Cloud Computing for the Humanities

Graham Wilcock
University of Helsinki
What is Cloud Computing?

- "Run your app in the cloud"
  - Using somebody else's computers
- Computing resources on-demand
  - Like electricity, or pizza delivery
- Platform-as-a-Service (PaaS)
  - Example: Google App Engine
Google App Engine

- "Run your web apps on Google’s infrastructure"
  - http://your-app-name.appspot.com

- My web app is AELRED:
  - App Engine Language Resource Editions
  - First version: Jane Austen novels
  - http://aelred-austen.appspot.com
App Engine Language Resource Editions

This web application demonstrates the use of natural language processing in a cloud computing environment. AELRED serves up English literary texts with a range of optional linguistic annotations including word frequencies, concordances, part-of-speech tagging, shallow parsing, and word sense definitions.

AELRED runs on the Google App Engine cloud computing framework. The raw texts (prior to being annotated) are the plain texts made available by Project Gutenberg. The annotations are created automatically by natural language processing techniques. Where possible, existing tools from NLTK Natural Language Toolkit are used. Word sense definitions are from WordNet.

In this initial demonstration version, the texts are the six completed novels of Jane Austen. The total number of words is about half a million. Click on the Jane Austen button below to continue.

Graham Wilcock

Available Authors

Jane Austen
Jane Austen

Available Books

- Sense and Sensibility (Published 1811)
- Pride and Prejudice (Published 1813)
- Mansfield Park (Published 1814)
- Emma (Published 1816)
- Persuasion (Published 1818)
- Northanger Abbey (Published 1818)
App Engine Language Resource Editions

Jane Austen: Northanger Abbey

List of Chapters

Chapter 1  No one who had ever seen Catherine Morland in her infancy would have supposed her born to be an heroine. Her situation in life, the character of...

Chapter 2  In addition to what has been already said of Catherine Morland's personal and mental endowments, when about to be launched into all the difficult...

Chapter 3  Every morning now brought its regular duties--shops were to be visited; some new part of the town to be looked at; and the pump-room to be attende...

Chapter 4  With more than usual eagerness did Catherine hasten to the pump-room the next day, secure within herself of seeing Mr. Tilney there before the mor...

Chapter 5  Catherine was not so much engaged at the theatre that evening, in returning the nods and smiles of Miss Thorpe, though they certainly claimed much...

Chapter 6  The following conversation, which took place between the two friends in the pump-room one morning, after an acquaintance of eight or nine days, is...

Chapter 7  Half a minute conducted them through the pump-yard to the archway, opposite Union Passage; but here they were stopped. Everybody acquainted with ...
App Engine Language Resource Editions

Jane Austen: Northanger Abbey: Chapter 1

- Tokenized Text  - Part-of-Speech Tags  - Phrase Chunks
- Word Frequencies  - Concordances  - WordNet Definitions

CHAPTER 1

No one who had ever seen Catherine Morland in her infancy would have supposed her born to be an heroine. Her situation in life, the character of her father and mother, her own person and disposition, were all equally against her. Her father was a clergyman, without being neglected, or poor, and a very respectable man, though his name was Richard -- and he had never been handsome. He had a considerable independence besides two good livings -- and he was not in the least addicted to locking up his daughters. Her mother was a woman of useful plain sense, with a good temper, and, what is more remarkable, with a good constitution. She had three sons before Catherine was born; and instead of dying in bringing the latter into the world, as anybody might expect, she still lived on -- lived to have six children more -- to see them growing up around her, and to enjoy excellent health herself. A family of ten children will be always called a fine family, where there are heads and arms and legs enough for the number; but the Morlands had little other right to the word, for they were in general very plain, and Catherine, for many years of her life, as plain as any. She had a thin awkward figure, a sallow skin without colour, dark lank hair, and strong features -- so much for her person; and not less unpromising for heroism seemed her mind. She was fond of all boy's plays, and greatly preferred cricket not merely to dolls, but to the more heroic enjoyments of infancy, nursing a dormouse, feeding a canary-bird, or
App Engine Language Resource Editions

Jane Austen: Northanger Abbey: Chapter 1

- Tokenized Text  - Part-of-Speech Tags  - Phrase Chunks
- Word Frequencies  - Concordances  - WordNet Definitions

