

# The MUIR Framework: Cross-Linking MOOC Resources to Enhance Discussion Forums

Ya-Hui An, Muthu Kumar Chandrasekaran,  
Min-Yen Kan, Yan Fu

[anyahui.120@gmail.com](mailto:anyahui.120@gmail.com)



电子科技大学  
University of Electronic Science and Technology of China

coursera

edX



UDACITY







VERIFIED



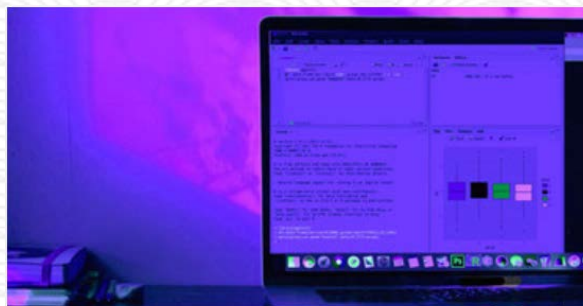
GTx  
Machine Learning



VERIFIED



ColumbiaX  
Machine Learning



VERIFIED



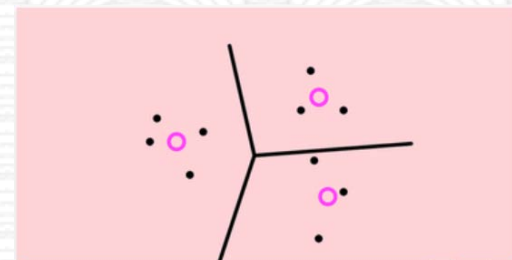
HarvardX  
Data Science: Machine Learning



VERIFIED



Microsoft  
Principles of Machine Learning: Python Edition

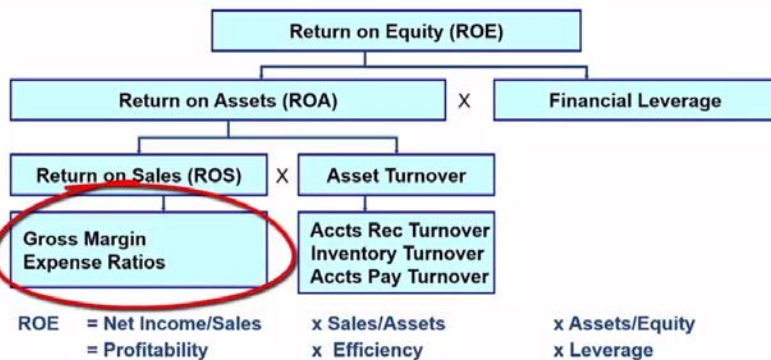


VERIFIED



UCSanDiegoX  
Machine Learning Fundamentals

## DuPont Ratio Analysis Framework



0:48 / 18:28

Ratio Analysis: Profitability and Turnover Ratios 1.5

## Week 1

Discuss and ask questions about Week 1.

SORT BY: Latest

Search



New Thread

### ROA interpretation

Mentor Replied

Last post by Jose DeLaCerdá · 2 months ago

110 views  
2 replies

### Accounts payable turnover

Last post by Jesus. M Cumplido Barreiro · 3 months ago

11 views  
1 replies

### Is there an answer key for us to check what we got wrong?

Staff Replied

Last post by Jenny Hopkins · 3 months ago

31 views  
1 replies

### Formula question on Chg. in Other Liabilities in CF statement

Created by 吳佳錄 · 4 months ago

10 views  
0 replies

### Question 5 - Week1 quiz

Instructor Replied

Last post by Keith Higgins · 5 months ago

291 views  
15 replies

### Balance Sheet

Created by Danyal Rashid · 5 months ago

13 views  
0 replies

### quiz results

Staff Replied

Last post by Muhammad Risyan Rizqi Khairi · 5 months ago

219 views  
7 replies

### DuPont Analysis spreadsheet

Last post by kemakolam kingsley · 6 months ago

71 views  
2 replies

### Cash flow from Financing activities

Created by Ankit Ainapure · 8 months ago

34 views  
0 replies

### Illustration of "saw-tooth" pattern when using activity ratios to forecast balance sheet items

Panel A					
Year	Sales	TATR	Avg Total Assets	Ending Total Assets	
0	100			40	
1	100	2.0	50	60	
2	100	2.0	50	40	
3	100	2.0	50	60	
4	100	2.0	50	40	
5	100	2.0	50	60	
Panel B					
Year	Sales	TATR	Avg Total Assets	Ending Total Assets	
0	100			40	
1	120	2.0	60	80	
2	144	2.0	72	64	
3	172	2.0	86	100	

