



GATE and UIMA in Language Technology Teaching

Graham Wilcock
University of Helsinki
graham.wilcock@helsinki.fi



Outline

- Current course materials
 - Shakespeare's Sonnets
 - GATE & ANNIE
- New course materials
 - UIMA & OpenNLP
 - UIMA & Stanford NLP
- Appendix
 - Eclipse & Stanford Eclipse

IBM: Tidwell XSLT tutorials

developerWorks : XML : Education - Online Courses - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://www.ling.helsinki.fi/kit/2006k/clk232/tutorials/XMLtoHTML.html

Project Gutenberg

IBM Search for: within All of dw Search Search help

IBM home Products & services Support & downloads My account

IBM: developerWorks : XML education

Part 1: Transforming XML into HTML

Doug Tidwell
Cyber Evangelist, developerWorks
Published: November 1999
Updated: November 2000

Part 1 of this [Transforming XML documents](#) tutorial shows you how to transform XML documents into HTML.

Let's meet our contestants
For our transformations, we'll use six source documents: a Shakespearean sonnet, a business letter, definitions for several technical terms, some spreadsheet data, a section from a technical manual, and a short section from Henry Fielding's 18th-century British novel *Tom Jones*. This will give us a wide range of document types to transform.

As we discuss our different transformations, some of these documents will be more relevant than others. For example, converting spreadsheet data into a pie chart would be useful to many people; converting a sonnet into a pie chart, while an intriguing exercise, probably has a more narrow appeal. We'll present a simple rendering of all our documents here. If you'd like to see the XML source for the documents and their DTDs, check out [Appendix A](#) - Sample XML documents.

A Shakespearean sonnet

This document is rendered line-by-line, with an indication of the rhyme scheme at the start of each line.

