borde, C. Bruneau, C. Brunet, G. Chassagnard, H. Claude,  
P. Dubois, A. Eisele, F. Genova, S. Lesteven, M. Neuville, F. Ochsenbein,  
E. Perret, P. Vannier, P. Vonflie

LISA VI - Pune, India  
February 14-17, 2010

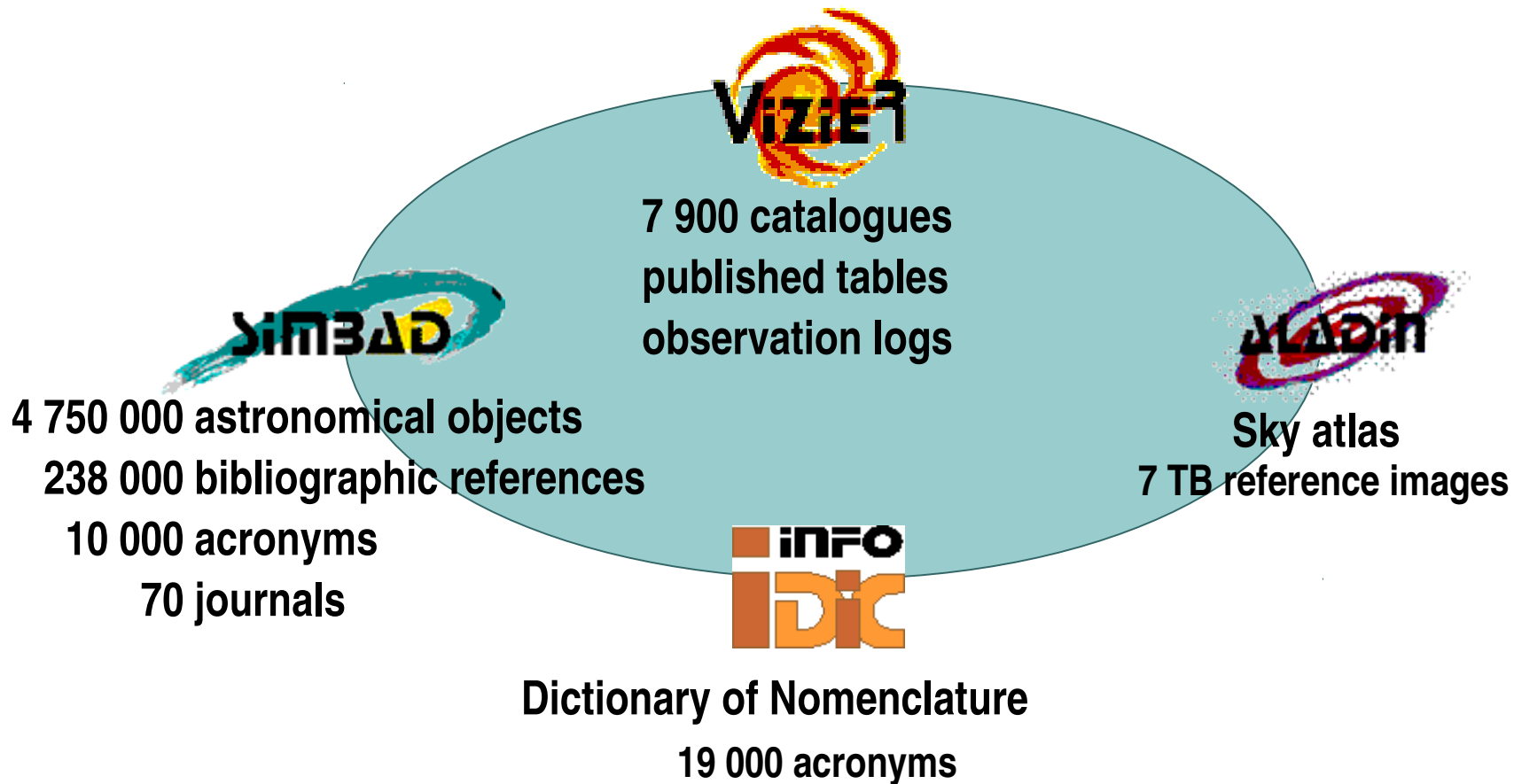


# Centre de Données de Strasbourg

- Founded in 1972
- Ongoing collaboration with other data centers
- On the Internet since 1991
- Involved in the publication of electronic tables of *Astronomy & Astrophysics* since 1993
- Develops and maintains on-line services