### Ratios and Forecasting

Quiz: Ratio Analysis and Forecasting Quiz 10 questions

Jul 16

### Earnings Management

Quiz: Earnings Management 10 questions

Jul 23

### Big Data and Prediction Models

Quiz: Big Data and Prediction Models 10 questions

Jul 30

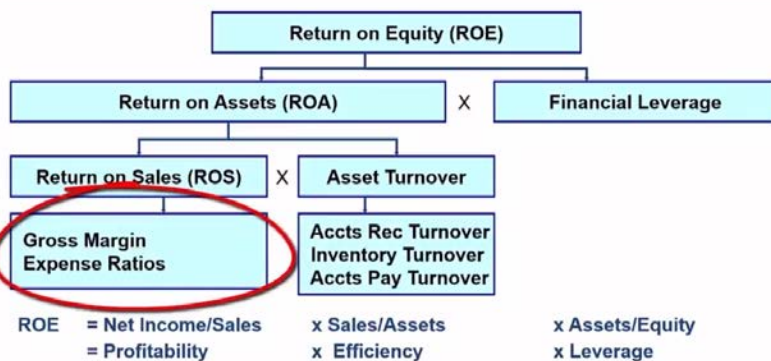
### Linking Non-financial Metrics to Financial Performance

Quiz: Linking Non-financial Metrics to Financial Performance 10 questions

Aug 6



## DuPont Ratio Analysis Framework



MUIR

Ratio Analysis: Profitability and Turnover Ratios 1.5

## Week 1

Discuss and ask questions about Week 1.