CHAPTER/INN 1/CD

No/DT one/NN who/WP had/VBD ever/RB seen/VBN Catherine/NNP Morland/NNP in/IN her/PRP$ infancy/NN would/MD have/VB supposed/VBN her/PRP$ born/NN to/TO be/VB an/DT heroine/NN ./f. Her/PRP$ situation/NN in/IN life/NN ./f, the/DT character/INN of/IN her/PRP$ father/NN and/CC mother/NN ./f, her/PRP$ own/JJ person/INN and/CC disposition/INN ./f, were/VBD all/DT equally/RB against/IN her/PRP$ ./f. Her/PRP$ father/NN was/VBD a/DT clergyman/NN ./f, without/IN being/VBG neglected/VBN ./f, or/CC poor/JJ ./f, and/CC a/DT very/RB respectable/JJ man/INN ./f, though/IN his/PRP$ name/INN was/VBD Richard/NNP --: and/CC he/PRP had/VBD never/RB been/VBN handsome/JJ ./f. He/PRP had/VBD a/DT considerable/JJ independence/INN besides/VBZ two/CD good/JJ livings/NNS --: and/CC he/PRP was/VBD not/RB in/IN the/DT least/JJS addicted/VBN to/TO locking/NIN up/IN his/PRP$ daughters/NNS ./f. Her/PRP$ mother/INN was/VBD a/DT woman/IN of/IN useful/JJ plain/INN sense/INN ./f, with/IN a/DT good/JJ temper/INN ./f, and/CC ./f, what/WP is/VBZ more/RBR remarkable/JJ ./f, with/IN a/DT good/JJ constitution/INN ./f. She/PRP had/VBD three/CD sons/NNS before/IN Catherine/NNP was/VBD born/VBN ./f: and/CC instead/RB of/IN dying/VBG in/IN bringing/NN the/DT latter/NN into/IN the/DT world/NN ./f, as/IN anybody/NN might/MD expect/VB ./f, she/PRP still/RB lived/VBD on/IN ./f: lived/VBN to/TO have/VB six/CD children/NNS more/JJR --: to/TO see/VB them/PRP growing/VBG up/RP around/IN

Graham Wilcock  Baltic HLT, Riga, 2010
Jane Austen: Northanger Abbey: Chapter 1

No/DT one/NN: Noun, singular or mass →/RB seen/VBN Catherine/NNP Morland/NNP in/IN her/PRP$: infancy/NN would/MD have/VB supposed/VBN her/PRP born/NN to/TO be/VB an/DT heroine/NN ./., Her/PRP$: situation/NN in/IN life/NN ./, the/DT character/NN of/IN her/PRP$: father/NN and/CC mother/NN ./, her/PRP$: own/JJ person/NN and/CC disposition/NN ./, were/VBD all/DT equally/RB against/IN her/PRP$:./. Her/PRP$: father/NN was/VBD a/DT clergyman/NN ./, without/IN being/VBG neglected/VBN ./, or/CC poor/JJ ./, and/CC a/DT very/RB respectable/JJ man/NN ./, though/IN his/PRP$: name/NN was/VBD Richard/NNP --.; and/CC he/PRP had/VBD ever/RB been/VBN handsome/JJ ./, He/PRP had/VBD a/DT considerable/JJ independence/NN besides/VBZ two/CD good/JJ livings/NNS --.; and/CC he/PRP was/VBD not/RB in/IN the/DT least/JJS addicted/VBN to/TO locking/NN up/IN his/PRP$: daughters/NNS ./, Her/PRP$: mother/NN was/VBD a/DT woman/NN of/IN useful/JJ plain/NN sense/NN ./, with/IN a/DT good/JJ temper/NN ./, and/CC ./, what/VP is/VBZ more/RBR remarkable/JJ ./, with/IN a/DT good/JJ constitution/NN ./, She/PRP had/VBD three/CD sons/HNS before/IN Catherine/NNP was/VBD born/VBN ./; and/CC instead/RR of/IN dying/VBG in/IN bringing/NN the/DT latter/NN into/IN the/DT world/NN ./, as/IN anybody/NN might/MD expect/VB ./, she/PRP still/RR lived/VBD on/IN --.; lived/VBN to/TO have/VB six/CD children/HNS more/JJS --.; to/TO see/VB them/PRP growing/VBG up/RR around/IN
App Engine Language Resource Editions

