

Estimating sentiment in Eli:  
Computational analysis of tone in student  
responses to student writing

Michael Wojcik

Rhetoric & Writing, Michigan State University  
Principal Software Systems Developer, Micro Focus International

CSE 842  
25 April 2011

- sentiment analysis and rhetorical tone
- Eli and student peer reviews
- UIMA
- data and method
- preliminary results
- further work

# Introduction

- Two main types of textual information.
  - Facts and Opinions
- Most current text information processing methods (e.g., web search, text mining) work with factual information.
- Sentiment analysis or opinion mining
  - computational study of opinions, sentiments and emotions expressed in text.
- Why opinion mining now? Mainly because of the Web; huge volumes of opinionated text.

# Easier and Harder Problems

- Reviews are easier.
    - Objects/entities are given (almost), and little noise
  - Forum discussions and blogs are harder.
    - Objects are not given, and a large amount of noise
- 
- Determining sentiments seems to be easier.
  - Determining objects and their corresponding features is harder.
  - Combining them is even harder.

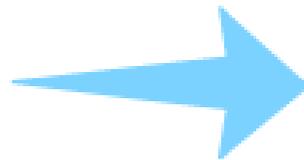
# Eli: a peer-review writing system

MSU WIDE and Red Cedar Solutions

Start

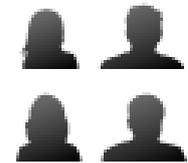
## Step 1: Review Created

The review coordinator selects reviewers, identifies review objects, and defines review criteria.



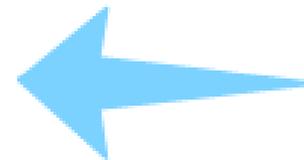
## Step 2: Reviewers Analyze

Reviewers examine review objects\* using framework or criteria specified by the coordinator.



## Step 4: Writer Responds

The writer(s) of the review objects give metric responses about the helpfulness of their feedback.



## Step 3: Reviewers Respond

Reviewers write suggestions\* and respond to metric evaluations as specified by the coordinator.

## Step 5: Coordinator Responds

The review coordinator can also respond to reviewer feedback, endorsing individual ideas.



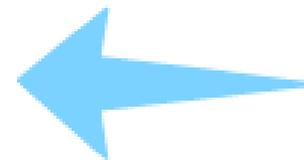
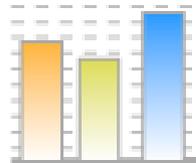
## Step 6: Writer Strategizes

Writers can add individual pieces of review feedback to an itemized, prioritized plan for revision.



## Step 8: Reviews Assessed

Data visualizations can show coordinators helpful reviewers and reviewers how to be more helpful.



## Step 7: Writer Revises

Writers compose new versions of their texts\* and flag feedback that directly contributed to it.



### Criteria Responses By Group Members

1. Indicate if the annotation includes the following features:

Response	Michael	Class	Endorsed
Text may be easily repurposed for final paper?	67 %	57 %	<b>NO</b>
A brief summary of the work?	100 %	93 %	<b>NO</b>
Situates work in a broader conversation?	67 %	70 %	<b>NO</b>
Complete Citation?	100 %	100 %	<b>NO</b>

### Ratings

4.0 How well does the abstract communicate the purpose and content of the article? 1 = not well and 5 = very well!

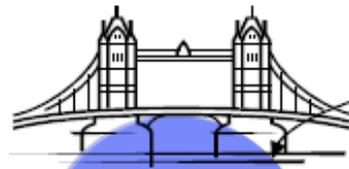
3.3 Annotation situates the work in a broader conversation or controversy 1=boo 5=awesome!

### Suggestions and Comments by Group Members

Suggestion	more specific examples in the summary section would make the specific discussions in the later sections easier to follow
Endorsed:	<b>NO</b>
Rate:	☆☆☆☆☆
Comments	I am supper interested in this article and I think you pull out the most useful and interesting points, its just not as clear as it could be for a person who has not read the article
Endorsed:	<b>NO</b>
Rate:	☆☆☆☆☆
S	

# UIMA: standard, framework, and tools for processing unstructured data

**Analytics** bridge the Unstructured & Structured worlds



UIMA

Unstructured Information



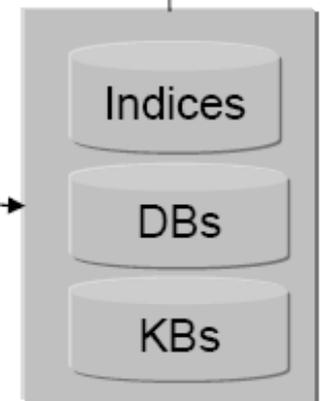
Text, Chat,  
Email, Audio,  
Video

*High-Value  
Most Current Content  
BUT ...*

*Buried in Huge Volumes  
Lots of Noise, Implicit Semantics  
Inefficient Search*

- Identify Semantic Entities, Induce Structure
- Chats, Phone Calls, Transfers
- People, Places, Org, Events
- Times, Topics, Opinions, Relationships
- Threats, Plots, etc.

Structured Information



*Explicit Structure  
Explicit Semantics  
Efficient Search  
Focused Content*

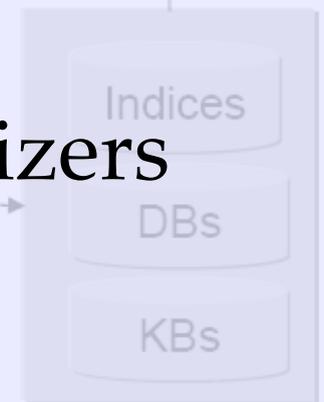
# UIMA: standard, framework, and tools for processing unstructured data

- invented at IBM
- standardized by OASIS
- maintained by Apache
- many free components
  - tokenizers, stemmers, entity recognizers
  - OpenNLP
  - Java, C++, other languages
- scale-out
- used by IBM Watson

Analytics bridge the  
Unstructured & Structured worlds



Structured  
Information



Explicit Structure  
Explicit Semantics  
Efficient Search  
Focused Content

Unstructured  
Information

Text  
Email, Audio,

Video

High-Value  
Most Current Content  
BUT ...  
Buried in Huge Volumes  
Lots of Noise, Inefficient Search

- Identify Semantic Entities, Induce Structure
- Chats, Phone Calls, Transfers
- People, Places, Org, Events
- Times, Topics, Opinions, Relationships
- Threats, Photos, etc.

