

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

# Constructing a Knowledge Base on Aging

## An Automated Approach

Mark Farrell

Bioinformatics Researcher

Center for Research and Education on Aging  
Lawrence Berkeley National Laboratory  
University of California, Berkeley

September 4th, 2014

# Outline

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

## 1 About Me

## 2 Understanding Aging

## 3 Knowledge Extraction

## 4 What Next?

# Who I am

## Mark Farrell

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?



# Where I'm from

## Atlantic Canada

Constructing a  
Knowledge  
Base on Aging

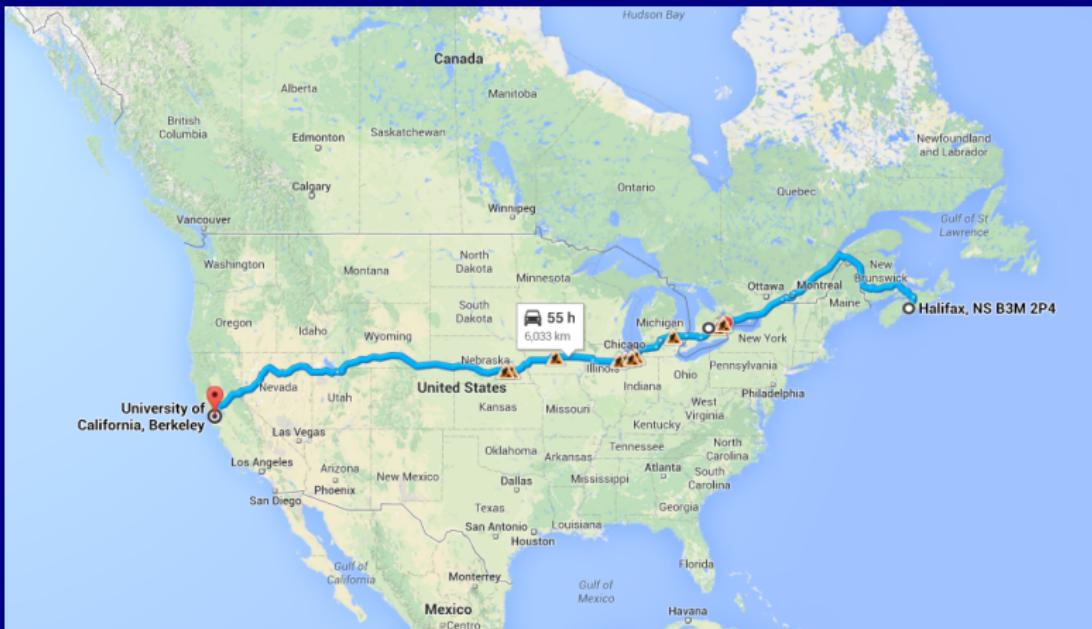
Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?



# What I'm Studying

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

Year 2

Program Bachelor of Computer Science

Faculty Mathematics

Institution University of Waterloo

Location Waterloo, Ontario, Canada



# How I Became a Researcher at CREA

Prof. Garan Emailed the University of Waterloo Computer Science Club

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?



# Why was I Interested in Research at CREA?

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

Hobby	CREA Research
Reverse Engineering	Natural Language Processing
Game Development	Artificial Intelligence
Health	Bioinformatics

# Outline

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

## 1 About Me

## 2 Understanding Aging

## 3 Knowledge Extraction

## 4 What Next?

# What does CREA want to do?

CREA Wants to Understand Aging and Increase Human Lifespan

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?