SORT BY: Latest

Search



New Thread

### ROA interpretation

Mentor Replied

Last post by Jose DeLaCerdá · 2 months ago

110 views  
2 replies

### Accounts payable turnover

Last post by Jesus. M Cumplido Barreiro · 3 months ago

11 views  
1 replies

### Is there an answer key for us to check what we got wrong?

Staff Replied

Last post by Jenny Hopkins · 3 months ago

31 views  
1 replies

### Formula question on Chg. in Other Liabilities in CF statement

Created by 吳佳錄 · 4 months ago

10 views  
0 replies

### Question 5 - Week1 quiz

Instructor Replied

Last post by Keith Higgins · 5 months ago

291 views  
15 replies

### Balance Sheet

Created by Danyal Rashid · 5 months ago

13 views  
0 replies

### quiz results

Staff Replied

Last post by Muhammad Risyan Rizqi Khairi · 5 months ago

219 views  
7 replies

### DuPont Analysis spreadsheet

Last post by kemakolam kingsley · 6 months ago

71 views  
2 replies

### Cash flow from Financing activities

Created by Ankit Ainapure · 8 months ago

34 views  
0 replies

Ratios and Forecasting

Quiz: Ratio Analysis and

Earnings Management

Quiz: Earnings Management 10 questions

Jul 23

Big Data and Prediction Models

Quiz: Big Data and Prediction Models 10 questions

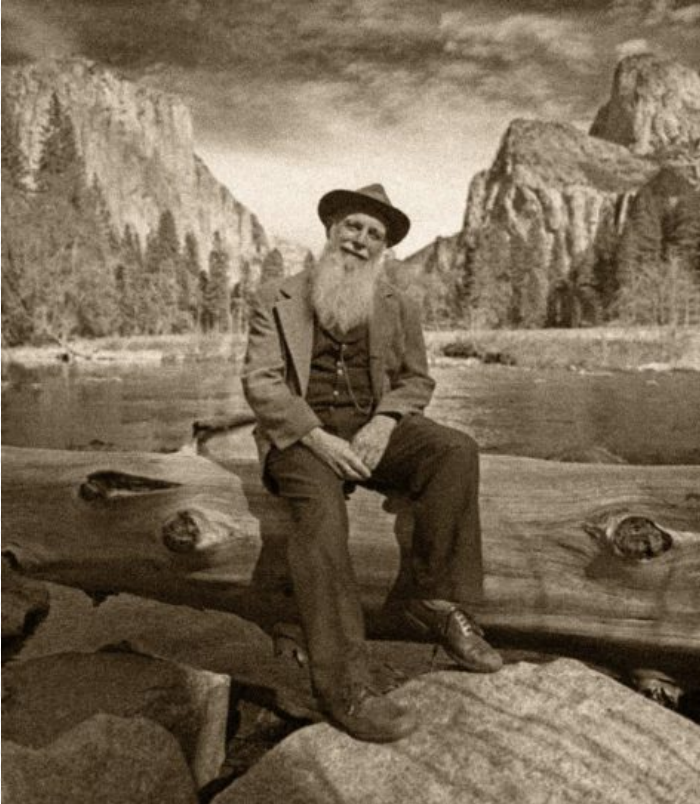
Jul 30

Linking Non-financial Metrics to Financial Performance

Quiz: Linking Non-financial Metrics to Financial Performance 10 questions

Aug 6

1	Illustration of "saw-tooth" pattern when using activity ratios to forecast balance sheet items				
2					
3					
4					
5					
6	Panel A				
7	Year	Sales	TATR	Avg Total Assets	Ending Total Assets
8	0	100			40
9	1	100	2.0	50	60
10	2	100	2.0	50	40
11	3	100	2.0	50	60
12	4	100	2.0	50	40
13	5	100	2.0	50	60
14					
15					
16	Panel B				
17	Year	Sales	TATR	Avg Total Assets	Ending Total Assets
18	0	100			40
19	1	120	2.0	60	80
20	2	144	2.0	72	64
21	3	172	2.0	88	100



*“When we try to pick out anything by itself, we find it hitched to everything else in the universe”*

*— John Muir*



# Wikification

## Discussion Forums

- Course Home

- Grades

- Discussion

Forums

- Messages

- Course Info

A

Asker Account • 2 days ago

### Question 9 of assignment 2

Below is the Accounts Receivable line from... How would you approach this problem?

B

Helper Account • 1 days ago

Hi, it may help you to review the spreadsheet for lecture 2.5, tab labeled "AR Allowance."

# Wikification

## Discussion Forums

A Asker Account • 2 days ago

### Question 9 of assignment 2

Below is the Accounts Receivable line from... How would you approach this problem?

B Helper Account • 1 days ago

Hi, it may help you to review the spreadsheet for lecture 2.5, tab labeled "AR Allowance."

Q9: In the problems below use matrix algebra on a computer to find  $a$ ,  $b$  and  $c$  uniquely if possible. If not possible explain why not. You are given that

$$a\vec{A} + b\vec{B} + c\vec{C} = \vec{D}$$





# Outline

Introduction

**Methods**

Evaluation

- Search Accuracy
- Mention Extraction Coverage

Discussion

Conclusion

# MUIR Workflow

MUIR:

**MOOOC Uniform Identifier for Resources**

B

Helper Account • 1 days ago

Hi, it may help you to review the spreadsheet for [lecture 2.5](#), tab labeled “AR Allowance.”



**Short Form** <http://www.muir.xref/coursera/accounting-analytics>  
**MUIR** [/videos/2-5](http://www.muir.xref/coursera/accounting-analytics/videos/2-5)



**Mention** [http://www.muir.xref/coursera/accounting-analytics/](http://www.muir.xref/coursera/accounting-analytics/I480320000/Brain&Ittner/videos/2-5)  
**MUIR** [I480320000/Brain&Ittner/videos/2-5](http://www.muir.xref/coursera/accounting-analytics/I480320000/Brain&Ittner/videos/2-5)



# MUIR Workflow

MUIR:  
**MOOOC Uniform Identifier for Resources**

Mention  
MUIR

[http://www.muir.xref/coursera/accounting-analytics/  
I480320000/Brain&Ittner/videos/2-5](http://www.muir.xref/coursera/accounting-analytics/I480320000/Brain&Ittner/videos/2-5)



Resource  
MUIR

[http://www.muir.xref/coursera/accounting-analytics/  
I480320000/Brain&Ittner/videos/expense-recognition-2-5](http://www.muir.xref/coursera/accounting-analytics/I480320000/Brain&Ittner/videos/expense-recognition-2-5)



Actual  
URL

[http://www.coursera.org/learn/accounting-analytics/  
lecture/IUzkX/expense-recognition-2-5](http://www.coursera.org/learn/accounting-analytics/lecture/IUzkX/expense-recognition-2-5)

# Resource Mention

- A **MOOC resource** is a teaching material associated with a course, i.e., videos, slides, assessments, exams, transcripts, readings, and additional resources.
- A **resource mention** is a semantically complete single/multi-word phrase that represents a MOOC resource.
- **Task:** Given a post, our system automatically extracts the resource mentions and category them into one of the resource types.