Showing rows 90 - 104 (3,682 total, Query took 0.0007 sec)

```
SELECT *
FROM 'activities_raterespond'
LIMIT 90 , 15
```

Profiling [[Edit](#)] [[Explain SQL](#)] [[Create PHP Code](#)] [[Refresh](#)]

Show : 15 row(s) starting from record # 105
 

 Page number: 7

in horizontal (rotated headers) mode and repeat headers after 100 cells

Sort by key: None

+ Options

			ID	ActivityID	ObjectID	DeliverableID	UserID	GroupID	Rating	Best	Comment	DateSubmitted	Sentiment
<input type="checkbox"/>			103	2	167	2	24	1	2	0	<p id="1">a bit short. there are no suggestions fo...	2007-10-08 16:37:28	NULL
<input type="checkbox"/>			104	2	181	3	14	1	4	0	<p id="1">Great Resume!</p>	2007-10-08 16:37:36	NULL
<input type="checkbox"/>			105	2	189	3	32	2	3	0	<p id="1">written well, and consistent with what i...	2007-10-08 16:37:43	NULL
<input type="checkbox"/>			106	2	192	2	18	1	3	0	<p id="1">Nice resume, supports what the qualifica...	2007-10-08 16:38:27	NULL
<input type="checkbox"/>			107	2	187	3	9	2	3	0	<p id="1">Great Reagan quote. Hope Team VISTA lik...	2007-10-08 16:38:30	NULL
<input type="checkbox"/>			108	2	184	2	19	1	3	0	<p id="1">No suggestion of ideas for the organizat...	2007-10-08 16:38:50	NULL
<input type="checkbox"/>			109	2	188	2	32	2	2	0	<p id="1">Confusing at times, and also could be mo...	2007-10-08 16:38:51	NULL
<input type="checkbox"/>			110	2	177	2	17	1	4	0	<p id="1">Nice job covering all bases. I like the...	2007-10-08 16:38:52	NULL
<input type="checkbox"/>			111	2	166	3	24	1	2	0	<p id="1">short statement of qualifications. does ...	2007-10-08 16:38:54	NULL
<input type="checkbox"/>			112	2	175	2	14	1	4	0	<p id="1">Really good. I like how she states why ...	2007-10-08 16:38:57	NULL
<input type="checkbox"/>			113	2	175	2	13	1	4	0	<p id="1">Your suggestions are great and your resu...	2007-10-08 16:39:07	NULL
<input type="checkbox"/>			114	2	178	3	29	1	4	1	<p id="1">Good job.</p>	2007-10-08 16:39:12	NULL
<input type="checkbox"/>			115	2	190	2	21	2	4	0	<p id="1">Very articulate</p>	2007-10-08 16:39:20	NULL
<input type="checkbox"/>			116	2	176	3	13	1	4	0	<p id="1">Nice.</p>	2007-10-08 16:39:23	NULL
<input type="checkbox"/>			117	2	176	3	14	1	4	0	<p id="1">Good resume!</p>	2007-10-08 16:39:25	NULL

With selected:

Show : 15 row(s) starting from record # 105
 

 Page number: 7

in horizontal (rotated headers) mode and repeat headers after 100 cells

Query results operations

[Print view](#)
[Print view \(with full texts\)](#)
[Export](#)
[CREATE VIEW](#)

Showing rows 90 - 104 (3,682 total, Query took 0.0007 sec)

```
SELECT *
FROM 'activities_raterespond'
LIMIT 90 , 15
```

 Profiling [[Edit](#)] [[Explain SQL](#)] [[Create PHP Code](#)] [[Refresh](#)]

Show : 15 row(s) starting from record # 105
 

 Page number: 7

in horizontal (rotated headers) mode and repeat headers after 100 cells

Sort by key: None

+ Options

<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	ID	ActivityID	ObjectID	DeliverableID	UserID	GroupID	Rating	Best	Comment	DateSubmitted	Sentiment
<input type="checkbox"/>			103	2	167	2	24	1	2	0	<p id="1">a bit short. there are no suggestions fo...	2007-10-08 16:37:28	NULL
<input type="checkbox"/>			104	2	181	3	14	1	4	0	<p id="1">Great Resume!</p>	2007-10-08 16:37:36	NULL
<input type="checkbox"/>			105	2	189	3	32	2	3	0	<p id="1">written well, and consistent with what i...	2007-10-08 16:37:43	NULL
<input type="checkbox"/>			106	2	192	2	18	1	3	0	<p id="1">Nice resume, supports what the qualifica...	2007-10-08 16:38:27	NULL
<input type="checkbox"/>			107	2	187	3	9	2	3	0	<p id="1">Great Reagan quote. Hope Team VISTA lik...	2007-10-08 16:38:30	NULL
<input type="checkbox"/>			108	2	184	2	19	1	3	0	<p id="1">No suggestion of ideas for the organizat...	2007-10-08 16:38:50	NULL
<input type="checkbox"/>			109	2	188	2	32	2	2	0	<p id="1">Confusing at times, and also could be mo...	2007-10-08 16:38:51	NULL
<input type="checkbox"/>			110	2	177	2	17	1	4	0	<p id="1">Nice job covering all bases. I like the...	2007-10-08 16:38:52	NULL
<input type="checkbox"/>			111	2	166	3	24	1	2	0	<p id="1">short statement of qualifications. does ...	2007-10-08 16:38:54	NULL
<input type="checkbox"/>			112	2	175	2	14	1	4	0	<p id="1">Really good. I like how she states why ...	2007-10-08 16:38:57	NULL
<input type="checkbox"/>			113	2	175	2	13	1	4	0	<p id="1">Your suggestions are great and your resu...	2007-10-08 16:39:07	NULL
<input type="checkbox"/>			114	2	178	3	29	1	4	1	<p id="1">Good job.</p>	2007-10-08 16:39:12	NULL
<input type="checkbox"/>			115	2	190	2	21	2	4	0	<p id="1">Very articulate</p>	2007-10-08 16:39:20	NULL
<input type="checkbox"/>			116	2	176	3	13	1	4	0	<p id="1">Nice.</p>	2007-10-08 16:39:23	NULL
<input type="checkbox"/>			117	2	176	3	14	1	4	0	<p id="1">Good resume!</p>	2007-10-08 16:39:25	NULL

 Check All / Uncheck All With selected: 

Show : 15 row(s) starting from record # 105
 

 Page number: 7

in horizontal (rotated headers) mode and repeat headers after 100 cells

Query results operations

[Print view](#)
[Print view \(with full texts\)](#)
[Export](#)
[CREATE VIEW](#)

Showing rows 90 - 104 (3,682 total, Query took 0.0007 sec)

```
SELECT *
FROM 'activities_raterespond'
LIMIT 90 , 15
```

 Profiling [[Edit](#)] [[Explain SQL](#)] [[Create PHP Code](#)] [[Refresh](#)]