# Why is it Difficult to Understand Aging?

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

NCBI Resources How To Sign in to NCBI

PubMed aging RSS Save search Advanced Search Help

Show additional filters Display Settings: Summary, 20 per page, Sorted by Recently Added Send to: Filters: Manage Filters

Results: 1 to 20 of 306272 <> First <- Prev Page 1 of 15314 Next > Last >>

[Proteostasis: Bad news and good news from the endoplasmic reticulum.](#)  
1. Noack J, Brambilla Pisoni G, Molinari M.  
Swiss Med Wkly. 2014 Aug 21;144:w14001. doi: 10.4414/smw.2014.14001. eCollection 2014.  
PMID: 25144910 [PubMed - as supplied by publisher] [Free Article](#)  
[Related citations](#)

[The Experiences of Making Infant Feeding Choices by African, Caribbean and Black HIV-Positive Mothers in Ontario, Canada.](#)  
2. Kapiriri L, Tharao WE, Muchenje M, Masinde KI, Siegel S, Ongoba F.  
World Health Popul. 2014;15(2):14-22.  
PMID: 25144786 [PubMed - as supplied by publisher] [Related citations](#)

[Famous Landmark Identification in Amnestic Mild Cognitive Impairment and Alzheimer's Disease.](#)  
3. Sheardova K, Laczó J, Vyhalek M, Andel R, Mokrisova I, Vicek K, Amlerova J, Hrt J.  
PLoS One. 2014 Aug 21;9(8):e105623. doi: 10.1371/journal.pone.0105623. eCollection 2014.  
PMID: 25144755 [PubMed - as supplied by publisher] [Free Article](#)  
[Related citations](#)

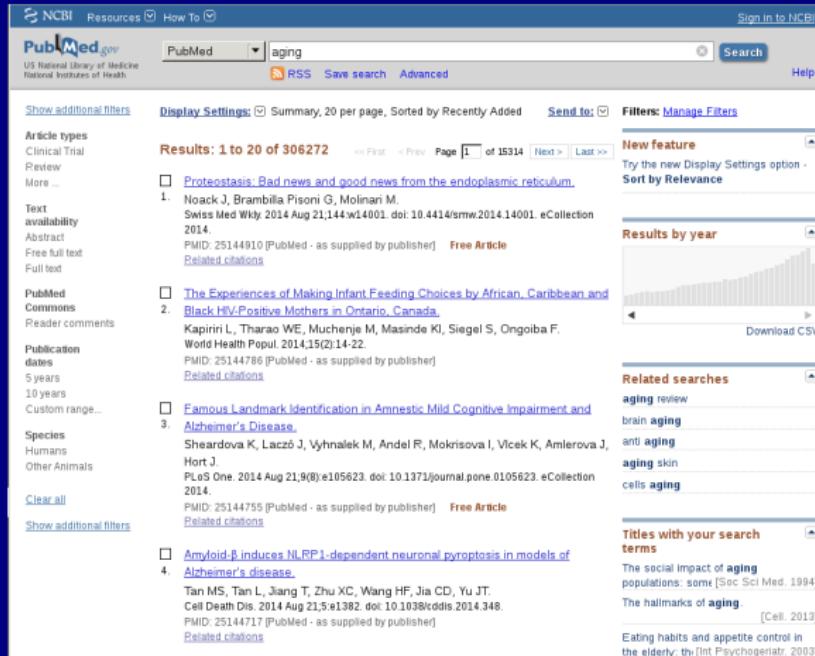
[Amyloid- \$\beta\$  induces NLRP1-dependent neuronal pyroptosis in models of Alzheimer's disease.](#)  
4. Tan MS, Wan L, Jiang T, Zhu XC, Wang HF, Jia CD, Yu JT.  
Cell Death Dis. 2014 Aug 21;5:e1382. doi: 10.1038/cddis.2014.348.  
PMID: 25144717 [PubMed - as supplied by publisher]  
[Related citations](#)

New feature Try the new Display Settings option. Sort by Relevance

Results by year Download CSV

Related searches aging review brain aging anti aging aging skin cells aging

Titles with your search terms The social Impact of aging populations. somt [Soc Sci Med. 1994] The hallmarks of aging [Cell. 2013] Eating habits and appetite control in the elderly. th [Int Psychogeriatr. 2003]



# Why is it Difficult to Understand Aging?

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?