# Step 1: Mention Extraction

**YES:** <m1> Is it just me or were some questions on Quiz 2 a surprise? There were a few questions that ...

**YES:** <m2> Hello, I just would like to note that on 12:30 in the answer to question 3 in the lecture 2.5, ...

**NO:** <m3> At around 5:00 in the lecture, we see that the regularization term in the cost function ...

**NO:** <m4> Hello. I want to use 'e' as a number for ex.2/week3. It didn't work, and I didn't find useful help with "help exponent".

Single, Concrete and With-in course mentions  
(SCI mentions)

# Step II: MUIR

*Institution*      *Source Platform*      *Course Name*

[http://www.muir.xref/coursera/accounting-analytics/](http://www.muir.xref/coursera/accounting-analytics/1480320000/Brain&Ittner/videos/expense-recognition-2-5)

[1480320000/Brain&Ittner/videos/expense-recognition-2-5](http://www.muir.xref/coursera/accounting-analytics/1480320000/Brain&Ittner/videos/expense-recognition-2-5)

*Session Date*      *Instructor(s)*      *Resource Type*      *Resource Title*

*Other Elements (optional categories)*



# MUIR Resource Types

- Surveyed resources on 29 global MOOC platforms
- Standardized into 7 types:
  1. Videos
  2. Slides
  3. Exams
  4. Assessments
  5. Transcripts
  6. Readings
  7. Additional Resources

# Outline

Introduction

Methods

**Evaluation**

- Search Accuracy
- Mention Extraction Coverage

Discussion

Conclusion



# Proof of Concept Evaluation

Resources from Coursera

**142 courses**

**102,661 posts**

**11,484 learning resources**

- Mention Extraction: How well can we extract mentions?
- MUIR Search: How well can the system find the correct resource?

# Outline

Introduction

Methods

**Evaluation**

- Search Accuracy
- Mention Extraction Coverage

Discussion

Conclusion



# Search Accuracy – Annotation

## Annotation I: Ground Truth

**Mention**



**Post**

Post Id

Course Name

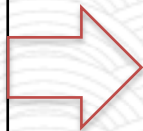
Session

Forum Name

Thread Id

Thread URL

Post Content



The screenshot shows a forum thread with the following content:

- Post 1:** User 'YS' asks a "quiz question about Dalmatian company". The content mentions "without getting too specific, i need a pointer." and "the question is about the updated net income for a given year when its capitalization is abandoned and the cost is immediately expensed. Knowing that the net income given is assuming 40% tax rate, my calculation yields an answer not among the choices (266)?". It has 0 Upvotes and a Reply button.
- Post 2:** User 'Muhammad Risyah Rizqi Khairi' asks "Did you calculate the after tax (net income) after being adjusted or pre-tax income after being adjusted? From what I see, your number is a pre-tax income, not net income." It has 0 Upvotes and a Reply button.
- Post 3:** User 'Peter Strucks' responds: "I was also really struggling with this question. I've managed to pass the exam but I never managed to get this question answered correctly. Any hints?". It has 0 Upvotes and a Reply button.
- Post 4:** User 'Tiji Mathew' starts with "I struggled with this question quite a bit and not finding an answer in the".

To give the realistic evaluation on the full complexity of the problem.

# Search Accuracy - Annotation

## Annotation 2

**Mention**



**Post**

Platform

Course Name

Session

Forum Name

Post Content

Data provided is identical to what the system sees.