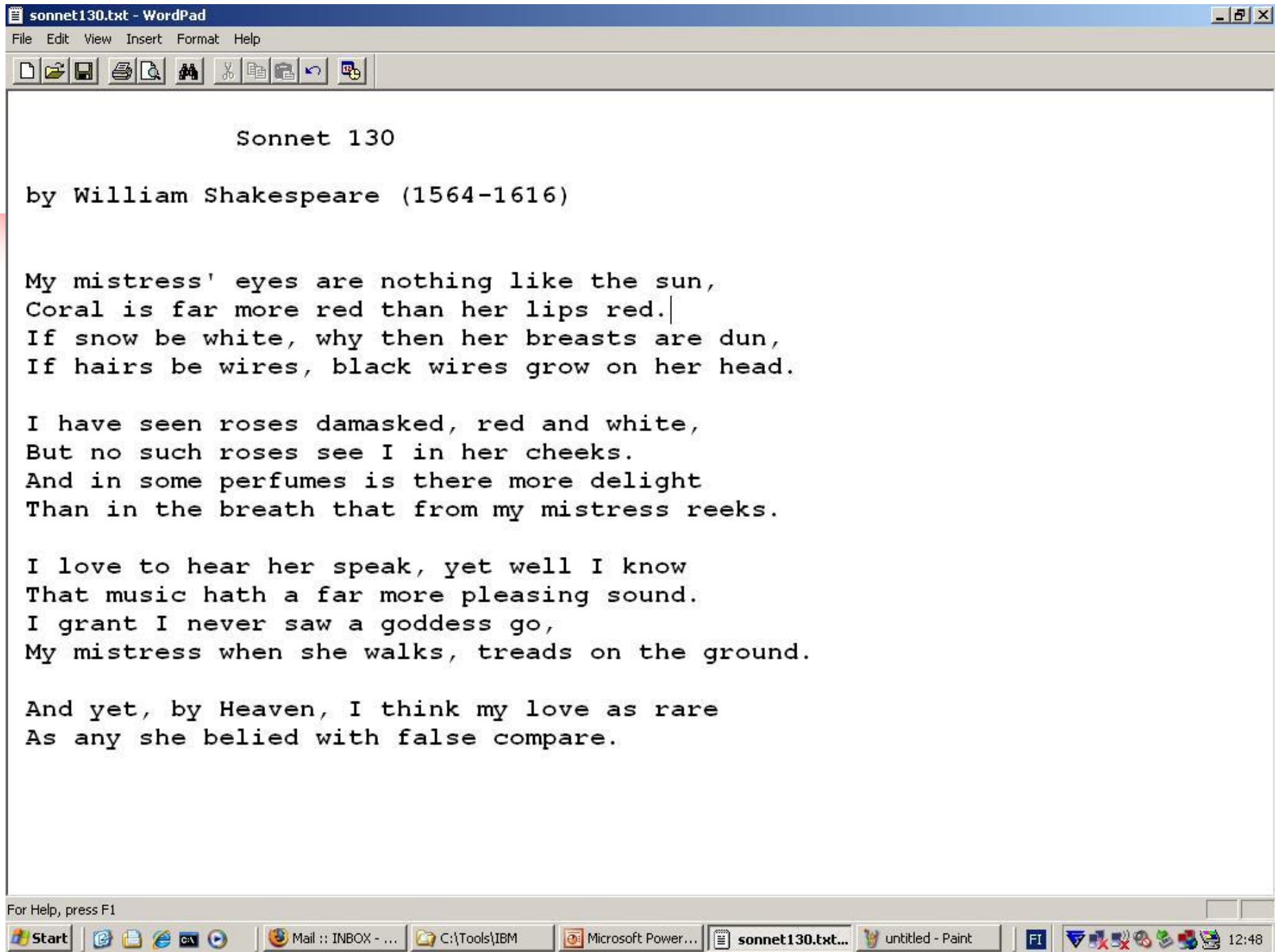
Sonnet 130

Author: William Shakespeare (*British, 1564-1616*)

A My mistress' eyes are nothing like the sun,
B Coral is far more red than her lips red.
A If snow be white, why then her breasts are dun,
B If hairs be wires, black wires grow on her head.
C I have seen roses damasked, red and white,
D But no such roses see I in her cheeks.
C And in some perfumes is there more delight

Done

Start C:\Documents and Se... Microsoft PowerPoint ... developerWorks : X... GutenbergSonnets.jp... FI 19:55



Gutenberg: Sonnets corpus

The screenshot shows a Mozilla Firefox browser window displaying the Project Gutenberg website. The address bar shows the URL <http://www.gutenberg.org/etext/1041>. The page title is "Shakespeare's Sonnets by William Shakespeare".

On the left side, there is a sidebar with the Project Gutenberg logo and navigation links: "Online Book Catalog", "Advanced Search", "Recent Books", "Top 100", "Offline Catalogs", "My Bookmarks", and "Main Page". There is also a "PayPal DONATE" button and a message: "Project Gutenberg needs your donation! [More Info](#)". At the bottom of the sidebar is the "hosted by ibiblio" logo.

The main content area features a banner for "Distributed Proofreaders" celebrating 10,000 titles. Below this is the title "Shakespeare's Sonnets by William Shakespeare" with links for "Help" and "Read online".

A "Bibliographic Record" table is displayed:

| Bibliographic Record | |
|-------------------------|---|
| Creator | Shakespeare, William, 1564-1616 |
| Title | Shakespeare's Sonnets |
| Language | English |
| LoC Class | PR: Language and Literatures: English literature |
| Subject | Poetry |
| EText-No. | 1041 |
| Release Date | 1997-09-01 |
| Copyright Status | Not copyrighted in the United States. If you live elsewhere check the laws of your country before downloading this ebook. |

Below the table, there is a section "Download this ebook for free" and a "Formats Available For Download" table:

| Edition | Format | Encoding | Compression | Size | Download Links |
|---------|--------|----------|-------------|------|----------------|
|---------|--------|----------|-------------|------|----------------|

The Windows taskbar at the bottom shows the Start button, several icons, and the system tray with the time 19:48.



Outline

- Current course materials
 - Shakespeare's Sonnets
 - GATE & ANNIE
- New course materials
 - UIMA & OpenNLP
 - UIMA & Stanford NLP
- Appendix
 - Eclipse & Stanford Eclipse



GATE & ANNIE

- Start with ANNIE
 - Ready-to-run NLP tools
 - ANNIE NE Recognizer
 - ANNIE POS Tagger
- Add JAPE annotations
 - Students write NP, PP rules



GATE

- Applications
 - ANNIE_00024
- Language Resources
 - Sonnets
 - Sonnet130
- Processing Resources
 - ANNIE OrthoMatcher...
 - ANNIE NE Transduce...
 - ANNIE POS Tagger_0...
 - ANNIE Sentence Split...
 - ANNIE Gazetteer_000...
 - ANNIE English Tokeni...
 - Document Reset PR...
- Data stores

Messages Sonnet130 ANNIE_00024

Annotation Sets Annotations Co-reference Editor Text

Sonnet 130

by William Shakespeare (1564-1616)

My mistress' eyes are nothing like the sun,
 Coral is far more red than her lips red.
 If snow be white, why then her breasts are dun,
 If hairs be wires, black wires grow on her head.

I have seen roses damasked, red and white,
 But no such roses see I in her cheeks.
 And in some perfumes is there more delight
 Than in the breath that from my mistress reeks.

I love to hear her speak, yet well I know
 That music hath a far more pleasing sound.
 I grant I never saw a goddess go,
 My mistress when she walks, treads on the ground.

And yet, by Heaven, I think my love as rare
 As any she belied with false compare.