Jane Austen: Northanger Abbey: Chapter 1

- Tokenized Text
- Part-of-Speech Tags
- Phrase Chunks
- Word Frequencies
- Concordances
- WordNet Definitions

CHAPTER 1

No one who had ever seen Catherine Morland in her infancy would have supposed her born to be an heroine. Her situation in life, the character of her father and mother, her own person and disposition, were all equally against her. Her father was a clergyman, without being neglected, or poor, and a very respectable man, though his name was Richard -- and he had never been handsome. He had a considerable independence besides two good livings -- and he was not in the least addicted to locking up his daughters. Her mother was a woman of useful plain sense, with a good temper, and, what is more remarkable, with a good constitution. She had three sons before Catherine was born; and instead of dying in bringing the latter into the world, as anybody might expect, she still lived on -- lived to have six children more -- to see them growing up around her, and to enjoy excellent health herself. A family of ten children will be always called a fine family, where there are heads and arms and legs enough for the number; but the Morlands had little other right to the word, for they were in general very plain, and Catherine, for many years of her life, as plain as any. She had a thin awkward figure, a sallow skin without colour, dark lank hair, and strong features -- so much for her person; and not less unpropitious for heroism seemed her mind. She was fond of all boy's plays, and greatly preferred cricket not merely to dolls, but to the more heroic enjoyments of infancy, nursing a dormouse, feeding a canary-bird, or...
App Engine Language Resource Editions

Jane Austen: Northanger Abbey: Chapter 1

Catherine Morland

caught 1 censure 1 chance 1 character 2 chickens 1 chief 1 chiefly 1 clean 1 cleanliness 1 clergyman 1 colour 1 complexion 1 compliance 1 composition 1 confinement 1 confirmation 1 conjectured 1 consequence 1 considerable 1 constitution 2 corporal 1 country 1 cradle 1 cricket 2 curl 1 dark 1 deficiency 1 delight 1 delightful 1 desert 1 design 1 detected 1 dies 1 dirt 1 dismissed 1 disposition 1 distaste 1 dolls 1 door 1 dormouse 1 drawing 3 dying 1 ears 1 elder 1 enjoy 1 enjoyments 1 equally 1 eventful 1 exceedingly 1 excellent 1 excited 1 expect 1 extraordinary 1 eyes 1 fable 1 fairly 1 families 1 family 9 father 1 fatigue 1 features 2 feeding 1 feels 1 fell 1 fifteen 3 figure 2 fine 1 finery 1 flower 1 flowers 1 forbidden 1 forlorn 1 forty 1 found 1 fourteen 1 fragrance 1 gained 3 garden 1 gathered 1 gave 1 general 1 generally 1 giant 1 girl 1 good 1 good-humoured 1 good-looking 1 gouty 1 great
App Engine Language Resource Editions

Jane Austen: Northanger Abbey: Chapter 1

CHAPTER 1

No one who had ever seen Catherine Morland in her infancy would have supposed her born to be an heroine. Her situation in life, the character of her father and mother, her own person and disposition, were all equally against her. Her father was a clergyman, without being neglected, or poor, and a very respectable man, though his name was Richard -- and he had never been handsome. He had a considerable independence besides two good livings -- and he was not in the least addicted to locking up his daughters. Her mother was a woman of useful plain sense, with a good temper, and, what is more remarkable, with a good constitution. She had three sons before Catherine was born; and instead of dying in bringing the latter into the world, as anybody might expect, she still lived on -- lived to have six children more -- to see them growing up around her, and to enjoy excellent health herself. A family of ten children will be always called a fine family, where there are heads and arms and legs enough for the number; but the Morlands had little other right to the word, for they were in general very plain, and Catherine, for many years of her life, as plain as any. She had a thin awkward figure, a sallow skin without colour, dark lank hair, and strong features -- so much for her person; and not less unprouitous for heroism seemed her mind. She was fond of all boy's plays, and greatly preferred cricket not merely to dolls, but to the more heroic enjoyments of infancy, nursing a dormouse, feeding a canary-bird, or
### Jane Austen: Northanger Abbey