To give an upper bound for how well mentions are actually resolved by our simple search logic



# Search Accuracy

Resources	# of instances	Annotation 1	Annotation 2
Videos	89	71.9%	57.3%
Slides	27	74.1%	33.3%
Exams	718	83.0%	53.3%
Assessments	12	50.0%	25.0%
Total	846	81.1%	54.6%

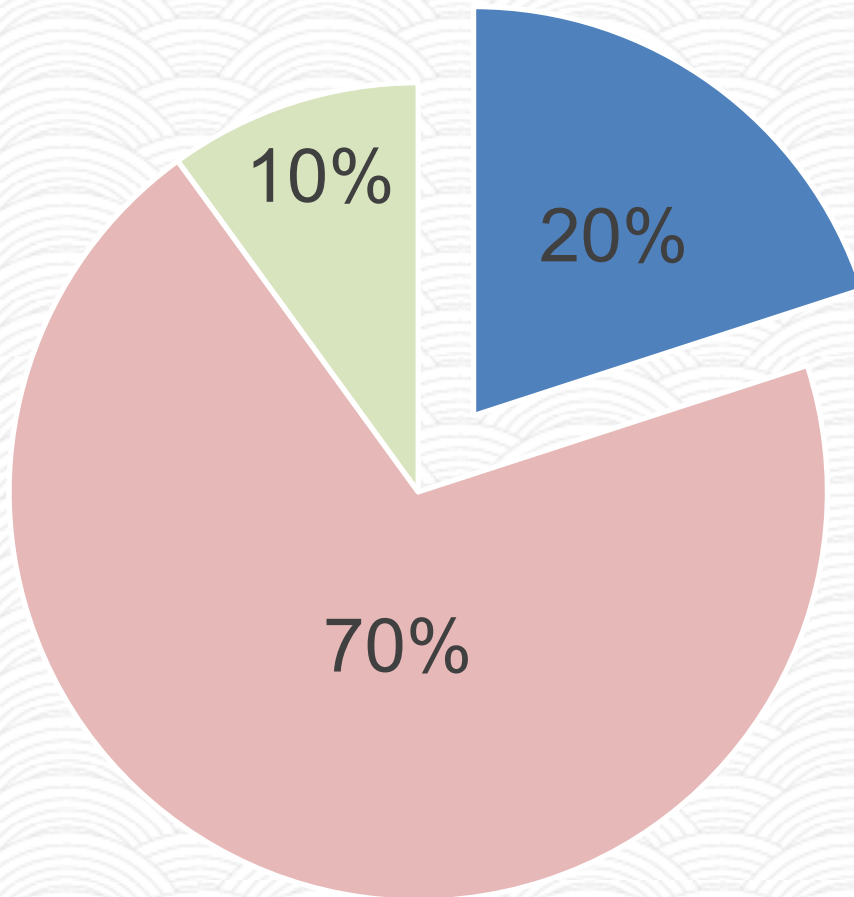
- Column 3 gives the result for upper bound that system can do --- 81.1%;
- Column 4 is accuracy on the ground truth test data --- 54.6%.

# Search Accuracy

- Observations:
  - Mentions easy to extract but hard to resolve;
  - The accuracy for four types of resource have different degrees of reduction;
  - But the results are encouraging: our prototype, even with its simple logic, can already handle almost 55% of learning resources.