Show : 15 row(s) starting from record # 105
 

 Page number: 7

in horizontal (rotated headers) mode and repeat headers after 10 cells

Sort by key: None

+ Options

			ID	ActivityID	ObjectID	DeliverableID	UserID	GroupID	Rating	Best	Comment	DateSubmitted	Sentiment
<input type="checkbox"/>			103	2	167	2	24	1	2	0	<p id="1">a bit short. there are no suggestions fo...	2007-10-08 16:37:28	NULL
<input type="checkbox"/>			104	2	181	3	14	1	4	0	<p id="1">Great Resume!</p>	2007-10-08 16:37:36	NULL
<input type="checkbox"/>			105	2	189	3	32	2	3	0	<p id="1">written well, and consistent with what i...	2007-10-08 16:37:43	NULL
<input type="checkbox"/>			106	2	192	2	18	1	3	0	<p id="1">Nice resume, supports what the qualifica...	2007-10-08 16:38:27	NULL
<input type="checkbox"/>			107	2	187	3	9	2	3	0	<p id="1">Great Reagan quote. Hope Team VISTA lik...	2007-10-08 16:38:30	NULL
<input type="checkbox"/>			108	2	184	2	19	1	3	0	<p id="1">No suggestion of ideas for the organizat...	2007-10-08 16:38:50	NULL
<input type="checkbox"/>			109	2	188	2	32	2	2	0	<p id="1">Confusing at times, and also could be mo...	2007-10-08 16:38:51	NULL
<input type="checkbox"/>			110	2	177	2	17	1	4	0	<p id="1">Nice job covering all bases. I like the...	2007-10-08 16:38:52	NULL
<input type="checkbox"/>			111	2	166	3	24	1	2	0	<p id="1">short statement of qualifications. does ...	2007-10-08 16:38:54	NULL
<input type="checkbox"/>			112	2	175	2	14	1	4	0	<p id="1">Really good. I like how she states why ...	2007-10-08 16:38:57	NULL
<input type="checkbox"/>			113	2	175	2	13	1	4	0	<p id="1">Your suggestions are great and your resu...	2007-10-08 16:39:07	NULL
<input type="checkbox"/>			114	2	178	3	29	1	4	1	<p id="1">Good job.</p>	2007-10-08 16:39:12	NULL
<input type="checkbox"/>			115	2	190	2	21	2	4	0	<p id="1">Very articulate</p>	2007-10-08 16:39:20	NULL
											<p id="2">Brings ...		
<input type="checkbox"/>			116	2	176	3	13	1	4	0	<p id="1">Nice.</p>	2007-10-08 16:39:23	NULL
<input type="checkbox"/>			117	2	176	3	14	1	4	0	<p id="1">Good resume!</p>	2007-10-08 16:39:25	NULL

 Check All / Uncheck All With selected: 

Show : 15 row(s) starting from record # 105
 

 Page number: 7

in horizontal (rotated headers) mode and repeat headers after 100 cells

Query results operations

[Print view](#)
[Print view \(with full texts\)](#)
[Export](#)
[CREATE VIEW](#)

 Showing rows 90 - 104 (3,682 total, Query took 0.0007 sec)

```
SELECT *
FROM 'activities_raterespond'
LIMIT 90 , 15
```

 Profiling [[Edit](#)] [[Explain SQL](#)] [[Create PHP Code](#)] [[Refresh](#)]


Show : 15 row(s) starting from record # 105



Page number: 7

horizontal (rotated headers) mode and repeat headers after 100 cells

Sort by key: None

+ Options

			ID	ActivityID	ObjectID	DeliverableID	UserID	GroupID	Rating	Best	Comment	DateSubmitted	Sentiment
<input type="checkbox"/>			103	2	167	2	24	1	2	0	<p id="1">a bit short. there are no suggestions fo...	2007-10-08 16:37:28	NULL
<input type="checkbox"/>			104	2	181	3	14	1	4	0	<p id="1">Great Resume!</p>	2007-10-08 16:37:36	NULL
<input type="checkbox"/>			105	2	189	3	32	2	3	0	<p id="1">written well, and consistent with what i...	2007-10-08 16:37:43	NULL
<input type="checkbox"/>			106	2	192	2	18	1	3	0	<p id="1">Nice resume, supports what the qualifica...	2007-10-08 16:38:27	NULL
<input type="checkbox"/>			107	2	187	3	9	2	3	0	<p id="1">Great Reagan quote. Hope Team VISTA lik...	2007-10-08 16:38:30	NULL
<input type="checkbox"/>			108	2	184	2	19	1	3	0	<p id="1">No suggestion of ideas for the organizat...	2007-10-08 16:38:50	NULL
<input type="checkbox"/>			109	2	188	2	32	2	2	0	<p id="1">Confusing at times, and also could be mo...	2007-10-08 16:38:51	NULL
<input type="checkbox"/>			110	2	177	2	17	1	4	0	<p id="1">Nice job covering all bases. I like the...	2007-10-08 16:38:52	NULL
<input type="checkbox"/>			111	2	166	3	24	1	2	0	<p id="1">short statement of qualifications. does ...	2007-10-08 16:38:54	NULL
<input type="checkbox"/>			112	2	175	2	14	1	4	0	<p id="1">Really good. I like how she states why ...	2007-10-08 16:38:57	NULL
<input type="checkbox"/>			113	2	175	2	13	1	4	0	<p id="1">Your suggestions are great and your resu...	2007-10-08 16:39:07	NULL
<input type="checkbox"/>			114	2	178	3	29	1	4	1	<p id="1">Good job.</p>	2007-10-08 16:39:12	NULL
<input type="checkbox"/>			115	2	190	2	21	2	4	0	<p id="1">Very articulate</p>	2007-10-08 16:39:20	NULL
<input type="checkbox"/>			116	2	176	3	13	1	4	0	<p id="1">Nice.</p>	2007-10-08 16:39:23	NULL
<input type="checkbox"/>			117	2	176	3	14	1	4	0	<p id="1">Good resume!</p>	2007-10-08 16:39:25	NULL

 Check All /  Uncheck All With selected: 


Show : 15 row(s) starting from record # 105



Page number: 7

in horizontal (rotated headers) mode and repeat headers after 100 cells

Query results operations

[Print view](#)
[Print view \(with full texts\)](#)
[Export](#)
[CREATE VIEW](#)

1741 4 I really like your outline. I think you need an attention getter. You need a conclusion. It's lacking :)

1748 1 I have no idea what your topic is. You look like you are going in the right way with your outline, and the ideas are all there. The crash itself could be a lot more detailed and specific, ex. What were the people thinking while the plane was going down? How did they survive the crash? You could also go into a lot more detail about the 72 days on the mountain, there had to be more than just eating each other and finding a place to survive, i.e. how did they stay warm? How many people survived and who stuck together? How else could they get food besides cannibalism?

1809 1 Wasn't available to review in class

1818 4 I think you had more guiding than maintaining, but overall, the letter was good.

2839 4 Section 508 and JimThatcher.com are cited in the memo, both appropriate for her argument that further research is needed on whether or not Home Depot needs to revise their website.

Cat analyzed the Southwest Airlines case as a great example of how websites are subject to ADA. "Southwest Airlines was subject to Section 508 to the standpoint that, as a place of public accommodation, Southwest Airlines was subject to the ADA". The next sentence after is an argument that websites need to be accessible to those with disabilities.

She also used the Target case as well to compare to Home Depot's predicament. She goes on to saying that further research needs to be done on whether Home Depot is considered a Federal agency and whether they need to recode their website if it is deemed inaccessible under Section 508 compliance and the ADA.

She followed the rubric and wrote a great preliminary report of what needs to be done next for SNG to research. I would say this was written really well, straight to the point with valid arguments based off of her analyses.

3035 4 so the comment below was supposed to go here. haha. I really like your memo, because I understand how and why you did what you did. your design gives me something to think about in regards to my own draft. way to go!

