

# Work in Progress

*The ADS Team*

ADS Users Group Meeting, 20-21 Nov. 2019



# Services Being Readied for Rollout

- **Reference Resolver**
  - Used to map reference strings extracted from papers into ADS citations
  - New framework uses state of the art techniques to parse and interpret reference strings vs. heuristics
- **Recommendations**
  - Tested different topic modeling techniques (LDA, w2v) and document similarity metrics against (anonymous) readership patterns
  - Preparing to test personalized recommendations for new articles based on each user's prior read history
- **myADS Notification Service**
  - Replaces and extends the personalized notifications created in Classic
  - Currently being enhanced to allow additional flexibility and personalizations

# Reference Resolver Process

## 1. Reference Extraction

Identify the references in the paper and extract individual ones

## 2. Reference Segmentation

Tokenize the string into its components (author, year, publication, page, etc)

## 3. Reference Resolution

Attempt to match the extracted metadata fields against an existing ADS record, and if found, assign a citation to it

*Note: 2 & 3 usually run iteratively*

No. 1, 2002

DECOMPOSITION OF GALAXY IMAGES

293

- Ho, L. C. 1999, ApJ, 516, 672  
Ho, L. C., Filippenko, A. V., & Sargent, W. L. W. 1997a, ApJ, 487, 568  
Ho, L. C., Filippenko, A. V., Sargent, W. L. W., & Peng, C. Y. 1997b, ApJS, 112, 391  
Ho, L. C., & Peng, C. Y. 2001, ApJ, 555, 650  
Ho, L. C., et al. 2000, ApJ, 541, 120  
Ho, L. C., et al. 2002, in preparation  
Holtzman, J. A., Burrows, C. J., Casertano, S., Hester, J. J., Trauger, J. T., Watson, A. M., Worthey, G. 1995, PASP, 107, 1065  
Jaffe, W., Ford, H. C., O'Connell, R. W., van den Bosch, F., & Ferrarese, L. 1994, AJ, 108, 1567  
Jedrzejewski, R. 1987, MNRAS, 226, 747  
Kent, S. M. 1985, ApJS, 59, 115  
Kent, S. M., Dame, T. M., & Fazio, G. 1991, ApJ, 378, 131  
Khosroshahi, H., Wadadekar, Y., & Kembhavi, A. 2000, ApJ, 533, 162  
Kormendy, J. 1977, ApJ, 217, 406  
Kormendy, J., Bender, R., Evans, A. S., & Richstone, D. 1998, AJ, 115, 1823  
Kormendy, J., & Bruzual A., G. 1978, ApJ, 223, L63  
Kotilainen, J. K., Ward, M. J., Boisson, C., Depoy, D. L., & Smith, M. G. 1997, MNRAS, 266, 195  
Prieto, M., Aguerri, J. A. L., Varela, A. M., & Muñoz-Tuñón, C. 2001, A&A, 367, 405  
Quillen, A. C., McDonald, C., Alonso-Herrero, A., Lee, A., Shaked, S., Rieke, M. J., & Rieke, G. H. 2001, ApJ, 547, 129  
Ravindranath, S., Ho, L. C., & Filippenko, A. V. 2002, ApJ, 566, 801  
Ravindranath, S., Ho, L. C., Peng, C. Y., Filippenko, A. V., & Sargent, W. L. W. 2001, AJ, 122, 653  
Rest, A., van den Bosch, F. C., Jaffe, W., Tran, H., Tsvetanov, Z., Ford, H. C., Davies, J., & Schafer, J. 2001, AJ, 121, 2431  
Riethorst, D., Fabbiano, G., & Garmann, J. 1997, MNRAS, 287, 1110  
Ryder, S. D., Purcell, G., Anderson, V., & Davis, D. 2000, in ASP Conf. Ser. 197, From the Early Universe to the Present, ed. F. Combes, G. Mamon, & V. Charmandaris (San Francisco: ASP), 405  
Sara, M., Rix, H.-W., Shields, J. C., Rudnick, G., Ho, L. C., McIntosh, D. H., Filippenko, A. V., & Sargent, W. L. W. 2001, ApJ, 550, 65  
Scorza, C. & Bender, R. 1995, A&A, 293, 20  
Sérsic, J. L. 1968, Atlas de Galaxias Australes (Córdoba: Obs. Astron., Univ. Nac. Córdoba)  
Shaw, M. A., & Gilmore, G. F. 1989, MNRAS, 237, 903  
Simon, F. & McQuinn, R. 1990, A&A, 237, 11