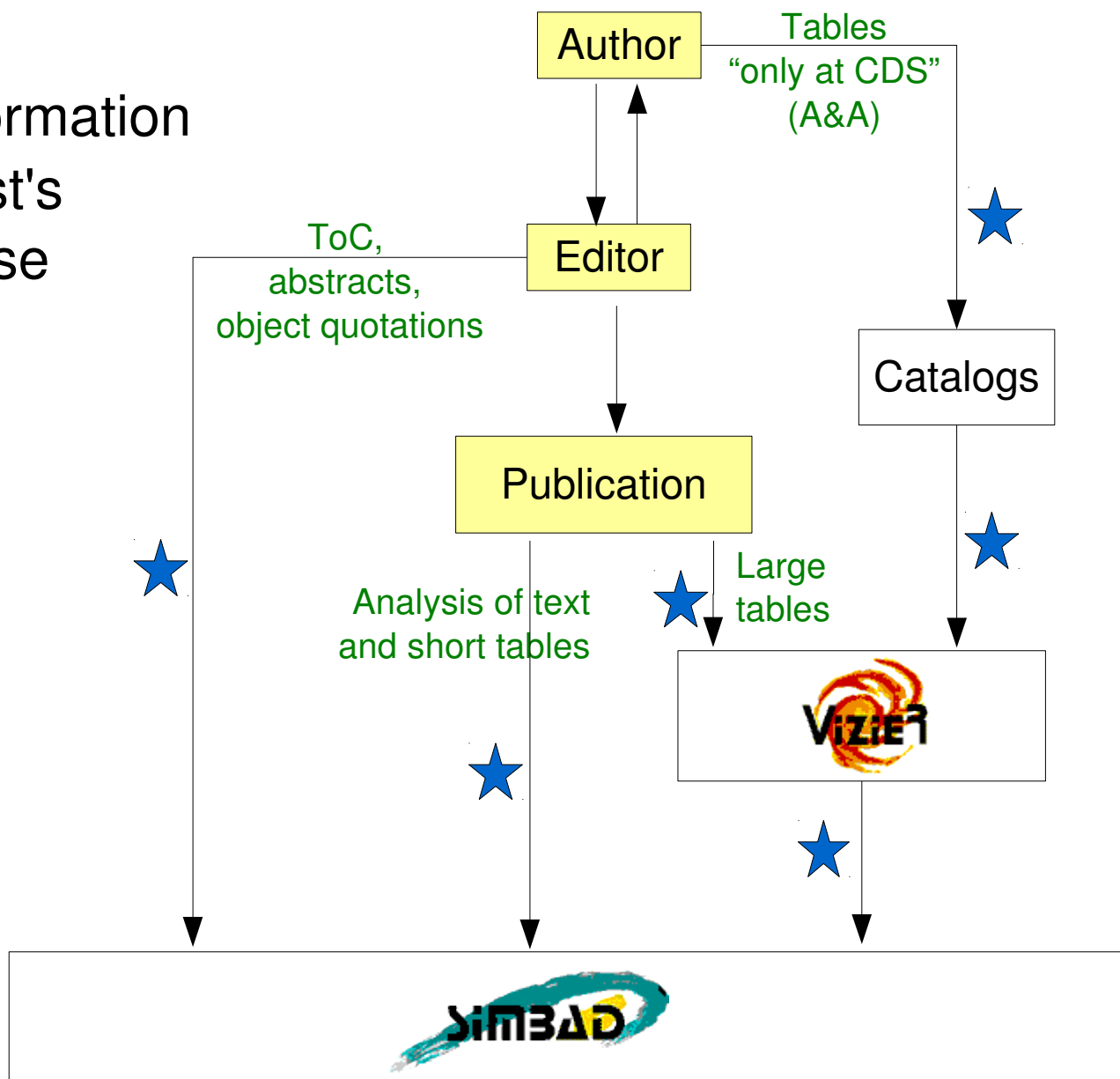


# The CDS on-line services

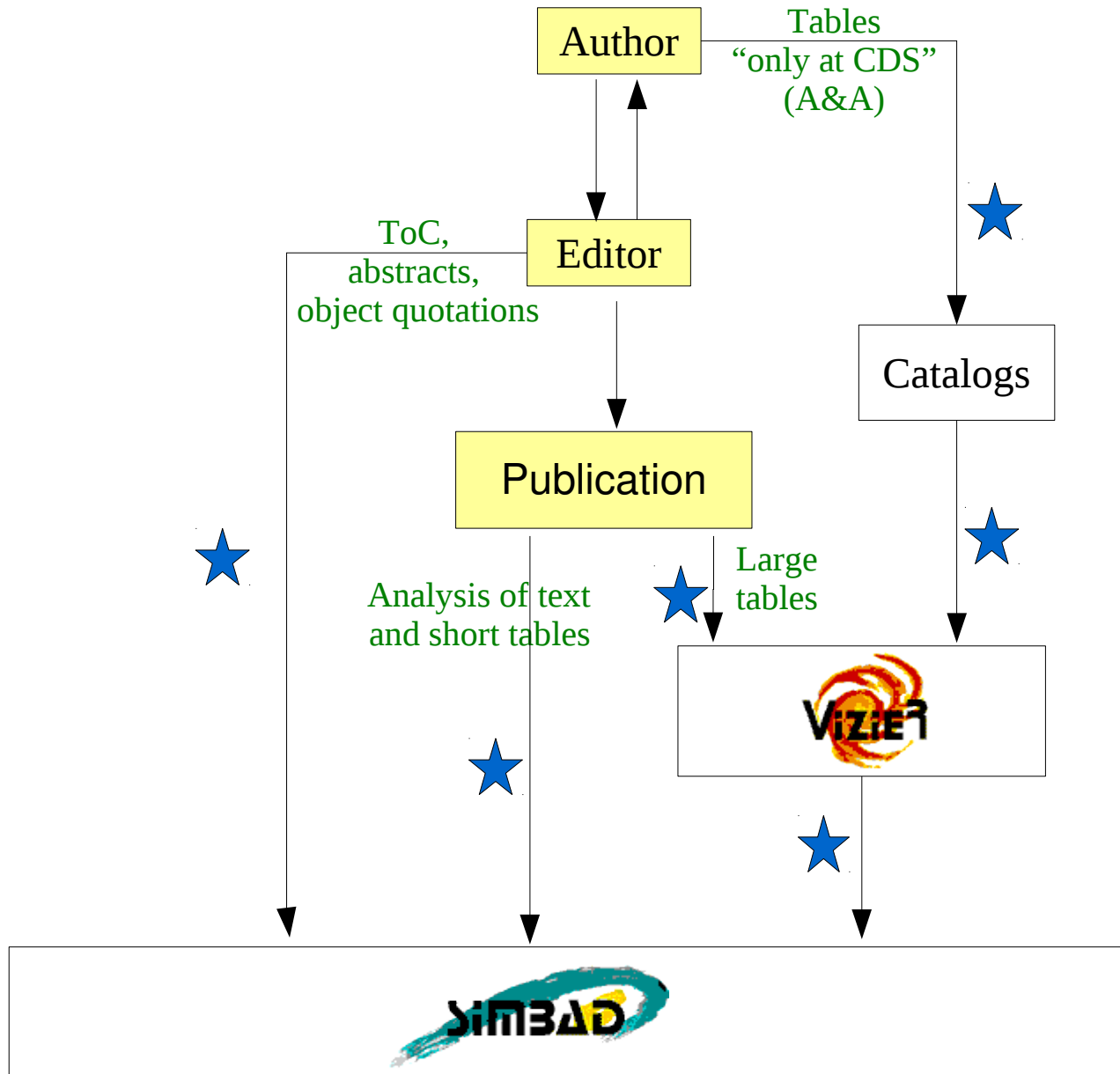


# Life cycle of an article

- ★ : information scientist's expertise

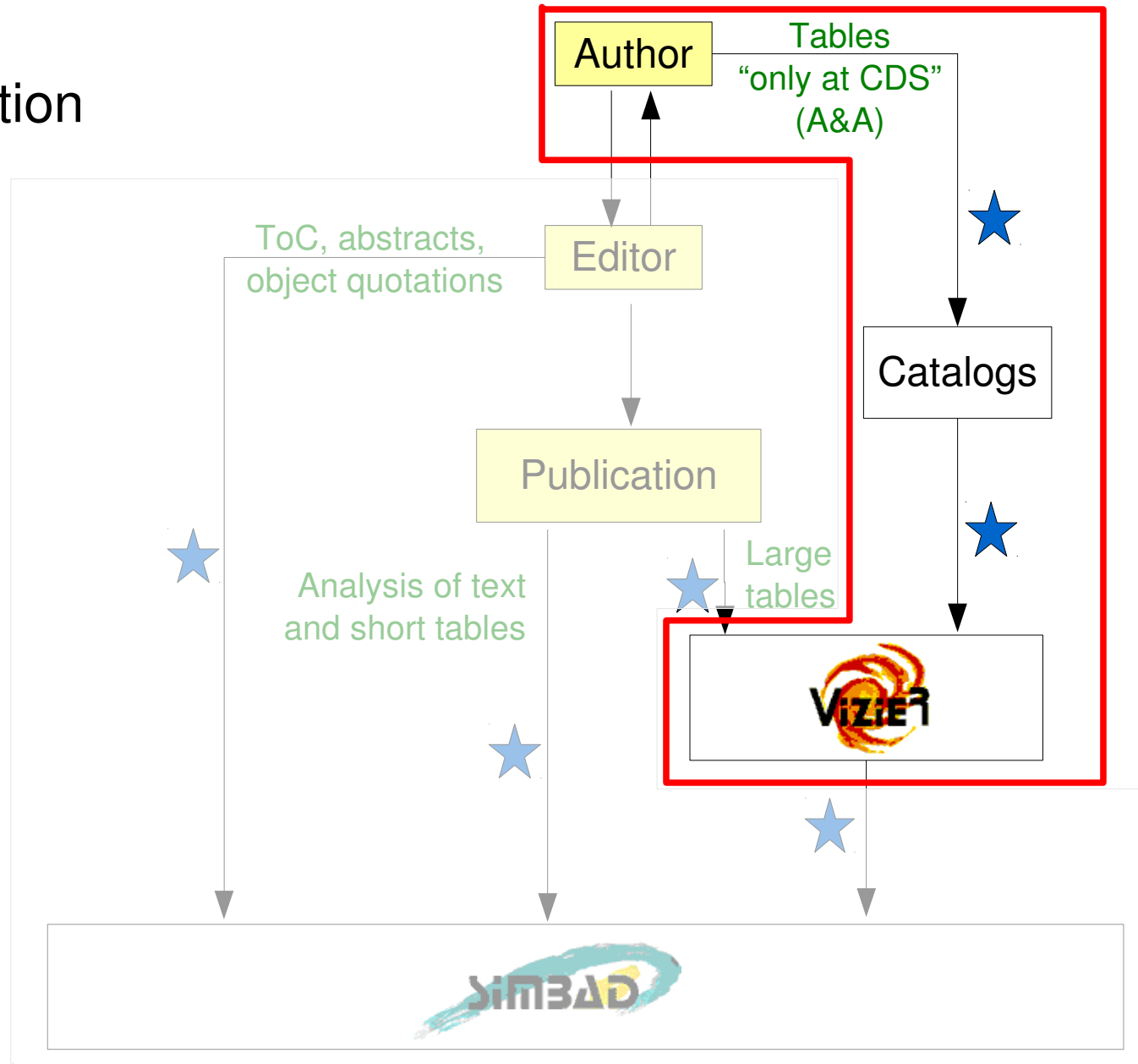