- Date
- FirstPerson
- JobTitle
- Lookup
- Person
- Sentence
- SpaceToken
- Split
- Token
- Unknown
- ▼ Original markups
 - paragraph

New

MimeType text/p

docNewLineType LF

gate.SourceURL file:/C

Document Editor Initialisation Parameters

ANNIE_00024 run in 0,491 seconds



GATE & ANNIE

- Start with ANNIE
 - Ready-to-run NLP tools
 - ANNIE NE Recognizer
 - ANNIE POS Tagger
- Add JAPE annotations
 - Students write NP, PP rules



GATE

- Applications
 - ANNIE_00024
- Language Resources
 - Sonnets
 - Sonnet130
- Processing Resources
 - PP-JAPE
 - NP-JAPE
 - ANNIE OrthoMatcher...
 - ANNIE NE Transduce...
 - ANNIE POS Tagger_0...
 - ANNIE Sentence Split...
 - ANNIE Gazetteer_000...
 - ANNIE English Tokeni...
 - Document Reset PR...
- Data stores

Messages ANNIE_00024 Sonnet130

Annotation Sets Annotations Co-reference Editor Text

Sonnet 130

by William Shakespeare (1564-1616)

My mistress' eyes are nothing like the sun,
 Coral is far more red than her lips red.
 If snow be white, why then her breasts are dun,
 If hairs be wires, black wires grow on her head.

I have seen roses damasked, red and white,
 But no such roses see I in her cheeks.
 And in some perfumes is there more delight
 Than in the breath that from my mistress reeks.

I love to hear her speak, yet well I know
 That music hath a far more pleasing sound.
 I grant I never saw a goddess go,
 My mistress when she walks, treads on the ground.

And yet, by Heaven, I think my love as rare
 As any she

- Date
- FirstPerson
- JobTitle
- Lookup
- Person
- Sentence
- SpaceToken
- Split
- Token
- Unknown
- Original markups
 - paragraph

New

Token

| | | |
|----------|-----------|---|
| category | WBD | X |
| kind | word | X |
| length | 6 | X |
| orth | lowercase | X |
| string | belied | X |

MimeType text/p

docNewLineType LF

gate.SourceURL file:/C

Views built!



GATE & ANNIE

- Start with ANNIE
 - Ready-to-run NLP tools
 - ANNIE NE Recognizer
 - ANNIE POS Tagger
- Add JAPE annotations
 - Students write NP, PP rules



GATE

- Applications
 - ANNIE_00024
- Language Resources
 - Sonnets
 - Sonnet130
- Processing Resources
 - PP-JAPE
 - NP-JAPE
 - ANNIE OrthoMatcher...
 - ANNIE NE Transduce...
 - ANNIE POS Tagger_0...
 - ANNIE Sentence Split...
 - ANNIE Gazetteer_000...
 - ANNIE English Tokeni...
 - Document Reset PR...
- Data stores

Messages Sonnet130 ANNIE_00024

Annotation Sets Annotations Co-reference Editor Text

Sonnet 130

by William Shakespeare (1564-1616)

My mistress' eyes are nothing like the sun,
 Coral is far more red than her lips red.
 If snow be white, why then her breasts are dun,
 If hairs be wires, black wires grow on her head.

I have seen roses damasked, red and white,
 But no such roses see I in her cheeks.
 And in some perfumes is there more delight
 Than in the breath that from my mistress reeks.

I love to hear her speak, yet well I know
 That music hath a far more pleasing sound.
 I grant I never saw a goddess go,
 My mistress when she walks, treads on the ground.

And yet, by Heaven, I think my love as rare
 As any she belied with false compare.

- Date
- FirstPerson
- JobTitle
- Lookup
- NP
- PP
- Person
- Sentence
- SpaceToken
- Split
- Token
- Unknown
- Original markups
 - paragraph

New

C MimeType text/p

C docNewLineType LF

C gate.SourceURL file:/C

C

Document Editor Initialisation Parameters

ANNIE_00024 run in 0,28 seconds



Outline

- Current course materials
 - Shakespeare's Sonnets
 - GATE & ANNIE
- New course materials
 - UIMA & OpenNLP
 - UIMA & Stanford NLP
- Appendix
 - Eclipse & Stanford Eclipse



UIMA & OpenNLP

- OpenNLP tools
 - First, run from command line
 - Install in UIMA (assignment)
- UIMA & OpenNLP
 - OpenNLP POS Tagger
 - OpenNLP NE Recognizer



UIMA & OpenNLP

- OpenNLP tools
 - First, run from command line
 - Install in UIMA (assignment)
- UIMA & OpenNLP
 - OpenNLP POS Tagger
 - OpenNLP NE Recognizer

Document Analyzer
File Help

Unstructured Information Management Architecture

Input Directory: C:\Tools\IBM\workspace\uima_examples\data\sonnet
 Output Directory: C:\Tools\IBM\workspace\uima_examples\data\processed
 Location of Analysis Engine XML Descriptor: workspace\uima_examples\opennlp_wrappers\descriptors\OpenNLPAggregate.xml
 XML Tag containing Text (optional):
 Language: en
 Character Encoding: UTF-8