Ryder, S. D., Purcell, G., Anderson, V., & Davis, D. 2000, in ASP Conf. Ser. 197, From the Early Universe to the Present, ed. F. Combes, G. Mamon, & V. Charmandaris (San Francisco: ASP), 405

VIEW

Abstract

Citations (1)

References (5)

Co-Reads

Similar Papers

Volume Content

Graphics

Metrics

Export Citation

### The Interaction of NGC 7421 with the Intracluster Medium

Show affiliations

Ryder, S. D.; Purcell, G.; Andersen, V.; Davis, D.

We examine various mechanisms that may account for the asymmetry of NGC 7421. A combination of intrinsic disk-halo misalignment, tidal interaction, and ram-pressure may be needed to fully explain the observed lopsidedness, and its longevity.

**Publication:** Dynamics of Galaxies: from the Early Universe to the Present, 15th IAP meeting held in Paris, France, July 9-13, 1999, Eds.: Françoise Combes, Gary A. Mamon, and Vassilis Charmandaris ASP Conference Series, Vol. 197, ISBN: 1-56381-024-2, 2000, p. 405.

**Pub Date:** 2000

**Bibcode:** 2000ASPC..197..405R

FULL TEXT SOURCES

My Institution  
ADS

Add paper to a library

# ADS Classic Reference Resolution: transformations + heuristics + bibcodes

```
ADS-References-Document-Parser -- emacs -- 88x81

sub parse {
  sv $refl = shift @_;
  sv $arg = shift @_;
  sv $BIBL = "";

  if (ref($arg) ) {
    read($arg,$BIBL,$MAX_SIZE) or
    carp "Error reading from file handle: $!";
    return ($);
  } else {
    $BIBL = join("", $arg @);
  }

  unless (length($BIBL)) {
    carp "No data to parse";
    return ($);
  }

  sv $pg = ($BIBL =~ s/(\r?)/g);
  sv $pages = scalar($pg);
  warn "Input document seems to have $pages pages";
  # require a minimum number of matches for the headers computing (1
  # based on the number of pages
  sv $min_matches = int(($pages - 1) / 2);
  $min_matches = 3 if ($min_matches < 3);

  sv $HEADERS;
  # headers
  while ($BIBL =~ /(\r?\n){0,50}(\r?\n){0,50}/g) {
    $HEADERS($header_to_regex($1))++;
  }
  # footers
  while ($BIBL =~ /(\r?\n){0,50}(\r?\n){0,50}/g) {
    $HEADERS($header_to_regex($1))++;
  }

  # XXX TODO: we should check that numbers in regexp end up watching
  # increasing page numbers (usually either by 1 or 2 counts)

  if ($HEADERS) {
    sv $REGEXP = sort { $HEADERS($b) <=> $HEADERS($a) } keys $HEADERS;
    warn "Regexp are:\n", (map { "\$$_count $HEADERS{$1}" }
      keys $HEADERS) if $DEBUG;

    warn "Not enough watching header/footers were found";
    if ($BIBL and $HEADERS{$count($BIBL)} = $min_matches) {
      foreach sv $regexp ($REGEXP) {
        if ($BIBL =~ $regexp) { $min_matches++; }
        sv $c = $BIBL =~ s/(\r?\n){0,50}(\r?\n){0,50}/g;
        warn "Applied regexp: $regexp ($HEADERS{$regexp} matches)";
        "Removed $c headers using $regexp" if $DEBUG;
      } else {
        last;
      }
    }
  } else {
    warn "No headers were found" if $DEBUG;
  }
  $BIBL =~ s/(\r?\n){0,50}(\r?\n){0,50}/g;

  sv $REFS;
  # attempt to find the reference section
  while ($REFS and ($BIBL =~ s/(\r?\n){0,50}(\r?\n){0,50}/g)) {
    warn "Ref. start marker: $!";
    $REFS = &trace_ref($BIBL);
  }

  # If we couldn't find a marker for the references, try to detect
  # at least well-formatted sections from the file
  while ($REFS and ($BIBL =~ s/(\r?\n){0,50}(\r?\n){0,50}/g)) {
    $REFS = &trace_ref($BIBL, style => "bracket_number");
  }

  sv $c = 1;
  return map { $c ? "[ $c" : " ] " } $c : () : $REFS;
}

sub { ... } Astro.pm 226 L137 (Perl)-----
```