# Life cycle of an article



# Electronic tables (1)

- ★ : information scientist's expertise



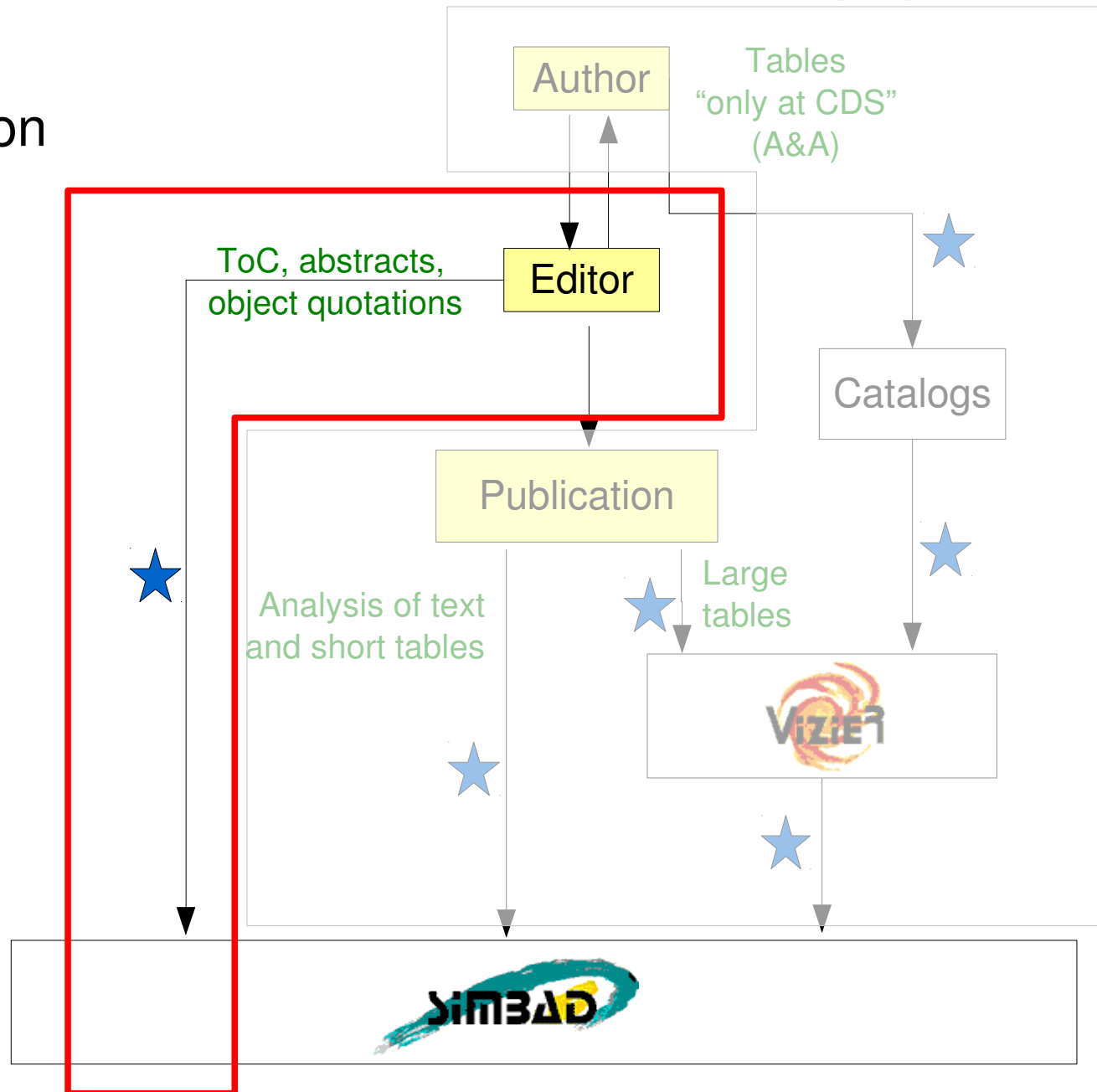


# Electronic tables (2)

- A&A: part of the publication process
  - Scientific editor
    - tables "only at CDS"
    - at editor's request:  
tables+description sent directly to the CDS by authors
  - CDS
    - descriptions homogenized and standardized
    - check (data consistency, may imply interactions with the authors)
    - become public upon publication
    - accessible via VizieR