# Eli data

- 3659 review documents (after removing test entries, etc)
- text and numerical rating assigned by reviewer
- wide variation in length, writing style, content
- I've done manual evaluation of some to establish a gold standard

# Sentiment rating

- Some early systems just use “positive” and “negative”
- Later researchers (eg Gamon & Aue 2005) add “neutral”
- I'm treating “positive” and “negative” as separate features – 2-bit rating:
  - neutral (0)
  - mostly negative (1)
  - mostly positive (2)
  - mixed (3)

# Bag of Words

- My baseline analyzer
- Looks for words that appear in either “negative” or “positive” vocabulary
- Vocabularies created by extracting words from known-negative and known-positive texts (100 each)
- Two modes:
  - Rate text based on relative number of negative/positive terms
  - Rate sentences based on words, and text based on sentences

Java - BowSentimentAnnotator/desc/BowSentimentAnnotator.xml - Eclipse

File Edit Navigate Search Project Run UIMA Window Help

Package Explorer

- BowSentimentAnnotator
  - src
  - JRE System Library [JavaSE-1.6]
  - uima
  - Referenced Libraries
  - conf
  - data
  - desc
  - doc
  - lib
  - metadata
  - resources
- GamonAueAnnotator
- NakagawaInuiKurohashiAnnotator
- OpenNlpTest
- RoomNumberAnnotator
- test
- uimaj-examples

BowSentimentAnnotator.xml

### Overview

**Implementation Details**

Implementation Language:  C/C++  Java

Engine Type:  Primitive  Aggregate

**Runtime Information**

This section describes information about how to run this component

updates the CAS

multiple deployment allowed

Outputs new CASes

Name of the Java class file:

**Overall Identification Information**

This section specifies the basic identification information for this descriptor

Name:

Version:

Vendor:

Description:

Document Analyzer

File Help



**Unstructured Information Management Architecture**  
*An Apache Project.*

Input Directory:

Output Directory:

Location of Analysis Engine XML Descriptor:

XML Tag containing Text (optional):

Language:

Character Encoding:

# Results for bag-of-words

- On small manually-rated data set (N=50)
- Treating “neutral” as absence of feature (recall)
- Considering negative/positive together:  
Precision & Recall both 0.6190 (so F1 also 0.6190)
- Considering negative/positive separately:
  - Positive:  
Precision = 0.7500 Recall = 0.7778 F1 = 0.7636
  - Negative:  
Precision = 0.6800 Recall = 0.7083 F1 = 0.6939

# Next steps

- Other algorithms:
  - Gamon & Aue's minimally-supervised vocabulary builder (compare to Yarowski's word-sense method)
  - Nakagawa *et al* [2010] subtree analysis
  - N-gram HMM (adapted from POS tagger)? MEMM (OpenNLP)?
- Manually annotate more data
- Typo / spelling correction (MED)?
- Compare sentiment to:
  - Reviewer's rating
  - Reviewer's helpfulness score
- Integrate into Eli

Estimating sentiment in Eli:  
Computational analysis of tone in student  
responses to student writing

Michael Wojcik

Rhetoric & Writing, Michigan State University  
Principal Software Systems Developer, Micro Focus International

CSE 842  
25 April 2011

I'm Michael Wojcik, from the Rhetoric program over in Arts&Letters, where I'm working on computational rhetoric. My project is looking at sentiment analysis of student peer review comments – these are comments that students write about one another's writing.

- sentiment analysis and rhetorical tone
- Eli and student peer reviews
- UIMA
- data and method
- preliminary results
- further work

I want to say a couple of words about sentiment analysis, the Eli system for writing review, the UIMA data-processing system, my data and method, some preliminary results, and where I'm going to take this from here.

## Introduction

- Two main types of textual information.
  - Facts and Opinions
- Most current text information processing methods (e.g., web search, text mining) work with factual information.
- Sentiment analysis or opinion mining
  - computational study of opinions, sentiments and emotions expressed in text.
- Why opinion mining now? Mainly because of the Web; huge volumes of opinionated text.

This is a slide from a presentation by Bing Liu, a major sentiment-analysis researcher. He defines sentiment analysis as trying to extract opinions, feelings, and emotions from text. Most of you probably know this is big business these days – there's a lot of research being done and a lot of money behind it because this information is very valuable to industry.

Now, I'm actually looking at rhetorical tone, which is defined as the perceived attitude of the author toward the subject; it's not quite the same thing as sentiment. But sentiment is a good heuristic for it.

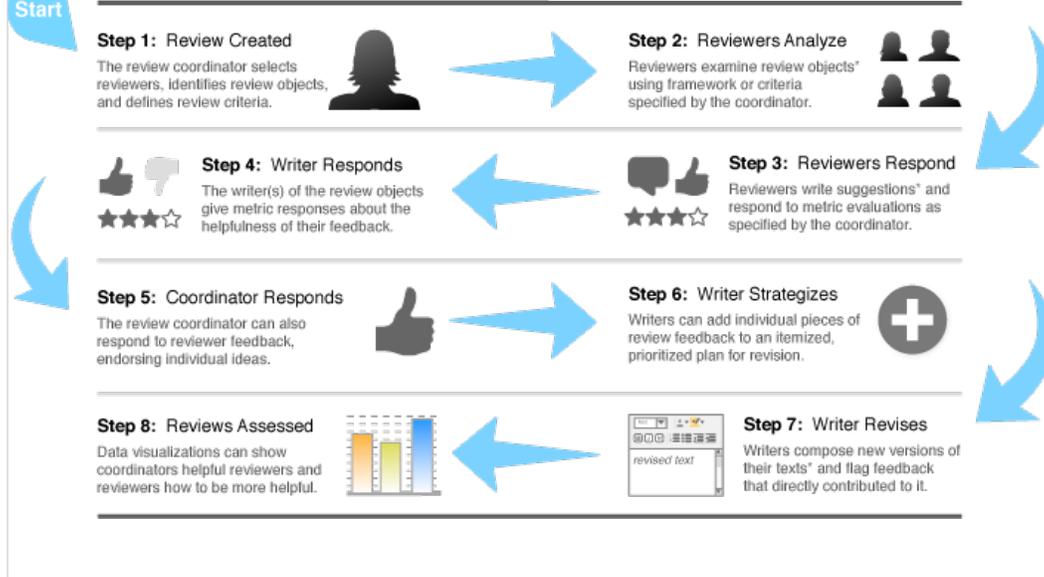
## Easier and Harder Problems

- Reviews are easier.
    - Objects/entities are given (almost), and little noise
  - Forum discussions and blogs are harder.
    - Objects are not given, and a large amount of noise
- 
- Determining sentiments seems to be easier.
  - Determining objects and their corresponding features is harder.
  - Combining them is even harder.
- 

And reviews are a good place to start, for a few reasons. One is that, as Bing Liu points out here, we typically know what a review is about. Also, reviews often have numerical ratings (“four out of five stars”), and we can see if our calculated sentiment correlates to those. And reviews often have pretty clear, simple opinions for human judges to extract so we have a good standard to compare to program output. Student peer reviews are a useful area because composition classes are a profit center for the university, so there's a lot of support for studying them. Peer reviews do have some challenges. Students are reluctant to express strong opinions about each other's work, so sentiment can be more subtle.