Ryder, S. D., Purcell, G., Anderson, V., & Davis, D. 2000, in ASP Conf. Ser. 197, From the Early Universe to the Present, ed. F. Combes, G. Mamon, & V. Charmandaris (San Francisco: ASP), 405

Ryder, S. D., Purcell, G., Anderson, V., & Davis, D. 2000, in ASP Conf. Ser. 197, From the Early Universe to the Present, ed. F. Combes, G. Mamon, & V. Charmandaris (San Francisco: ASP), 405

2000ASPC..197..405R

# New ADS Resolver:

## Machine learning and structured metadata

### 1. Reference Extraction

Will use a mix of open source tools depending on input formats (PDF, LaTeX, XML)

### 2. Reference Segmentation

Relies on Conditional Random Fields models trained on input documents

### 3. Reference Resolution

Matches extracted metadata fields against ADS API via probabilistic model to identify citation

Ryder, S. D., Purcell, G., Anderson, V., & Davis, D. 2000, in ASP Conf. Ser. 197, From the Early Universe to the Present, ed. F. Combes, G. Mamon, & V. Charmandaris (San Francisco: ASP), 405

Ryder, S. D., Purcell, G., Anderson, V., & Davis, D. 2000, in ASP Conf. Ser. 197, From the Early Universe to the Present, ed. F. Combes, G. Mamon, & V. Charmandaris (San Francisco: ASP), 405

```
"bibcode": "2000ASPC..197..405R",  
"author": ["Ryder, S. D.",  
            "Purcell, G.",  
            "Andersen, V.",  
            "Davis, D."],  
"year": "2000",  
"series": "Astronomical Society of the Pacific  
Conference Series",  
"volume": "197",  
"pub": "Dynamics of Galaxies: from the Early  
Universe to the Present",  
"editor": ["Combes, Francoise",  
            "Mamon, Gary A.",  
            "Charmandaris, Vassilis"],  
"page": ["405"]
```

*Steps 2+3 currently being validated*

# Personalized Recommendations

Lots of ways to do this, so start simple and test against users:

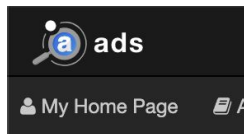
*“ADS Oracle, given what I have read recently, what new things should I read today?”*

You may want to read articles which...

- are **similar** to the ones you read (new)
- have been read by **other people with similar interests** (trending)
- **cite** papers that you have recently read (reviews)
- **are cited** by papers that you have recently read (useful)

<https://qa.adsabs.harvard.edu/static/tmp/oracle/>

# myADS Not



[Customize Settings](#)

## USER PREFERENCES

[Import Libraries](#)  
[Library Link Server](#)  
[ORCID Settings](#)  
[Application](#)

## NOTIFICATION SETTINGS

[MyADS](#)

## ACCOUNT SETTINGS

[Change Email](#)  
[Change Password](#)  
[API Token](#)  
[Delete Account](#)

## arXiv Notification Template

Set of Keywords \*

star or planet

Boolean "AND" is assumed, but can be overridden by using explicit logical operators between keywords

arXiv Groups \* (must

☒ astro-ph: As

☒ astro-pl

☒ astro-ph

☐ astro-ph

☒ astro-pl

☐ astro-ph

☒ astro-pl

☒ cond-mat: C

☐ gr-qc: Gener

☐ math-ph: Ma

☐ nucl-ex: Nucl

☐ nucl-th: Nucl

☐ econ: Econo

☐ eess: Electric

☐ physics: Phyt

☐ quant-ph: Qu

☐ math: Mather

☐ nlin: Nonlinea

☐ cs: Computer

☐ q-bio: Quanti

☐ q-fin: Quantit

☐ stat: Statistic

☐ Select All

Select the groups to c

Create Notification

### Get Notifications for this Query?

Name

Stuff I want

☒ Only latest results

Frequency

Daily

Weekly

Create Notification

### Get Notifications for this Query?

Name

Stuff I want

☒ Only latest results

Frequency

Daily

Weekly

The notification **Stuff I want** created successfully!