Run Interactive

siapi.jar - UIMA_HOME\lib - C:\Tool
 jVinci.jar - UIMA_HOME\lib\vinci - C
 resources
 maxent-2.4.0.jar - C:\Tools\NLP\op
 trove.jar - C:\Tools\NLP\opennlp-1
 opennlp-tools-1.3.0.jar - C:\Tools\l
 data
 deploy
 descriptors
 ecore_src
 opennlp_wrappers
 descriptors
 OpenNLPAggregate.xml
 OpenNLPNEDetector.xml
 OpenNLPParser.xml
 OpenNLPOSTagger.xml
 OpenNLPSentenceDetector.xml
 OpenNLPTokenAndSentenc
 OpenNLPTokenizer.xml

Outputs new
Name of the Java
Overview Aggregat
Problems Javadoc
UIMA Document Anal

Analysis Results
These are the Analyzed Documents.
Select viewer type and double-click file to open.
sonnet130.txt.xmi

Results Display Format: Java Viewer JV user colors HTML XML

Annotation Results for sonnet130.txt.xmi in C:\Tools\IBM\workspace\uima_examples\data\processed

Click In Text to See Annotation Detail

Annotations

- Token
 - Token ("belied")
 - begin = 645
 - end = 651
 - posTag = VBN
 - componentId = OpenNLP Tokenizer
 - Sentence
 - Sentence ("And yet, by Heaven, I think my l")
 - begin = 590
 - end = 673
 - componentId = OpenNLP Sentence Dete

Sonnet 130
by William Shakespeare (1564-1616)

My mistress' eyes are nothing like the sun,
Coral is far more red than her lips red,
If hairs be wires, black wires grow on her head.

I have seen roses damasked, red and white,
But no such roses see I in her cheeks.
And in some perfumes is there more delight
Than in the breath that from my mistress reeks.

I love to hear her speak, yet well I know
That music hath a far more pleasing sound.
I grant I never saw a goddess go,
My mistress when she walks, treads on the ground.

And yet, by Heaven, I think my love as rare
As any she belied with false compare.

Legend
 Docum... Sentence Token

Select All Deselect All Hide Unselected

Start Mail :: INBOX - Mozilla... Java - OpenNLPAggre... Document Analyzer untitled - Paint 16:43



UIMA & OpenNLP

- OpenNLP tools
 - First, run from command line
 - Install in UIMA (assignment)
- UIMA & OpenNLP
 - OpenNLP POS Tagger
 - OpenNLP NE Recognizer

Document Analyzer
File Help

Unstructured Information Management Architecture

Input Directory: C:\Tools\IBM\workspace\uima_examples\data\sonnet
 Output Directory: C:\Tools\IBM\workspace\uima_examples\data\processed
 Location of Analysis Engine XML Descriptor: \pace\uima_examples\opennlp_wrappers\descriptors\OpenNLPAggregateNER.x
 XML Tag containing Text (optional):
 Language: en
 Character Encoding: UTF-8

Results Display Format: Java Viewer JV user colors HTML XML

Run **Interactive**

- resources
- maxent-2.4.0.jar - C:\Tools\A
- trove.jar - C:\Tools\NLP\open
- opennlp-tools-1.3.0.jar - C:\T
- data
- deploy
- descriptors
- ecore_src
- opennlp_wrappers
 - descriptors
 - OpenNLPAggregate.x
 - OpenNLPAggregateNE
 - OpenNLPAggregateSt
 - OpenNLPNEDetector..
 - OpenNLPParser.xml
 - OpenNLPPOSTagger..
 - OpenNLPSentenceDet
 - OpenNLPTokenAndSe
 - OpenNLPTokenizer.xn
 - OpenNLPRadme.html

Annotation Results for sonnet130.txt.xml in C:\Tools\IBM\workspace\uima_examples\data\proce...

Sonnet 130

by **William Shakespeare** (1564-1616)

My mistress' eyes are nothing like the sun,
Coral is far more red than her lips red.
If snow be white, why then her breasts are dun,
If hairs be wires, black wires grow on her head.

I have seen roses damasked, red and white,
But no such roses see I in her cheeks.
And in some perfumes is there more delight
Than in the breath that from my mistress reeks.

I love to hear her speak, yet well I know
That music hath a far more pleasing sound.
I grant I never saw a goddess go,
My mistress when she walks, treads on the ground.

And yet, by Heaven, I think my love as rare
As any she belied with false compare.

Legend

Docu... Person Sente... Token

Click In Text to See Annotation Detail

Annotations

- Person
 - Person ("William Shakespeare")
 - begin = 29
 - end = 48
 - componentId = OpenNLP NE Detector

OpenNLPAggregateNER.xml - uima_examples/opennlp_wrappers/d

Select All **Deselect All** **Hide Unselected**

Windows taskbar: Start, C:\Tools\IBM, Java - OpenNLPAgr..., Document Analyzer, untyped - Paint, 19:44



UIMA & OpenNLP

- OpenNLP tools
 - First, run from command line
 - Install in UIMA (assignment)
- UIMA & OpenNLP
 - OpenNLP POS Tagger
 - OpenNLP NE Recognizer
- Add Java annotators
 - Students write NP, PP annotators?