# Eli: a peer-review writing system

MSU WIDE and Red Cedar Solutions



I'm working with data from Eli, a web-based peer review system developed at Michigan State, at the center for Writing in Digital Environments, and being commercialized by Red Cedar Solutions. It implements an entire review and revision workflow, and it accumulates various metrics about review, including a composite "helpfulness" score for each reviewer. But that's based all on workflow – it doesn't look at the content of review comments. So that's what I'm looking at adding to Eli.

AL 841: Professional Writing Theory & Research - Mozilla Firefox

http://elireview.com/unit/student/writing-response/module\_id/60/activity\_id/99

linguistic data consor...  
 Disable Cookies CSS Forms Images Information Miscellaneous Outline Resize Tools View Source Options

Criteria Responses By Group Members

1. Indicate if the annotation includes the following features:

Response	Michael	Class	Endorsed
Text may be easily repurposed for final paper?	67 %	57 %	NO
A brief summary of the work?	100 %	93 %	NO
Situates work in a broader conversation?	67 %	70 %	NO
Complete Citation?	100 %	100 %	NO

Ratings

4.0 How well does the abstract communicate the purpose and content of the article? 1 = not well and 5 = very well!

3.3 Annotation situates the work in a broader conversation or controversy 1=boo 5=awesome!

Suggestions and Comments by Group Members

Suggestion

more specific examples in the summary section would make the specific discussions in the later sections easier to follow

Endorsed: **NO** Rate: ★★★★★

Comments

I am supper interested in this article and I think you pull out the most useful and interesting points, its just not as clear as it could be for a person who has not read the article

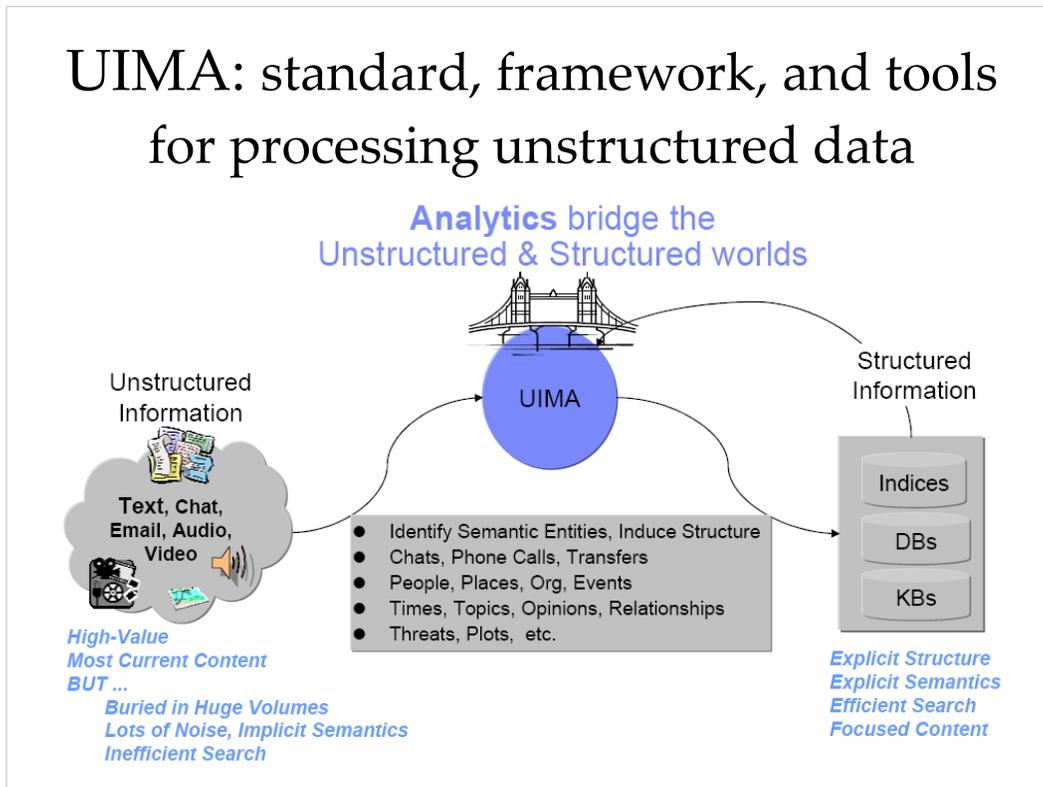
Endorsed: **NO** Rate: ★★★★★

Done

JSESSIONID=undefined Apr 22 04/22/2011 21:50:26 385M 0% #0: 0/0

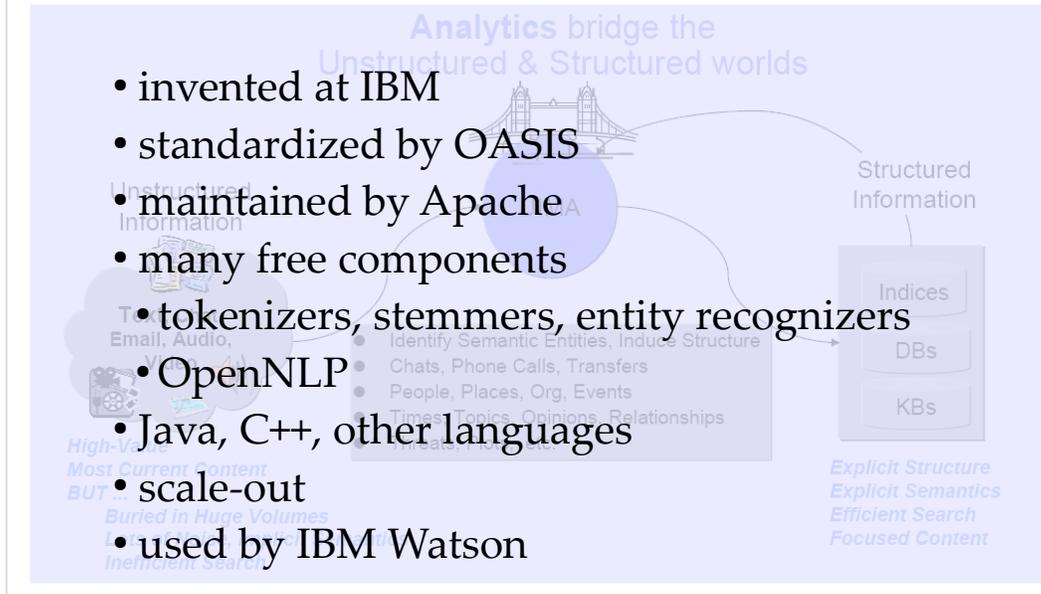
This is one of Eli's views. I know this is hard to read, but basically authors get reviews of their writing with numerical evaluations, comments, and other kinds of feedback.