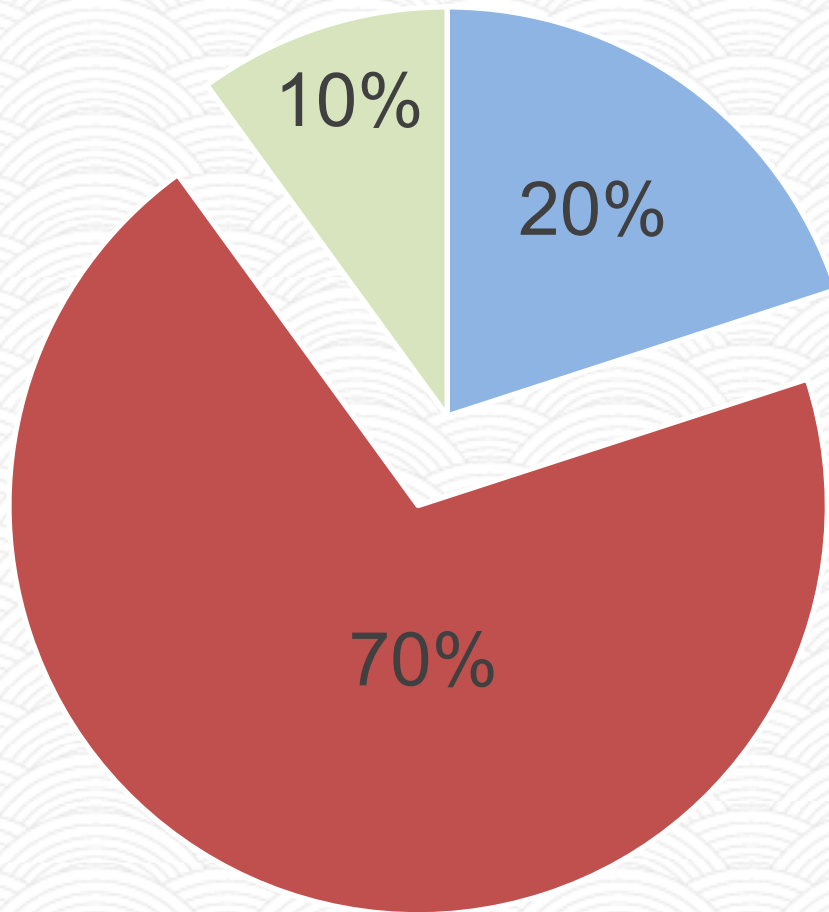


# Search Accuracy - Error Analysis



- Mentions needing context to resolve against multiple matches
  - “lecture 4 ”, could refer to “lecture 4.1 ”, “lecture 4.2 ”, because both have textual components ‘4’ in slug name

# Search Accuracy - Error Analysis

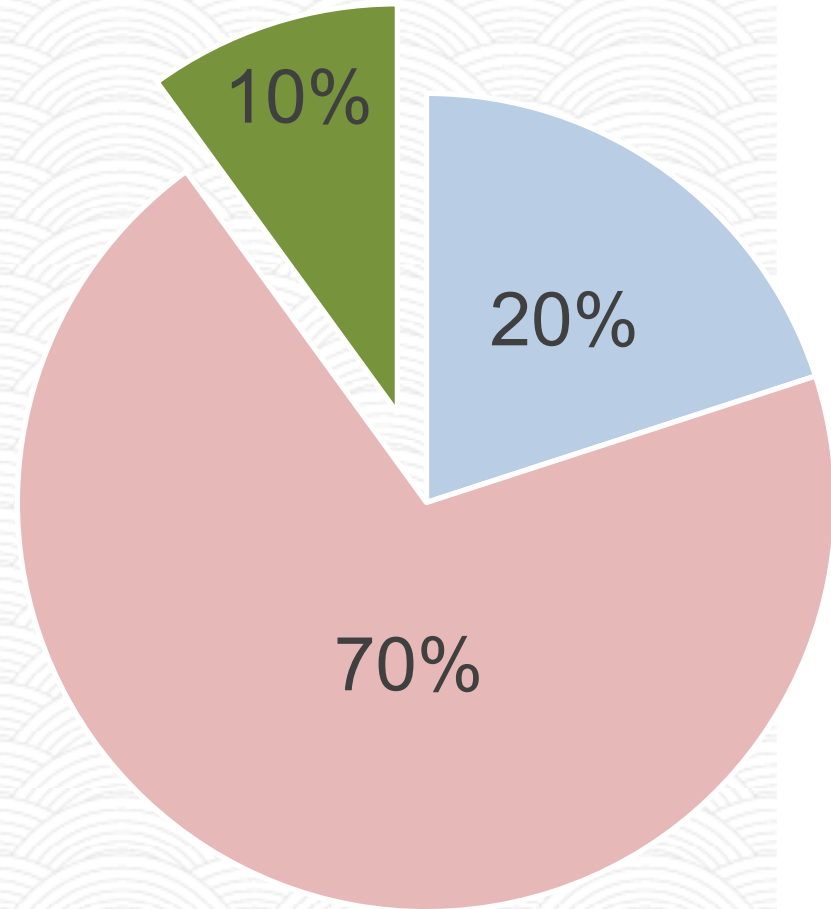


- **Multiple potential targets**
  - mention states “*question 3*”, but there are multiple quizzes, all of which have a Question 3, the target is ambiguous



# Search Accuracy - Error Analysis

- **Errors in extraction**
  - “*the transcript of lecture 2*”,  
partial mention extraction