Outline

- Current course materials
 - Shakespeare's Sonnets
 - GATE & ANNIE
- New course materials
 - UIMA & OpenNLP
 - UIMA & Stanford NLP
- Appendix
 - Eclipse & Stanford Eclipse

The logo consists of a vertical black line intersected by a horizontal black line. To the left of the intersection are three overlapping squares: a yellow one at the top, a red one in the middle, and a blue one at the bottom. The text "Stanford NLP Group" is positioned to the right of the logo.

Stanford NLP Group

- Stanford NLP tools
 - Stanford POS Tagger
 - Stanford NE Recognizer
 - Stanford Parser



UIMA & Stanford NLP

- Stanford NE Recognizer
 - Use Stanford NER-GUI
- UIMA & Stanford NER
 - Install in UIMA (assignment)
 - UIMA wrapper by F. Laws

Stanford Named Entity Recognizer

File Edit Classifier

Sonnet 130

by William Shakespeare (1564-1616)

My mistress' eyes are nothing like the sun,
 Coral is far more red than her lips red.
 If snow be white, why then her breasts are dun,
 If hairs be wires, black wires grow on her head.

I have seen roses damasked, red and white,
 But no such roses see I in her cheeks.
 And in some perfumes is there more delight
 Than in the breath that from my mistress reeks.

I love to hear her speak, yet well I know
 That music hath a far more pleasing sound.
 I grant I never saw a goddess go,
 My mistress when she walks, treads on the ground.

And yet, by Heaven, I think my love as rare
 As any she belied with false compare.

PERSON

ORGANIZATION

LOCATION

MISC

Run NER

```

c:\ Stanford NE Recognizer
C:\Tools\NLP\stanford\stanford-ner-2007-01-29>java -mx360m -cp "stanford-ner.jar
;" edu.stanford.nlp.ie.crf.NERGUI
Loading classifier from C:\Tools\NLP\stanford\stanford-ner-2007-01-29\classifier
s\ner-eng-ie.crf-4-conll.ser.gz ... done [3,9 sec].
content type: text/rtf
ORGANIZATION: Sonnet
PERSON: William Shakespeare
PERSON: Coral
PERSON: Heaven
Loading classifier from C:\Tools\NLP\stanford\stanford-ner-2007-01-29\classifier
s\ner-eng-ie.crf-3-all2006.ser.gz ... done [6,7 sec].
content type: text/rtf
PERSON: William Shakespeare
ORGANIZATION: Coral
Loading classifier from C:\Tools\NLP\stanford\stanford-ner-2007-01-29\classifier
s\ner-eng-ie.crf-4-conll-distsim.ser.gz ... done [56,5 sec].
content type: text/rtf
ORGANIZATION: Sonnet
PERSON: William Shakespeare
PERSON: Coral
LOCATION: Heaven
  
```





UIMA & Stanford NLP

- Stanford NE Recognizer
 - Use Stanford NER-GUI
- UIMA & Stanford NER
 - Install in UIMA (assignment)
 - UIMA wrapper by F. Laws

Document Analyzer

File Help

Unstructured Information Management Architecture

Input Directory: C:\Tools\IBM\workspace\uima_examples\data\sonnet

Output Directory: C:\Tools\IBM\workspace\uima_examples\data\processed

Location of Analysis Engine XML Descriptor: C:\Tools\IBM\workspace\StanfordNLP\de\florianlaws\uima\StanfordNERAnn

XML Tag containing Text (optional):

Language: en

Character Encoding: UTF-8

Run Interact

edu.stanford.nlp.maxent
 edu.stanford.nlp.objectbank
 edu.stanford.nlp.optimization
 edu.stanford.nlp.process
 edu.stanford.nlp.sequences
 edu.stanford.nlp.stats
 edu.stanford.nlp.util
 JRE System Library [jre1.5.0_11]
 uima_core.jar - C:\Tools\IBM\workspace\uima\lib
 uima_jcas_builtin_types.jar - C:\Tools\IBM\workspace\uima\lib
 uima_cpe.jar - C:\Tools\IBM\workspace\uima\lib
 uima_search.jar - C:\Tools\IBM\workspace\uima\lib
 uima_tools.jar - C:\Tools\IBM\workspace\uima\lib
 uima_examples.jar - C:\Tools\IBM\workspace\uima_examples\lib
 LICENSE.txt
 README.txt
 uima_examples

Overview Aggre
 Problems Javado
 UIMA Document A
 Loading cle
 getting doc
 analyzing..
 collecting
 ORGANIZATIO

StanfordNLP

Analysis Results

These are the Analyzed Documents.
 Select viewer type and double-click file to open.

sonnet130.txt.xmi

Results Display Format: Java Viewer JV user colors HTML XML

Annotation Results for sonnet130.txt.xmi in C:\Tools\IBM\workspace\uima_examples\data\processed

Click In Text to See Annotation Detail

Sonnet 130

by William Shakespeare (1564-1616)

My mistress' eyes are nothing like the sun,
 Coral is far more red than her lips red.
 If snow be white, why then her breasts are dun,
 If hairs be wires, black wires grow on her head.