[? About](#) [Account](#)

# myADS emails



## myADS Personal Notification Service

Kelly Lockhart

Wed Nov 20 07:04:34 2019

daily arXiv e-prints database



Cornell University  
arXiv

### myADS Notifications

[View in browser](#)

[Update myADS Settings](#)

[myADS Help](#)

### ADS Query Forms

[The New ADS](#)

[Astronomy](#)

[Physics](#)

[arXiv e-prints](#)

[FAQ](#)

### Current Tables of Contents

[Recent astro-ph](#)

### Search Recent Papers

Author:   
Subject:

### \*. NUCLEAR STAR CLUSTER - Recent Papers

- [2019arXiv191108475R](#) Raghunathan,+: The Gravitational Lensing Signatures of BOSS Voids in the Cosmic Microwave Background
- [2019arXiv191108473F](#) Fonseca,+: Relaxion Fluctuations (Self-stopping Relaxion) and Overview of Relaxion Stopping Mechanisms
- [2019arXiv191108472F](#) Fonseca,+: Axion Fragmentation
- [2019arXiv191108471A](#) Arico,+: Modelling the large scale structure of the Universe as a function of cosmology and baryonic physics
- [2019arXiv191108462A](#) Antonucci,+: Metis: the Solar Orbiter visible light and ultraviolet coronal imager
- [2019arXiv191108450L](#) Lacchin,+: Chemical evolution of ultra-faint dwarf galaxies: testing the IGIMF
- [2019arXiv191108436K](#) Keys,+: High-resolution spectropolarimetric observations of the temporal evolution of magnetic fields in photospheric bright points
- [2019arXiv191108431Q](#) Quarles,+: Obliquity Evolution of Circumstellar Planets in Sun-like Stellar Binaries
- [2019arXiv191108407O](#) Ofengeim,+: Bulk viscosity in neutron stars with hyperon cores
- [2019arXiv191108391W](#) Wehbe,+: The impact of atmospheric dispersion in the performance of high-resolution spectrographs
- [2019arXiv191108390C](#) Cohen,+: Bayesian Characterization of Main Sequence Binaries in the Old Open Cluster NGC 188
- [2019arXiv191108387C](#) Chrimes,+: Binary population synthesis models for core-collapse gamma-ray burst progenitors



astrophysics  
data system

## myADS Personal Notification Service

Daily email

November 20, 2019

### "nuclear star cluster" - Recent Papers

- [2019arXiv191108475R](#) Raghunathan, S,+: The Gravitational Lensing Signatures of BOSS Voids in the Cosmic Microwave Background
- [2019arXiv191108473F](#) Fonseca, N,+: Relaxion Fluctuations (Self-stopping Relaxion) and Overview of Relaxion Stopping Mechanisms
- [2019arXiv191108472F](#) Fonseca, N,+: Axion Fragmentation
- [2019arXiv191108471A](#) Arico, G,+: Modelling the large scale structure of the Universe as a function of cosmology and baryonic physics
- [2019arXiv191108462A](#) Antonucci, E,+: Metis: the Solar Orbiter visible light and ultraviolet coronal imager
- [2019arXiv191108450L](#) Lacchin, E,+: Chemical evolution of ultra-faint dwarf galaxies: testing the IGIMF
- [2019arXiv191108436K](#) Keys, P,+: High-resolution spectropolarimetric observations of the temporal evolution of magnetic fields in photospheric bright points
- [2019arXiv191108431Q](#) Quarles, B,+: Obliquity Evolution of Circumstellar Planets in Sun-like Stellar Binaries
- [2019arXiv191108407O](#) Ofengeim, D,+: Bulk viscosity in neutron stars with hyperon cores
- [2019arXiv191108391W](#) Wehbe, B,+: The impact of atmospheric dispersion in the performance of high-resolution spectrographs