# UIMA: standard, framework, and tools for processing unstructured data



So to implement the sentiment-analysis algorithms I'm experimenting with, I'm using Apache UIMA. UIMA is a standard architecture for processing unstructured data. Basically, it provides ways to add structure and metadata to unstructured text, images, audio and video, and so on.

## UIMA: standard, framework, and tools for processing unstructured data



UIMA was invented at IBM, standardized by OASIS, and is now maintained by Apache. It has lots of free components – basic things like tokenizers, entity recognizers, and wrappers for various OpenNLP functions. You can write your own components in Java or C++, or use bridges for other languages. It has various scale-out mechanisms for service or grid computing. And it was part of the Watson Jeopardy-winning system (and you can get a semantic query engine for it from IBM).

I could say a lot more if I had time. Check it out.

localhost > eli > activities\_raterespond

Showing rows 90 - 104 (3,682 total, Query took 0.0007 sec)

```

SELECT *
FROM 'activities_raterespond'
LIMIT 90, 15

```

Profiling [ Edit ] [ Explain SQL ] [ Create PHP Code ] [ Refresh ]

Show: 15 row(s) starting from record # 105  
in horizontal (rotated headers) mode and repeat headers after 100 cells

Sort by key: None

ID	ActivityID	ObjectID	DeliverableID	UserID	GroupID	Rating	Best	Comment	DateSubmitted	Sentiment
103	2	167	2	24	1	2	0	<p id="1">a bit short, there are no suggestions fo...	2007-10-08 16:37:28	NULL
104	2	181	3	14	1	4	0	<p id="1">Great Resume</p>	2007-10-08 16:37:36	NULL
105	2	189	3	32	2	3	0	<p id="1">written well, and consistent with what L...	2007-10-08 16:37:43	NULL
106	2	192	2	18	1	3	0	<p id="1">Nice resume, supports what the qualifica...	2007-10-08 16:38:27	NULL
107	2	187	3	9	2	3	0	<p id="1">Great Reagan quote. Hope Team VISTA lik...	2007-10-08 16:38:30	NULL
108	2	184	2	19	1	3	0	<p id="1">No suggestion of ideas for the organizati...	2007-10-08 16:38:50	NULL
109	2	188	2	32	2	2	0	<p id="1">Confusing at times, and also could be mo...	2007-10-08 16:38:51	NULL
110	2	177	2	17	1	4	0	<p id="1">Nice job covering all bases. I like the...	2007-10-08 16:38:52	NULL
111	2	166	3	24	1	2	0	<p id="1">short statement of qualifications. does ...	2007-10-08 16:38:54	NULL
112	2	175	2	14	1	4	0	<p id="1">Really good. I like how she states why ...	2007-10-08 16:38:57	NULL
113	2	175	2	13	1	4	0	<p id="1">Your suggestions are great and your resu...	2007-10-08 16:39:07	NULL
114	2	178	3	29	1	4	1	<p id="1">Good job.</p>	2007-10-08 16:39:12	NULL
115	2	190	2	21	2	4	0	<p id="1">Very articulate</p>	2007-10-08 16:39:20	NULL
116	2	176	3	13	1	4	0	<p id="2">Brings ... <p id="1">Nice.</p>	2007-10-08 16:39:23	NULL
117	2	178	3	14	1	4	0	<p id="1">Good resume</p>	2007-10-08 16:39:25	NULL

Check All / Uncheck All With selected:

Show: 15 row(s) starting from record # 105  
in horizontal (rotated headers) mode and repeat headers after 100 cells

Query results operations  
[Print view](#) [Print view \(with full texts\)](#) [Export](#) [CREATE VIEW](#)

I started by extracting reviews from a snapshot of the Eli database.

localhost > eli > activities\_raterespond

Showing rows 90 - 104 (3,682 total, Query took 0.0007 sec)

```

SELECT *
FROM 'activities_raterespond'
LIMIT 90, 15

```

Profiling [ Edit ] [ Explain SQL ] [ Create PHP Code ] [ Refresh ]

Show: 15 row(s) starting from record # 105 in horizontal (rotated headers) mode and repeat headers after 100 cells Page number: 7

Sort by key: None

ID	ActivityID	ObjectID	DeliverableID	UserID	GroupID	Rating	Best	Comment	DateSubmitted	Sentiment
103	2	167	2	24	1	2	0	<p id="1">a bit short, there are no suggestions fo...	2007-10-08 16:37:28	NULL
104	2	181	3	14	1	4	0	<p id="1">Great Resume!</p>	2007-10-08 16:37:36	NULL
105	2	189	3	32	2	3	0	<p id="1">written well, and consistent with what L...	2007-10-08 16:37:43	NULL
106	2	192	2	18	1	3	0	<p id="1">Nice resume, supports what the qualifica...	2007-10-08 16:38:27	NULL
107	2	187	3	9	2	3	0	<p id="1">Great Reagan quote. Hope Team VISTA lik...	2007-10-08 16:38:30	NULL
108	2	184	2	19	1	3	0	<p id="1">No suggestion of ideas for the organizat...	2007-10-08 16:38:50	NULL
109	2	188	2	32	2	2	0	<p id="1">Confusing at times, and also could be mo...	2007-10-08 16:38:51	NULL
110	2	177	2	17	1	4	0	<p id="1">Nice job covering all bases. I like the...	2007-10-08 16:38:52	NULL
111	2	166	3	24	1	2	0	<p id="1">short statement of qualifications. does ...	2007-10-08 16:38:54	NULL
112	2	175	2	14	1	4	0	<p id="1">Really good. I like how she states why ...	2007-10-08 16:38:57	NULL
113	2	175	2	13	1	4	0	<p id="1">Your suggestions are great and your resu...	2007-10-08 16:39:07	NULL
114	2	178	3	29	1	4	1	<p id="1">Good job.</p>	2007-10-08 16:39:12	NULL
115	2	190	2	21	2	4	0	<p id="1">Very articulate</p>	2007-10-08 16:39:20	NULL
116	2	176	3	13	1	4	0	<p id="2">Brings ... <p id="1">Nice.</p>	2007-10-08 16:39:23	NULL
117	2	178	3	14	1	4	0	<p id="1">Good resume</p>	2007-10-08 16:39:25	NULL

Check All / Uncheck All With selected: 15 row(s) starting from record # 105 in horizontal (rotated headers) mode and repeat headers after 100 cells Page number: 7

Query results operations: [Print view](#) [Print view \(with full texts\)](#) [Export](#) [CREATE VIEW](#)

The table containing the reviews contains the text. It has some HTML markup which I stripped off after unloading.

The screenshot shows a MySQL database interface with a table named 'activities\_raterespond'. A red arrow points to the 'Rating' column. The table contains 17 rows of data, each with a unique ID, activity ID, object ID, deliverable ID, user ID, group ID, rating, best flag, comment, date submitted, and sentiment.