# Tables of contents (1)

- ★ : information scientist's expertise





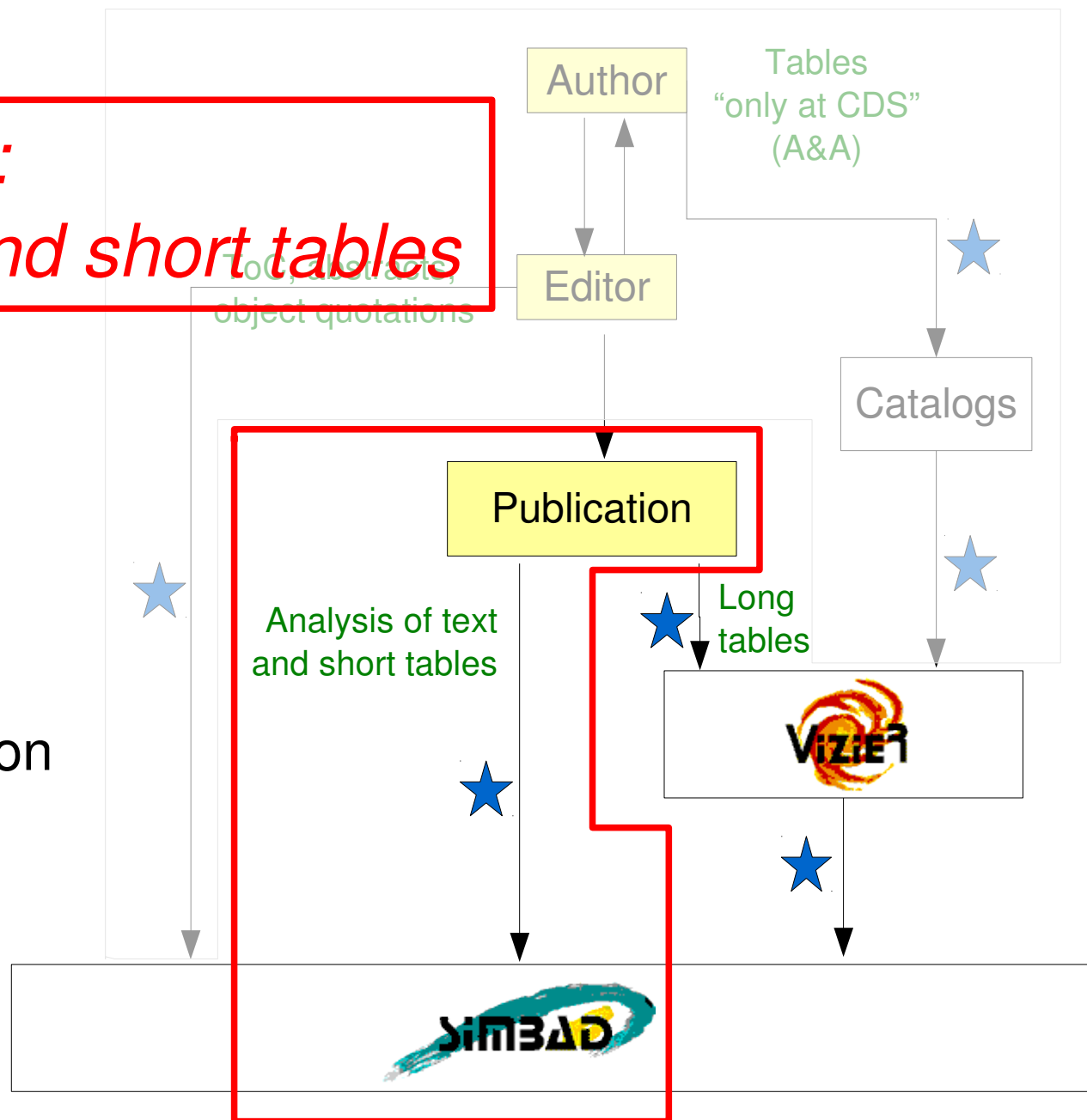


# Tables of contents (2)

- Publisher
  - Tables of Contents + Abstracts + Keywords
  - Object Quotations (A&A), tagged by the authors
- CDS
  - Check/control/validation
  - SIMBAD update