# Outline

Introduction

Methods

**Evaluation**

- Search Accuracy
- **Mention Extraction Coverage**

Discussion

Conclusion

# Mention Coverage Evaluation

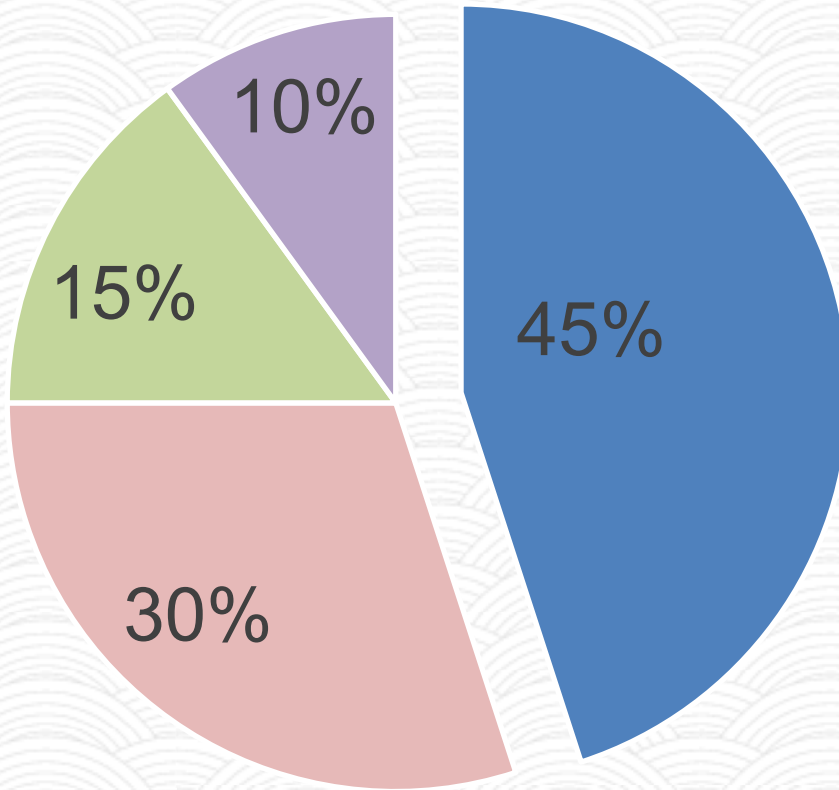
Annotation ID	# of Posts	# of Posts identified as having mentions	# Extracted by our Wikifier	# Correct	Coverage
Annotator 1	1087	156	5	5	14.4%
Annotator 2		175		5	16.1%
Overall		196 (Union)		5	18.0%

- Randomly sample ~1,000 posts to check the actual coverage;
- 18% posts or more contain resource mentions;
- Our <“keyword” + number> pattern has a 100% precision but low recall, covering only about 2.6% of possible mentions.



- How can we improve mention extraction coverage?
- We examine the causes for the coverage disparity, where the parenthetical percentage is determined over the same sampled data.

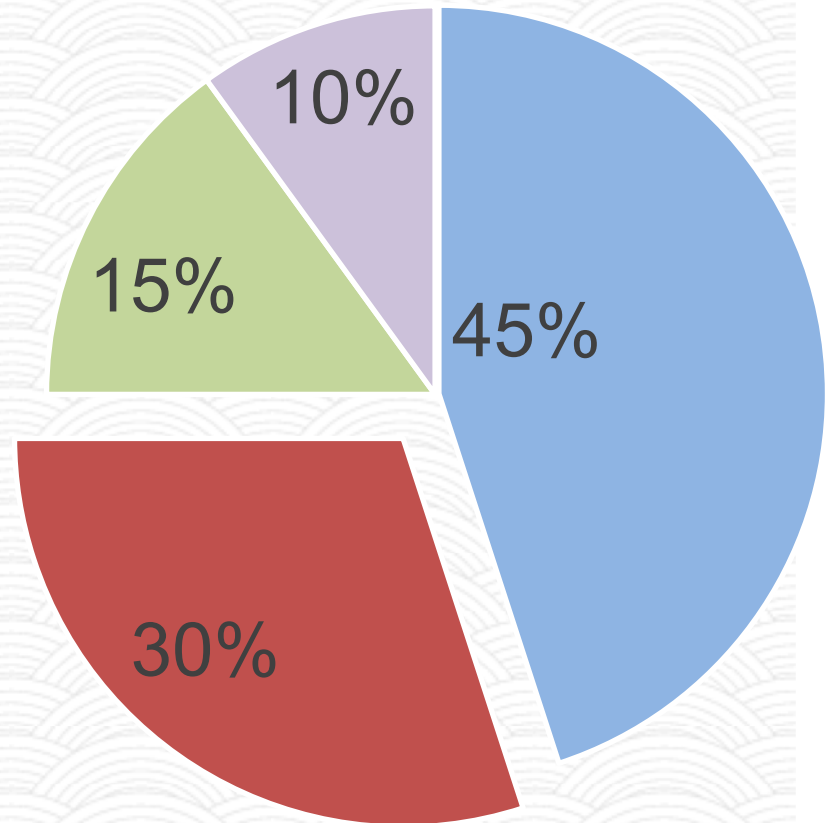
# Coverage Analysis



- Implicit Contextual Expressions
  - Pronouns, e.g., “this video”, “that assignment”