# How can CREA Understand Aging?

Build a Knowledge Base that Describes and Reasons about Aging

Constructing a  
Knowledge  
Base on Aging

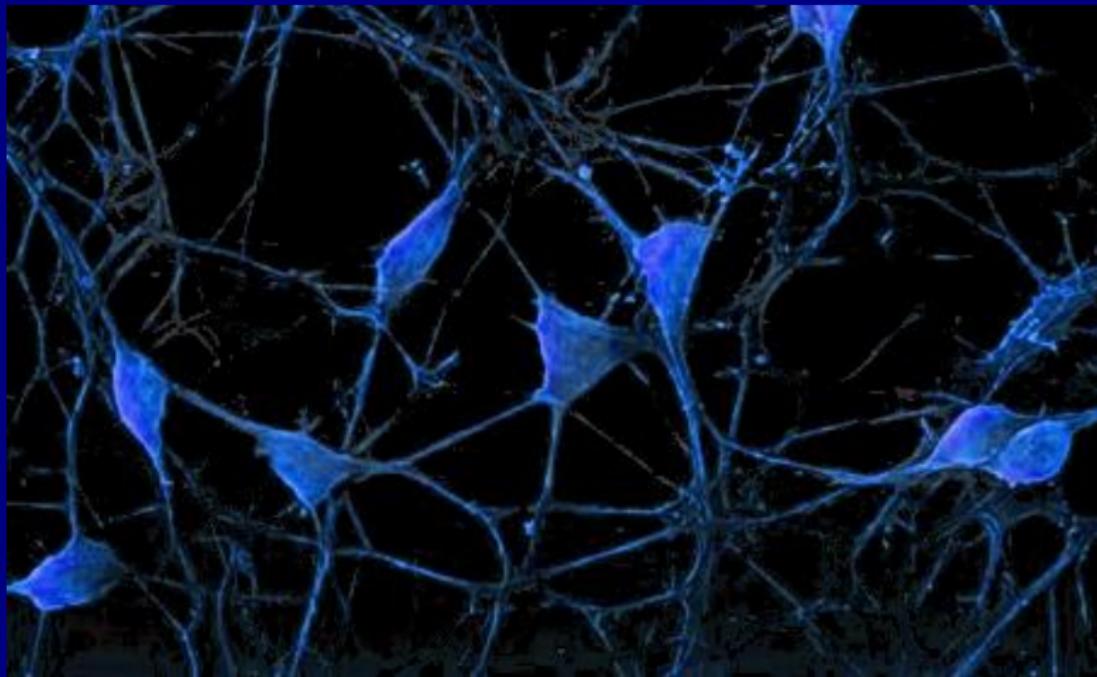
Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?