# How data are entered into SIMBAD (1)

*Two paths:  
1- Text and short tables*



- ★ : information scientist's expertise

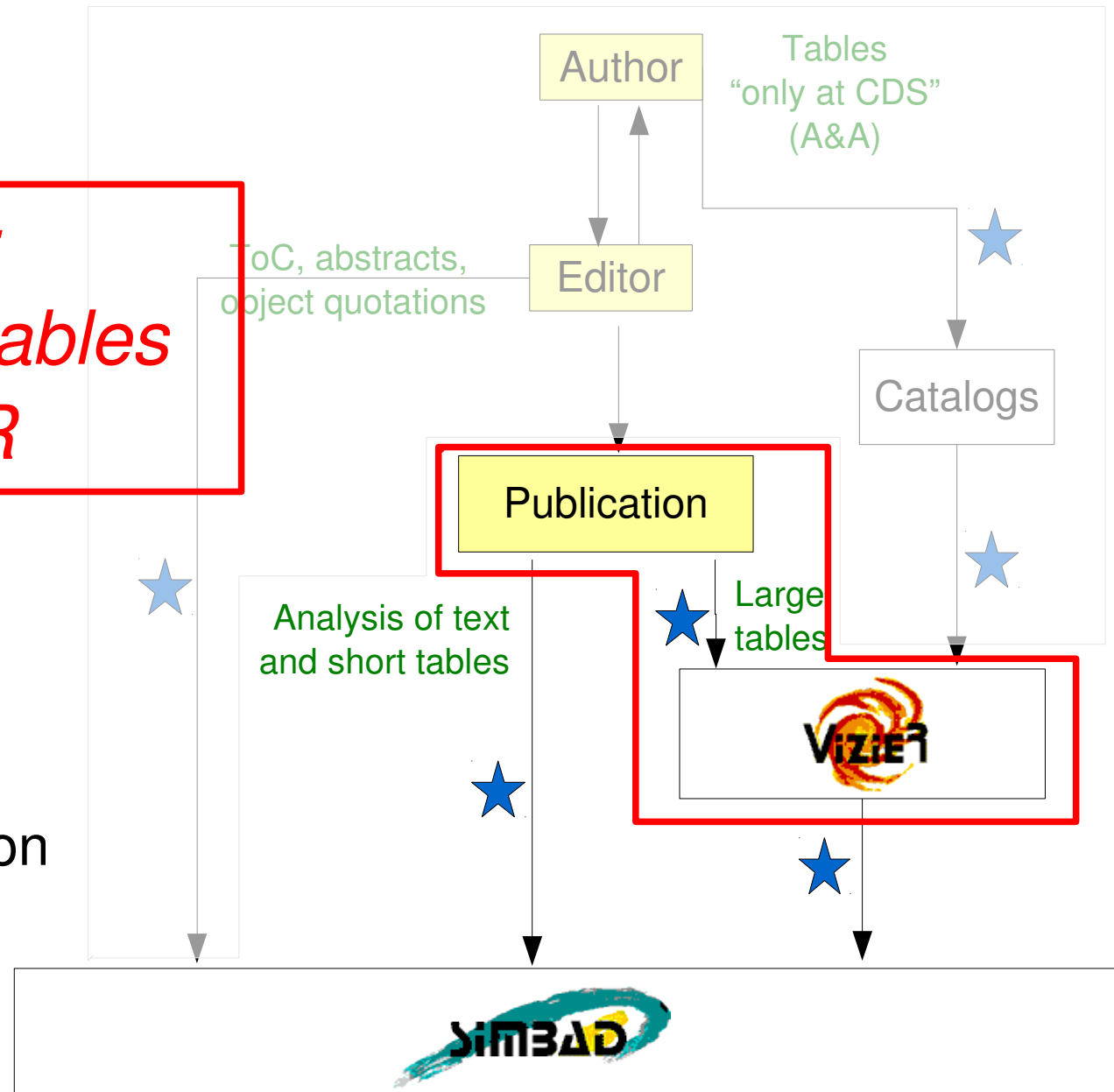


# Detailed article analysis

- Objects in text and short tables
  - Identification of astronomical objects and related data
    - by reading the paper
      - with the help of a dedicated in-house software (DJIN, see S. Lesteven presentation) since 2008 and only for the main journals
  - SIMBAD update
    - cross identification: does the object already exist in SIMBAD (possibly under another name)?