<table>
<thead>
<tr>
<th>Left Context</th>
<th>Word</th>
<th>Right Context</th>
</tr>
</thead>
<tbody>
<tr>
<td>uld have supposed her born to be an</td>
<td>heroine</td>
<td>...</td>
</tr>
<tr>
<td>seventeen she was in training for a</td>
<td>heroine</td>
<td>...</td>
</tr>
<tr>
<td>. But when a young lady is to be a</td>
<td>heroine</td>
<td>...</td>
</tr>
<tr>
<td>ons which the first separation of a</td>
<td>heroine</td>
<td>...</td>
</tr>
<tr>
<td>ess delight in being fine ; and our</td>
<td>heroine</td>
<td>...</td>
</tr>
<tr>
<td>omfort ; and now was the time for a</td>
<td>heroine</td>
<td>...</td>
</tr>
<tr>
<td>s simple praise than a true-quality</td>
<td>heroine</td>
<td>...</td>
</tr>
<tr>
<td>fortune was more favourable to our</td>
<td>heroine</td>
<td>...</td>
</tr>
<tr>
<td>itting them to be read by their own</td>
<td>heroine</td>
<td>...</td>
</tr>
<tr>
<td>id pages with disgust . Alas If the</td>
<td>heroine</td>
<td>...</td>
</tr>
<tr>
<td>one novel be not patronized by t</td>
<td>heroine</td>
<td>...</td>
</tr>
<tr>
<td>nces which peculiarly belong to the</td>
<td>heroine</td>
<td>...</td>
</tr>
<tr>
<td>Every young lady may feel for my</td>
<td>heroine</td>
<td>...</td>
</tr>
</tbody>
</table>
Jane Austen: Northanger Abbey: Chapter 1

CHAPTER 1

No one who had ever seen Catherine Morland in her infancy would have supposed her born to be an heroine. Her situation in life, the character of her father and mother, her own person and disposition, were all equally against her. Her father was a clergyman, without being neglected, or poor, and a very respectable man, though his name was Richard -- and he had never been handsome. He had a considerable independence besides two good livings -- and he was not in the least addicted to locking up his daughters. Her mother was a woman of useful plain sense, with a good temper, and, what is more remarkable, with a good constitution. She had three sons before Catherine was born; and instead of dying in bringing the latter into the world, as anybody might expect, she still lived on -- lived to have six children more -- to see them growing up around her, and to enjoy excellent health herself. A family of ten children will be always called a fine family, where there are heads and arms and legs enough for the number; but the Morlands had little other right to the word, for they were in general very plain, and Catherine, for many years of her life, as plain as any. She had a thin awkward figure, a sallow skin without colour, dark lank hair, and strong features -- so much for her person; and not less unpropitious for heroism seemed her mind. She was fond of all boy's plays, and greatly preferred cricket not merely to dolls, but to the more heroic enjoyments of infancy, nursing a dormouse, feeding a canary-bird, or...
Jane Austen: Northanger Abbey: Chapter 1

CHAPTER 1

No one who had ever seen Catherine Morland in her infancy would have supposed her born to be an heroine. Her situation in life, the character of her father and mother, her own person and dispositions were all equally against her. It is true that she was an clergyman's daughter, without being neglected; she was well educated, and she had a very respectable income; though his name was Richard -- and he had never been handsome. He had a considerable independence besides two good livings -- and he was not in the least addicted to locking up his daughters. Her mother was a woman of useful plain sense, with a good temper, and, what is more remarkable, with a good constitution. She had three sons before Catherine was born; and instead of dying in bringing the latter into the world, as anybody might expect, she still lived on -- lived to have six children more -- to see them growing up around her, and to enjoy excellent health herself. A family of ten children will be always called a fine family, where there are heads and arms and legs enough for the number; but the Morlands had little other right to the word, for they were in general very plain, and Catherine, for many years of her life, as plain as any. She had a thin awkward figure, a sallow skin without colour, dark lank hair, and strong features -- so much for her person; and not less unpropitious for heroism seemed her mind. She was fond of all boy's plays, and greatly preferred cricket not merely to dolls, but to the more heroic enjoyments of infancy, nursing a dormouse, feeding a canary-bird, or...
Key Ideas: Easy, Big, Free