# Coverage Analysis

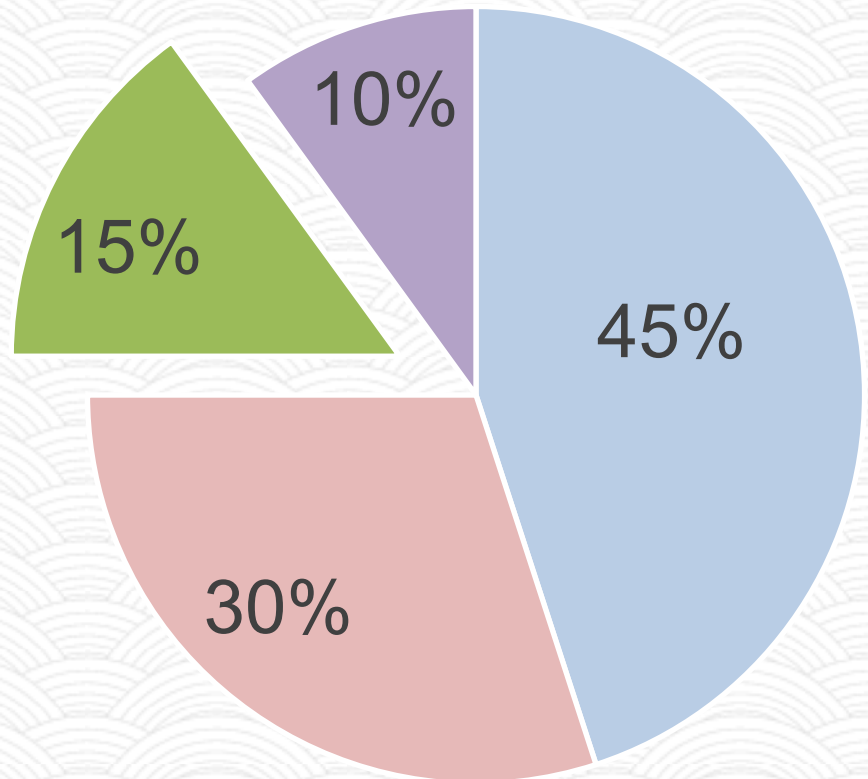
- **Named Reference**
  - resource name, e.g.,  
“the problem “Hashing  
with chains””





# Coverage Analysis

- **Informal Expressions**
  - colloquial expressions, e.g., “*programming assignment of ML*”, “*ex2/week3*”



# Discussion

- Both clearly point forward improving coverage of mention extraction;
- To incorporate contextual knowledge to improve the extraction and resolution.

# Conclusion

Let's make linked open data a reality for learners

- Wikification takes us closer to seamless interlinking;
- Introduce the MUIR framework;
- Proof of concept for simple mention types
  - We need better **mention extraction** and **search resolution**.



## Hidden Slides



电子科技大学  
University of Electronic Science and Technology of China

# Late Breaking Results Performance

Performance	Marco			Micro		
	Precision	Recall	F1 score	Precision	Recall	F1 score
Bi-LSTM	63.52	38.39	47.03	71.53	47.32	56.96
CRF	<b>77.57</b>	65.19	70.36	<b>78.08</b>	69.09	73.31
Bi-LSTM-CRF	73.09	65.65	68.98	76.43	68.98	72.52
Bi-LSTM-CRF + CE	71.52	67.66	69.41	75.40	72.38	73.90
Bi-LSTM-CRF + CE (pre-trained)	71.36	74.57	<b>72.67</b>	73.76	77.62	75.64
Bi-LSTM-CRF + PCE + Encoder	70.71	<b>75.06</b>	72.28	72.91	<b>79.20</b>	<b>75.92</b>