# How data are entered into SIMBAD (2)

*Two paths:  
2. Large tables  
2.1 VizieR*





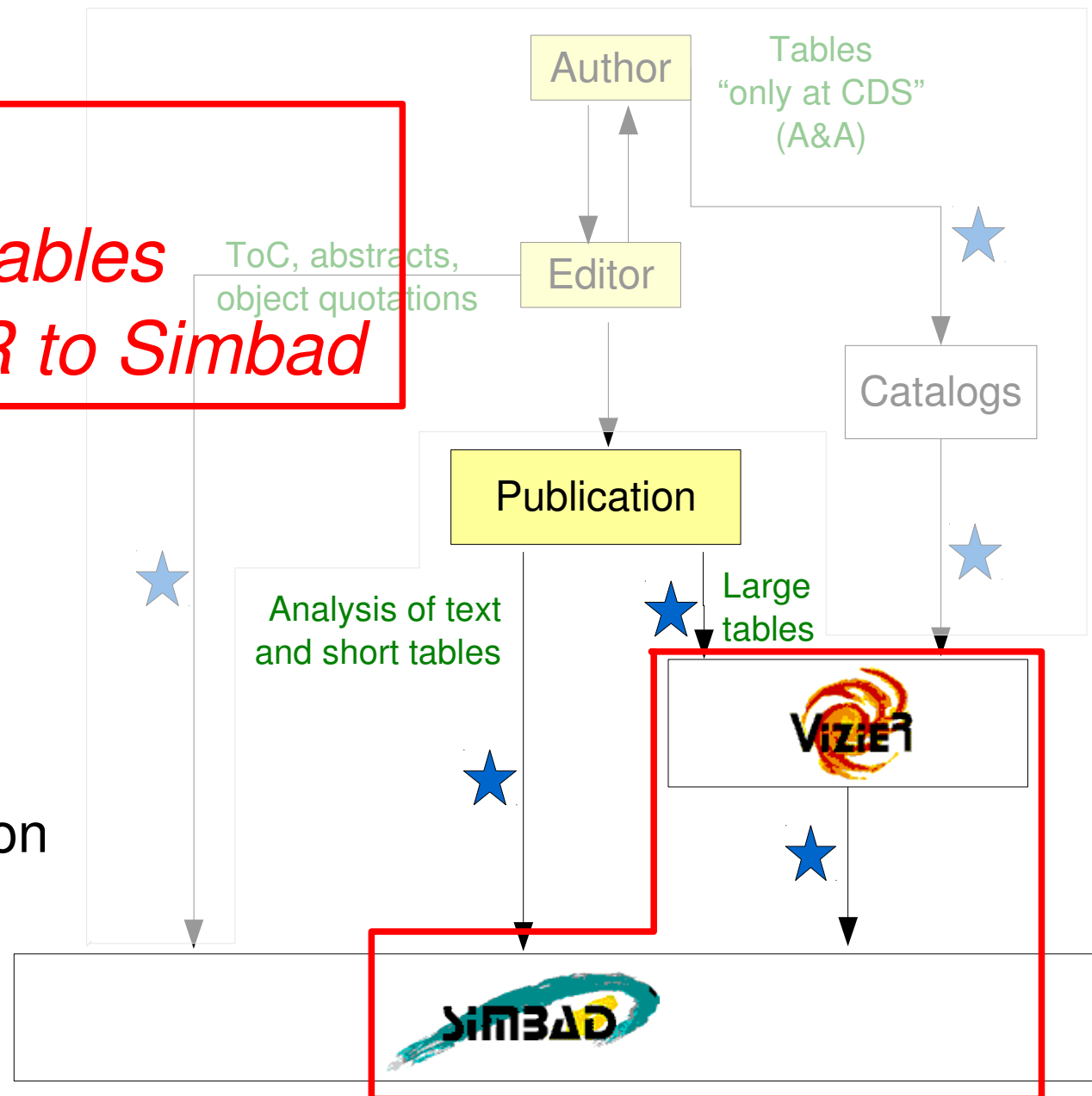
# Large tables

## First into VizieR

- Analysis of large tables
  - Descriptions homogenized and standardized
    - if no description file is available, one is created
  - Checks for data consistency
  - VizieR update