- Easy: use Python
  - NLTK Natural Language Toolkit
  - Django HTML Template Engine
- Big: Google’s scalable infrastructure
  - BigTable non-relational datastore
  - MapReduce data-intensive processing
- Free: App Engine has free quotas
  - Only pay if high demand for app
NLTK Natural Language Toolkit

- Open source Python tools
  - Taggers, chunkers, parsers, classifiers ...
- Many major corpora and resources
  - Brown Corpus, Penn Treebank, WordNet ...
- Excellent free online textbook
  - *Natural Language Processing with Python*
  - Stephen Bird, Ewan Klein, Edward Loper
App Engine code must be pure Python

Normal "import nltk" does not work

- Some NLTK code is not pure Python
- E.g. uses Numpy with C for speed

Use "import aelred" instead

- Aelred code is pure Python
- Other customization, e.g. tokenization
Meet Django

Django is a high-level Python Web framework that encourages rapid development and clean, pragmatic design.

Developed four years ago by a fast-moving online-news operation, Django was designed to handle two challenges: the intensive deadlines of a newsroom and the stringent requirements of the experienced Web developers who wrote it. It lets you build high-performing, elegant Web applications quickly.

Django focuses on automating as much as possible and adhering to the DRY principle.

Dive in by reading the overview and tutorial.

When you're ready to code, read the installation guide and tutorial.

The Django framework

Object-relational mapper
Django Web App Framework

- Open source Python
  - Model-View-Controller design pattern
  - Models defined easily by Python classes
- HTML Template Engine
  - Web pages generated using contexts
  - Excellent ”template inheritance” facility
- Free online textbook
  - Django: The Book
Google BigTable Datastore

- Non-relational database
  - Different thinking from SQL databases
  - Designed for massive scalability
- My current way of using the datastore:
  - Serialize complex objects to YAML
  - Store/retrieve YAML as big text strings
MapReduce Algorithms

- Data-intensive distributed processing
  - Different thinking from usual algorithms
  - Designed for massive scalability
- My current way of using MapReduce:
  - Iterate over all entities in datastore
  - Delete entity, or update and save
MapReduce: Simplified Data Processing on Large Clusters
Jeffrey Dean and Sanjay Ghemawat

Abstract

MapReduce is a programming model and an associated implementation for processing and generating large data sets. Users specify a map function that processes a key/value pair to generate a set of intermediate key/value pairs, and a reduce function that merges all intermediate values associated with the same intermediate key. Many real world tasks are expressible in this model, as shown in the paper.

Programs written in this functional style are automatically parallelized and executed on a large cluster of commodity machines. The run-time system takes care of the details of partitioning the input data, scheduling the program's execution across a set of machines, handling machine failures, and managing the required inter-machine communication. This allows programmers without any experience with parallel and distributed systems to easily utilize the resources of a large distributed system.

Our implementation of MapReduce runs on a large cluster of commodity machines and is highly scalable: a typical MapReduce computation processes many terabytes of data on thousands of machines. Programmers find the system easy to use: hundreds of MapReduce programs have been implemented and upwards of one thousand MapReduce jobs are executed on Google's clusters every day.

Appeared at:
OSDI'04: Sixth Symposium on Operating System Design and Implementation,

Download: PDF Version
Slides: HTML Slides
Overview

AppEngine-MapReduce is an open-source library for doing MapReduce-style computations on the Google App Engine platform.

Status

This is an early experimental release of mapper implementation.

As of today, it contains a mapper implementation that provides a fast and efficient way to iterate over datastore entities and blob files. In the future, we will expand the library to provide full MapReduce support, including the shuffle and reduce stages.

We have both Python and Java version available under the Source tab.

Check the changelog page for recent changes in the library.

Getting Started

- Check out the Getting Started documents for Python and Java. If you have experience with Hadoop, you may also be interested in our transition guide for Hadoop programmers.

Screenshots

Mapper library contains a UI for managing and overseeing mapper execution.