ID	ActivityID	ObjectID	DeliverableID	UserID	GroupID	Rating	Best	Comment	DateSubmitted	Sentiment
103	2	167	2	24	1	2	0	<p id="1">a bit short, there are no suggestions fo...	2007-10-08 16:37:28	NULL
104	2	181	3	14	1	4	0	<p id="1">Great Resume!</p>	2007-10-08 16:37:36	NULL
105	2	189	3	32	2	3	0	<p id="1">written well, and consistent with what L...	2007-10-08 16:37:43	NULL
106	2	192	2	18	1	3	0	<p id="1">Nice resume, supports what the qualifica...	2007-10-08 16:38:27	NULL
107	2	187	3	9	2	3	0	<p id="1">Great Reagan quote. Hope Team VISTA lik...	2007-10-08 16:38:30	NULL
108	2	184	2	19	1	3	0	<p id="1">No suggestion of ideas for the organizat...	2007-10-08 16:38:50	NULL
109	2	188	2	32	2	2	0	<p id="1">Confusing at times, and also could be mo...	2007-10-08 16:38:51	NULL
110	2	177	2	17	1	4	0	<p id="1">Nice job covering all bases. I like the...	2007-10-08 16:38:52	NULL
111	2	166	3	24	1	2	0	<p id="1">short statement of qualifications. does ...	2007-10-08 16:38:54	NULL
112	2	175	2	14	1	4	0	<p id="1">Really good. I like how she states why ...	2007-10-08 16:38:57	NULL
113	2	175	2	13	1	4	0	<p id="1">Your suggestions are great and your resu...	2007-10-08 16:39:07	NULL
114	2	178	3	29	1	4	1	<p id="1">Good job.</p>	2007-10-08 16:39:12	NULL
115	2	190	2	21	2	4	0	<p id="1">Very articulate</p>	2007-10-08 16:39:20	NULL
116	2	176	3	13	1	4	0	<p id="2">Brings ... <p id="1">Nice.</p>	2007-10-08 16:39:23	NULL
117	2	178	3	14	1	4	0	<p id="1">Good resume</p>	2007-10-08 16:39:25	NULL

There's also this numerical rating, assigned by reviewers along with their comments, which we expect will correlate somewhat with sentiment;

localhost > eli > activities\_raterespond

Showing rows 90 - 104 (3,682 total, Query took 0.0007 sec)

```

SELECT *
FROM 'activities_raterespond'
LIMIT 90 - 10

```

Profiling [ Edit ] [ Explain SQL ] [ Create PHP Code ] [ Refresh ]

Show: 15 row(s) starting from record # 105

horizontal (rotated headers) mode and repeat headers after 100 cells

Sort by key: None

ID	ActivityID	ObjectID	DeliverableID	UserID	GroupID	Rating	Best	Comment	DateSubmitted	Sentiment
103	2	167	2	24	1	2	0	<p id="1">a bit short, there are no suggestions fo...	2007-10-08 16:37:28	NULL
104	2	181	3	14	1	4	0	<p id="1">Great Resume</p>	2007-10-08 16:37:36	NULL
105	2	189	3	32	2	3	0	<p id="1">written well, and consistent with what L...	2007-10-08 16:37:43	NULL
106	2	192	2	18	1	3	0	<p id="1">Nice resume, supports what the qualifica...	2007-10-08 16:38:27	NULL
107	2	187	3	9	2	3	0	<p id="1">Great Reagan quote. Hope Team VISTA lik...	2007-10-08 16:38:30	NULL
108	2	184	2	19	1	3	0	<p id="1">No suggestion of ideas for the organizat...	2007-10-08 16:38:50	NULL
109	2	188	2	32	2	2	0	<p id="1">Confusing at times, and also could be mo...	2007-10-08 16:38:51	NULL
110	2	177	2	17	1	4	0	<p id="1">Nice job covering all bases. I like the...	2007-10-08 16:38:52	NULL
111	2	166	3	24	1	2	0	<p id="1">short statement of qualifications. does ...	2007-10-08 16:38:54	NULL
112	2	175	2	14	1	4	0	<p id="1">Really good. I like how she states why ...	2007-10-08 16:38:57	NULL
113	2	175	2	13	1	4	0	<p id="1">Your suggestions are great and your resu...	2007-10-08 16:39:07	NULL
114	2	178	3	29	1	4	1	<p id="1">Good job.</p>	2007-10-08 16:39:12	NULL
115	2	190	2	21	2	4	0	<p id="1">Very articulate</p>	2007-10-08 16:39:20	NULL
116	2	176	3	13	1	4	0	<p id="2">Brings ... <p id="1">Nice.</p>	2007-10-08 16:39:23	NULL
117	2	178	3	14	1	4	0	<p id="1">Good resume</p>	2007-10-08 16:39:25	NULL

Check All / Uncheck All With selected:

Show: 15 row(s) starting from record # 105

horizontal (rotated headers) mode and repeat headers after 100 cells

Query results operations

Print view Print view (with full texts) Export CREATE VIEW

and there's a unique ID column (the database is third-normal), so we have a convenient label for each comment.

1741 4 I really like your outline. I think you need an attention getter. You need a conclusion. It's lacking :)

1748 1 I have no idea what your topic is. You look like you are going in the right way with your outline, and the ideas are all there. The crash itself could be a lot more detailed and specific, ex. What were the people thinking while the plane was going down? How did they survive the crash? You could also go into a lot more detail about the 72 days on the mountain, there had to be more than just eating each other and finding a place to survive, i.e. how did they stay warm? How many people survived and who stuck together? How else could they get food besides cannibalism?

1809 1 Wasn't available to review in class

1818 4 I think you had more guiding than maintaining, but overall, the letter was good.

2839 4 Section 508 and JimThatcher.com are cited in the memo, both appropriate for her argument that further research is needed on whether or not Home Depot needs to revise their website.

Cat analyzed the Southwest Airlines case as a great example of how websites are subject to ADA. "Southwest Airlines was subject to Section 508 to the standpoint that, as a place of public accommodation, Southwest Airlines was subject to the ADA". The next sentence after is an argument that websites need to be accessible to those with disabilities.

She also used the Target case as well to compare to Home Depot's predicament. She goes on to saying that further research needs to be done on whether Home Depot is considered a Federal agency and whether they need to recode their website if it is deemed inaccessible under Section 508 compliance and the ADA.

She followed the rubric and wrote a great preliminary report of what needs to be done next for SNG to research. I would say this was written really well, straight to the point with valid arguments based off of her analyses.

3035 4 so the comment below was supposed to go here. haha. I really like your memo, because I understand how and why you did what you did. your design gives me something to think about in regards to my own draft way to go!

There's a lot of variation among the reviews. Some are just a few words; some are a few paragraphs. Sometimes they don't have any opinion content – it's all factual.

## Eli data

- 3659 review documents (after removing test entries, etc)
- text and numerical rating assigned by reviewer
- wide variation in length, writing style, content
- I've done manual evaluation of some to establish a gold standard

After extraction and some preprocessing I ended up with a bit less than 3700 documents, each with the text from one review and the numerical rating assigned by the reviewer. (The rating range is 1 to 5, but no reviewer has ever assigned a 5.)

As you saw on the previous slide, there's a wide variation in length, writing style, vocabulary, topic, etc.

I've been working on manually evaluating sentiment in a test set taken from the data, so I can use it as a gold standard for comparisons. I've also grouped some and annotated some others to create training data. Of course this is a lot of work, so I haven't done very many yet.