I have seen roses damasked, red and white,
 But no such roses see I in her cheeks.
 And in some perfumes is there more delight
 Than in the breath that from my mistress reeks.

I love to hear her speak, yet well I know
 That music hath a far more pleasing sound.
 I grant I never saw a goddess go,
 My mistress when she walks, treads on the ground.

And yet, by Heaven, I think my love as rare
 As any she belied with false compare.

Annotations

- Location
 - Location ("Heaven")
 - begin = 602
 - end = 608
 - componentId = Stanford NER Detector

Legend

Docum... Location Organi... Person

Select All Deselect All Hide Unselected

Start C:\Tools\IBM Java - StanfordNERA... Document Analyzer untitled - Paint 0:18



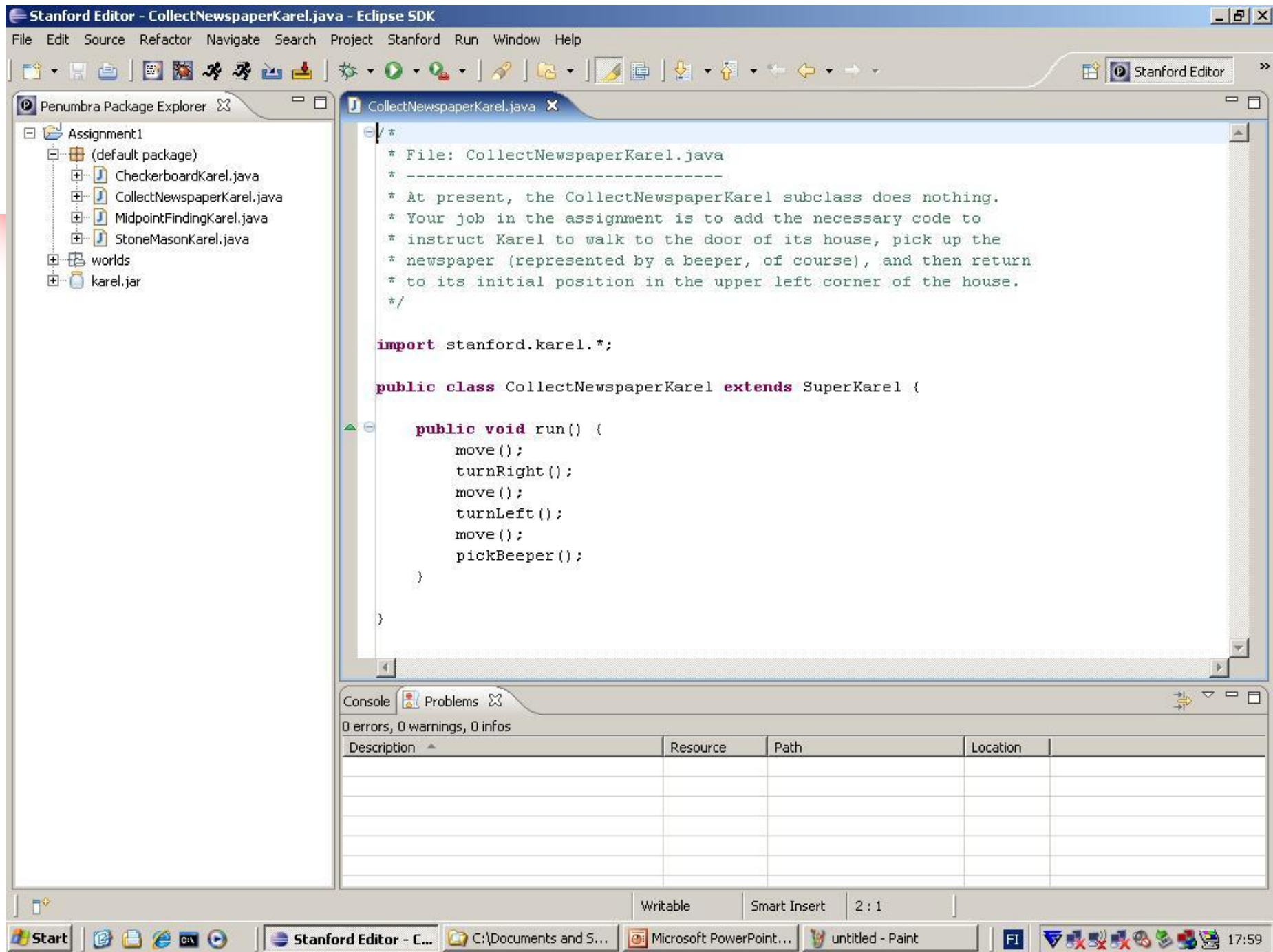
Outline

- Current course materials
 - Shakespeare's Sonnets
 - GATE & ANNIE
- New course materials
 - UIMA & OpenNLP
 - UIMA & Stanford NLP
- Appendix
 - Eclipse & Stanford Eclipse