# How data are entered into SIMBAD (3)

*Two paths:  
2. Large tables  
2.2 VizieR to Simbad*



- ★ : information scientist's expertise

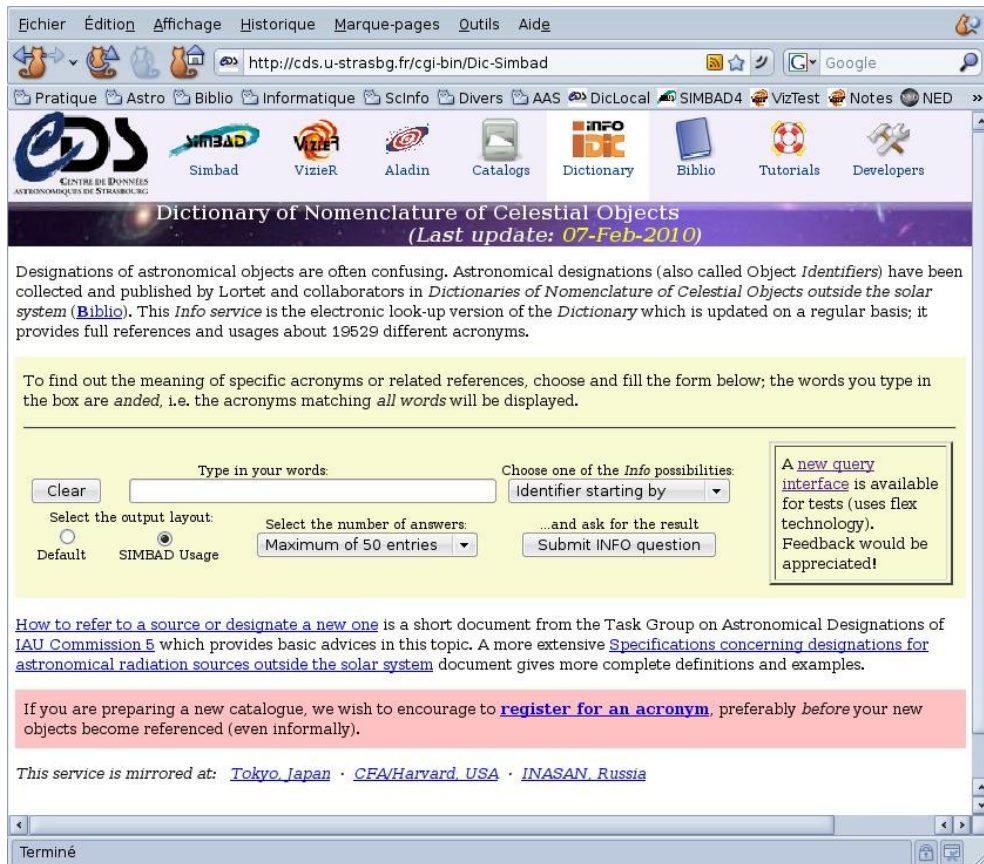


# Large tables

## From VizieR to Simbad

- Table processing
  - Identification of astronomical objects and related data
    - in-house tool “Raccord” (“*junction*”/”*join*”)
    - cross match with SIMBAD
  - SIMBAD update
  - See F. Woelfel et al. (2006, LISA V, Boston)

# Our cornerstone: The *Dictionary of Nomenclature of Celestial Objects*



Fichier Édition Affichage Historique Marque-pages Outils Aide

http://cds.u-strasbg.fr/cgi-bin/Dic-Simbad

Pratique Astro Biblio Informatique Scinfo Divers AAS DicLocal SIMBAD4 VizTest Notes NED

Centre de Données Astronomiques de Strasbourg

Dictionary

Dictionary of Nomenclature of Celestial Objects  
(Last update: 07-Feb-2010)

Designations of astronomical objects are often confusing. Astronomical designations (also called Object *Identifiers*) have been collected and published by Lortet and collaborators in *Dictionaries of Nomenclature of Celestial Objects outside the solar system* ([Biblio](#)). This *Info* service is the electronic look-up version of the *Dictionary* which is updated on a regular basis; it provides full references and usages about 19529 different acronyms.

To find out the meaning of specific acronyms or related references, choose and fill the form below; the words you type in the box are *anded*, i.e. the acronyms matching *all* words will be displayed.

Clear

Type in your words:

Choose one of the *Info* possibilities:  
Identifier starting by

Select the output layout:  
 Default  SIMBAD Usage

Select the number of answers:  
Maximum of 50 entries

...and ask for the result  
Submit INFO question

A new query interface is available for tests (uses flex technology). Feedback would be appreciated!

[How to refer to a source or designate a new one](#) is a short document from the Task Group on Astronomical Designations of [IAU Commission 5](#) which provides basic advices in this topic. A more extensive [Specifications concerning designations for astronomical radiation sources outside the solar system](#) document gives more complete definitions and examples.

If you are preparing a new catalogue, we wish to encourage to [register for an acronym](#), preferably *before* your new objects become referenced (even informally).

This service is mirrored at: [Tokyo Japan](#) · [CFA/Harvard USA](#) · [INASAN, Russia](#)

Terminé

- Confusing designations
- 19 000 acronyms
- IAU Commission V: register an acronym
- Essential input to DJIN
- ... It keeps me busy!

<http://cds.u-strasbg.fr/cgi-bin/Dic-Simbad>



# What special skills do we need? (1)

- Knowledge of astronomy
  - to identify astronomical object names
  - to be familiar with the astronomical zoo (stars, binaries, cepheids, clusters of stars, planetary nebulae, interacting galaxies, voids...)
  - to recognize the parameters relevant for SIMBAD (redshifts, magnitudes, spectral types, metallicity...)
  - to follow the evolution of the discipline

# What special skills do we need? (2)

- Information technology skills
  - linux environment
  - scripting and programming (in awk, perl)
- Sharing expertise
  - close interactions
  - discussion of complex cases with astronomers

# Conclusion-1

- CDS: a team of complementary skills
  - information scientists
  - astronomers
  - software engineers
- Collaboration with journal editors quite effective, data and bibliography fully linked
- Widely used services with a high added-value
  - ~300,000 queries/day

# Conclusion-2

- The critical elements for success
  - evolution with the scientific knowledge
  - constant interaction with astronomers
  - contents & tools: always the user's needs in mind
  - close collaboration with all others actors (journals, ADS, NED)