## Sentiment rating

- Some early systems just use “positive” and “negative”
- Later researchers (eg Gamon & Aue 2005) add “neutral”
- I'm treating “positive” and “negative” as separate features – 2-bit rating:
  - neutral (0)
  - mostly negative (1)
  - mostly positive (2)
  - mixed (3)

How am I representing sentiment? My system is pretty simple, but a bit more sophisticated than some early ones that just identified a sample as “positive” or “negative”. Later researchers (such as Gamon & Aue [“ow-eh”]) added “neutral” as a possibility.

I'm treating “positive” and “negative” as separate features for each text, which gives me a two-bit system and four possibilities: neutral, mostly negative, mostly positive, and mixed – that is, the passage contains both positive and negative sentiment in roughly equivalent amounts.

## Bag of Words

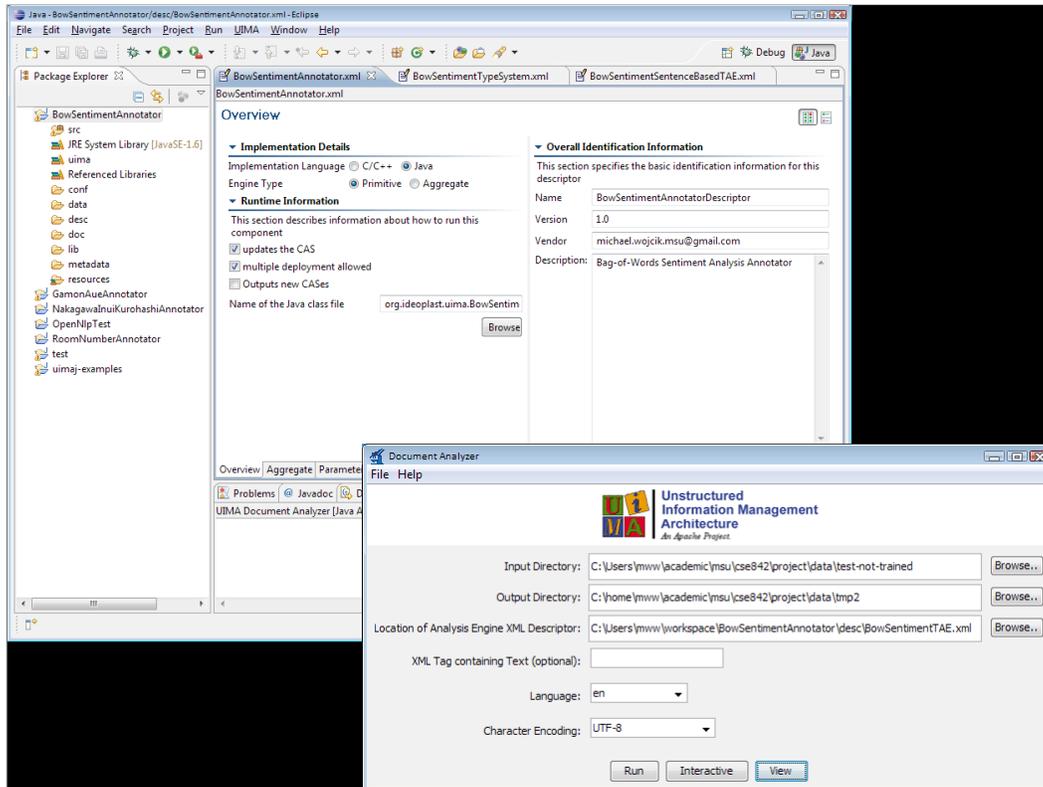
- My baseline analyzer
- Looks for words that appear in either “negative” or “positive” vocabulary
- Vocabularies created by extracting words from known-negative and known-positive texts (100 each)
- Two modes:
  - Rate text based on relative number of negative/positive terms
  - Rate sentences based on words, and text based on sentences

My baseline sentiment analyzer is a bag-of-words algorithm.

So it doesn't look at any structure – it just looks for words in the text that appear in a “negative vocabulary list” or a “positive vocabulary list”.

I created the vocabularies by extracting 100 known-negative reviews and 100 known-positive reviews from my data set, and running a program which built a list of all unique words in each set, with counts. Then the program eliminated words that appeared very infrequently from each set, and if a word appeared in both sets, it eliminated it either from both sets or from the set in which it was less frequent, depending on the ratio of normalized counts.

This analyzer supports two modes. It can rate a document based just on the count of positive and negative words, but it can also rate sentences in the document, and then rate the document based on positive and negative sentence counts. That's more accurate, which confirms Gamon & Aue's claim that most sentences only contain either negative or positive sentiment.



OK, I'm going to do a quick demo.

Demo:

- Show descriptors & code in Eclipse if time.
- Run Document Analyzer.
- Show first 3 docs; if time, can find #2268 for an example of “mixed”.

## Results for bag-of-words

- On small manually-rated data set (N=50)
- Treating “neutral” as absence of feature (recall)
- Considering negative/positive together:  
Precision & Recall both 0.6190 (so F1 also 0.6190)
- Considering negative/positive separately:
  - Positive:  
Precision = 0.7500 Recall = 0.7778 F1 = 0.7636
  - Negative:  
Precision = 0.6800 Recall = 0.7083 F1 = 0.6939

Since I currently don't have a very large gold-standard data set (only 50 texts), I don't have a very accurate test bed. But with that data, if I treat “neutral” in the gold-standard set as the absence of a feature for recall purposes, and take negative, positive, and mixed as three distinct features, I get a precision and recall (and F1) of 0.6190.

If I consider “positive sentiment” and “negative sentiment” as independent features, I get an F1 of 0.7636 for positive sentiment and 0.6939 for negative sentiment.

So that's not very good in an absolute sense, but it gives me a good baseline to compare other algorithms to.

## Next steps

- Other algorithms:
  - Gamon & Aue's minimally-supervised vocabulary builder (compare to Yarowski's word-sense method)
  - Nakagawa *et al* [2010] subtree analysis
  - N-gram HMM (adapted from POS tagger)? MEMM (OpenNLP)?
- Manually annotate more data
- Typo / spelling correction (MED)?
- Compare sentiment to:
  - Reviewer's rating
  - Reviewer's helpfulness score
- Integrate into Eli

Where do I go from here? The next step is certainly to do more algorithms. I identified a bunch of interesting ones in my proposal, and I won't get to all of them. I've also experimented with adapting my part-of-speech tagger into a sentiment detector (a Hidden Markov Model approach), and I'm wondering about using OpenNLP to build a max-ent sentiment detector.

I need to manually annotate more documents, for training and testing.

I originally thought I'd have to implement a spelling-correction filter. As it turns out, the reviews are cleaner than I expected, but it still might be useful.

I want to see if sentiment correlates to the numerical ratings that reviewers assign, and long-term we want to know if there's any correlation with the reviewer's helpfulness rating as calculated by Eli.

And finally, the ultimate goal is to integrate this system back into Eli so we can add all sorts of capabilities.