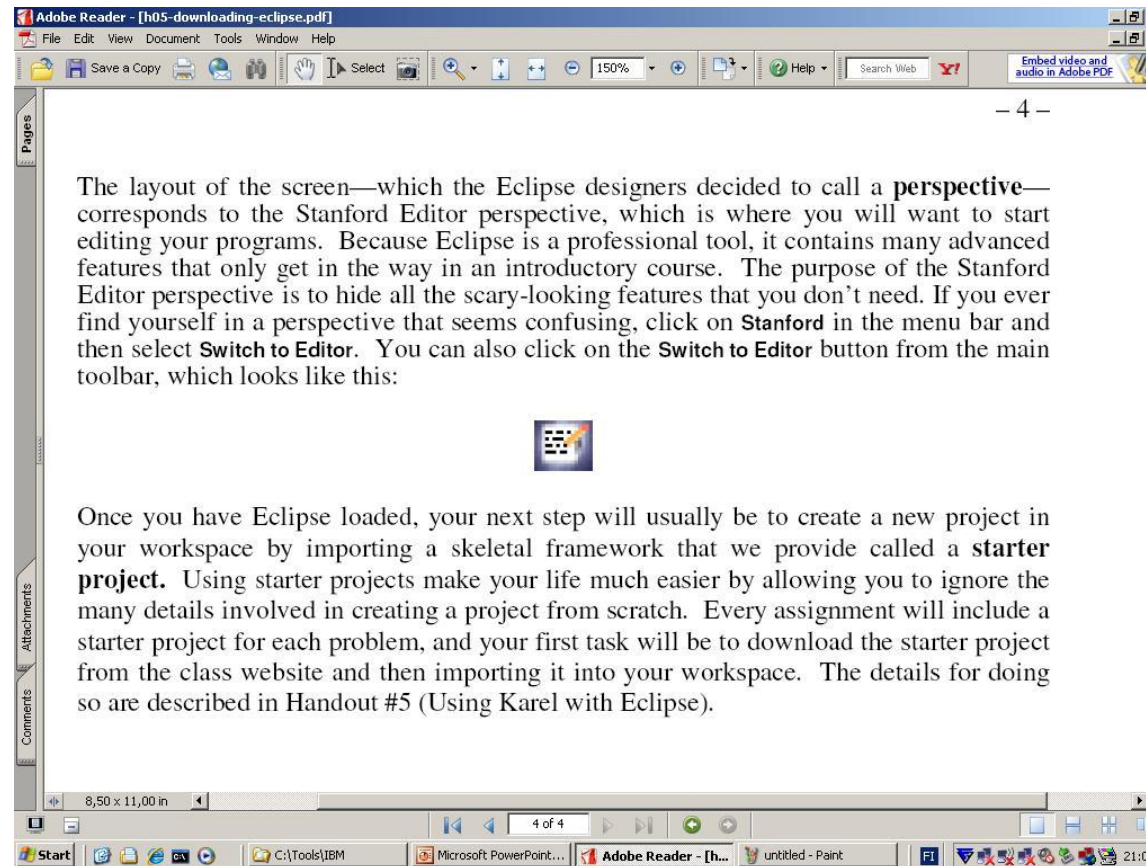


Eclipse & Stanford Eclipse

- jEdit vs. Eclipse
 - Students currently use jEdit
 - Eclipse learning curve is harder
- Stanford Eclipse
 - Stanford CS: simplified Eclipse
 - *Karel the Robot Learns Java*
 - Starter projects for assignments



Starter projects



Document Analyzer

File Help

Unstructured Information Management Architecture

An Apache Incubator Project

Input Directory: C:\Tools\Stanford\eclipse\uima_workspace\uimaj-examples\data\sonnet

Output Directory: C:\Tools\Stanford\eclipse\uima_workspace\uimaj-examples\data\sonnet\proc

Location of Analysis Engine XML Descriptor: C:\Tools\Apache\uima-2.1.0\examples\opennlp_wrappers\descriptors\OpenNLP

XML Tag containing Text (optional):

Language: en

Character Encoding: UTF-8

Run I

- trove.jar - C:\Tools\NLP\opennlp-1.3.0
- opennlp-tools-1.3.0.jar - C:\Tools\NLP
- data
- deploy
- descriptors
- ecore_src
- opennlp_wrappers
 - descriptors
 - OpenNLPAggregate.xml
 - OpenNLPNEDetector.xml
 - OpenNLPParser.xml
 - OpenNLPOSTagger.xml
 - OpenNLPSentenceDetector.xml
 - OpenNLPTokenAndSentenceA...
 - OpenNLPTokenizer.xml
 - OpenNLPReadme.html
- run_configuration

Analysis Results

These are the Analyzed Documents.
Select viewer type and double-click file to open.

- sonnet130.txt.xmi

Results Display Format: Java Viewer JV user colors HTML XML

Annotation Results for sonnet130.txt.xmi in C:\Tools\Stanford\eclipse\uima_workspace\uimaj-examples\data\s...

Click In Text to See Annotation Detail

Annotations

- Token
 - Token ("belied")
 - begin = 645
 - end = 651
 - posTag = VBN
 - componentId = OpenNLP Tokenizer
- Sentence
 - Sentence ("And yet, by Heaven, I think my love as
 - begin = 590
 - end = 673
 - componentId = OpenNLP Sentence Detector

Legend

- Docum...
- Sentence
- Source...
- Token

Annotation Results for sonnet130.txt.xmi in C:\Tools\Stanford\eclipse\uima_workspace\uimaj-examples\data\s...

Sonnet 130

by William Shakespeare (1564-1616)

My mistress' eyes are nothing like the sun,
Coral is far more red than her lips red.
If snow be white, why then her breasts are dun,
If hairs be wires, black wires grow on her head.

I have seen roses damasked, red and white,
But no such roses see I in her cheeks.
And in some perfumes is there more delight
Than in the breath that from my mistress reeks.

I love to hear her speak, yet well I know
That music hath a far more pleasing sound.
I grant I never saw a goddess go,
My mistress when she walks, treads on the ground.

And yet, by Heaven, I think my love as rare
As any she belied with false compare.

Overview

Problem: 0 errors, 42

Description

Wa

Start

C:\Tools\Stanford\ec... Apache UIMA OpenN... Java - OpenNLPSent...

Document Analyzer F1 18:17

Document Analyzer

File Help

Unstructured Information Management Architecture
An Apache Incubator Project

Input Directory: C:\Tools\Stanford\eclipse\uima_workspace\uimaj-examples\data\sonnet

Output Directory: C:\Tools\Stanford\eclipse\uima_workspace\uimaj-examples\data\sonnet\proc

Location of Analysis Engine XML Descriptor: C:\Tools\Stanford\eclipse\uima_workspace\StanfordNLP\de\florianlaws\uima\...

XML Tag containing Text (optional):

Language: en

Character Encoding: UTF-8

Run Interact

- edu.stanford.nlp.maxent
- edu.stanford.nlp.objectbank
- edu.stanford.nlp.optimization
- edu.stanford.nlp.process
- edu.stanford.nlp.sequences
- edu.stanford.nlp.stats
- edu.stanford.nlp.util
- JRE System Library [jre1.5.0_11]
- uima-core.jar - C:\Tools\Apache\...
- uima-document-annotation.jar - C:\Tools\...
- uima-cpe.jar - C:\Tools\Apache\...
- uima-examples.jar - C:\Tools\Apache\...
- uima-tools.jar - C:\Tools\Apache\...
- uima-adapter-vinci.jar - C:\Tools\Apache\...
- uima-adapter-soap.jar - C:\Tools\Apache\...
- LICENSE.txt
- README.txt
- uimaj-examples

```
<type>Stri
<multiValu
<mandatory
</configur
</configur
<configura
<nameValue
<name>Clas
<value>
```

Analysis Results

These are the Analyzed Documents.
Select viewer type and double-click file to open.

- sonnet130.txt.xmi

Results Display Format: Java Viewer JV user colors HTML XML

Annotation Results for sonnet130.txt.xmi in C:\Tools\Stanford\...

Sonnet 130

by William Shakespeare (1564-1616)

My mistress' eyes are nothing like the sun,
Coral is far more red than her lips red.
If snow be white, why then her breasts are dun,
If hairs be wires, black wires grow on her head.

I have seen roses damasked, red and white,
But no such roses see I in her cheeks.
And in some perfumes is there more delight
Than in the breath that from my mistress reeks.

I love to hear her speak, yet well I know
That music hath a far more pleasing sound.
I grant I never saw a goddess go,
My mistress when she walks, treads on the ground.

And yet, by Heaven, I think my love as rare
As any she belied with false compare.

Click In Text to See Annotation Detail

- Annotations
 - Location
 - Location ("Heaven")
 - begin = 602
 - end = 608
 - componentId = Stanford NER Detector

Legend

Docum...
 Location
 Organiz...
 Person

Select All Deselect All Hide Unselected

de.florianlaws.uima - StanfordNLP



Summary

- Current course materials
 - Shakespeare's Sonnets
 - GATE & ANNIE
- New course materials
 - UIMA & OpenNLP
 - UIMA & Stanford NLP
- Appendix
 - Eclipse & Stanford Eclipse