# How can CREA Understand Aging?

Build a Knowledge Base that Describes and Reasons about Aging

Constructing a  
Knowledge  
Base on Aging

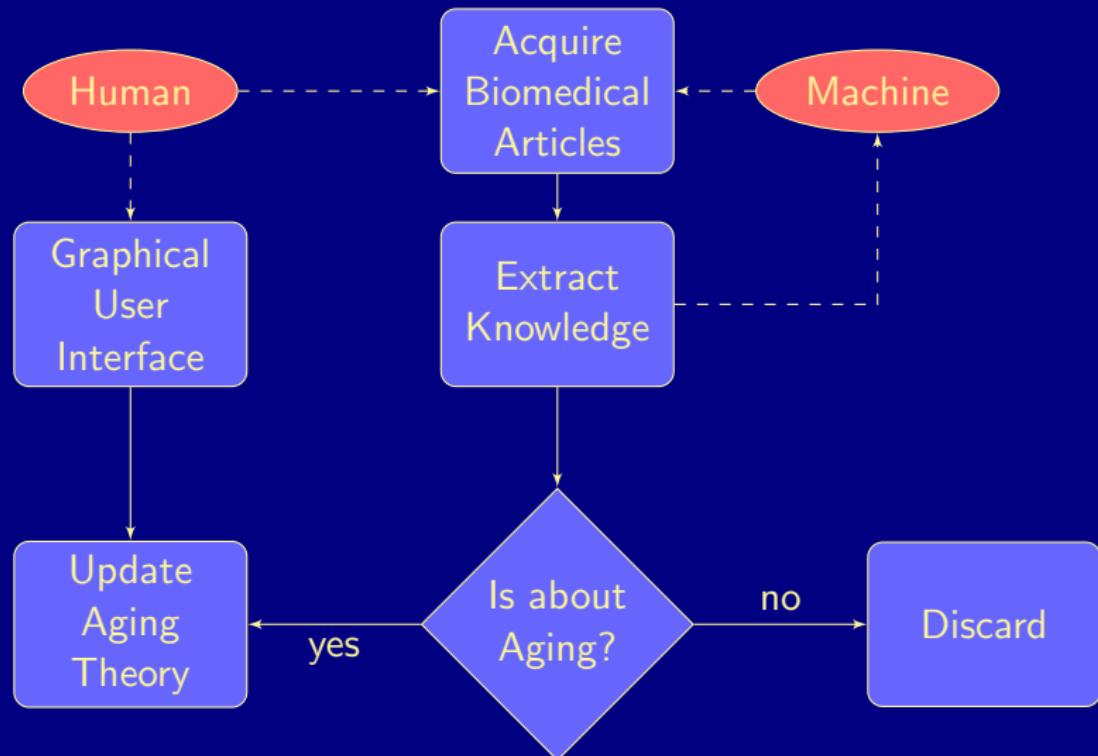
Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?



# Outline

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

## 1 About Me

## 2 Understanding Aging

## 3 Knowledge Extraction

## 4 What Next?

# Knowledge Extraction and Visualization

## Overview of Progress

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

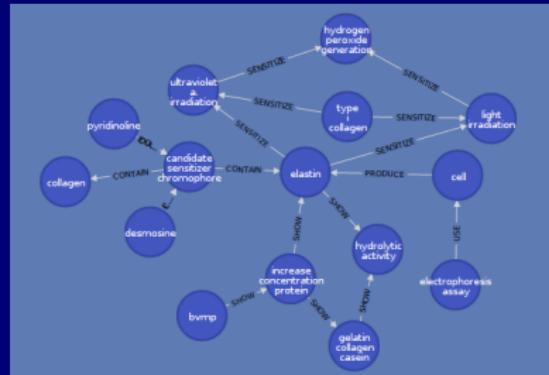
Knowledge  
Extraction

What Next?

## Input

... The pyridinoline and desmosine were examined as candidate sensitizer chromophores contained in collagen and elastin, respectively. ...

## Output



# Knowledge Extraction

## Method

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

Tokenization Input a text document and read it, one sentence at a time.

# Knowledge Extraction

## Method

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

**Tokenization** Input a text document and read it, one sentence at a time.

**Parsing** For each sentence, generate a constituent tree that describes its phrase structure.

# Knowledge Extraction

## Method

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

**Tokenization** Input a text document and read it, one sentence at a time.

**Parsing** For each sentence, generate a constituent tree that describes its phrase structure.

**Compilation** Extract facts by pattern matching on each constituent tree.

# Constituent Tree Tags

## Word Level

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

Tag	Meaning	Example
DT	Determiner	the
IN	Preposition	of
JJ	Adjective	blue
RB	Adverb	quickly
CC	Coordinating Conjunction	and
NN	Singular Noun	monkey
NNS	Plural Noun	monkeys
VB	Base Verb	fall
VBD	Singular Present Verb	falls
VBN	Past Tense Verb	fell
VBN	Past Participle Verb	fallen
VBG	Gerund Verb	falling
...	...	...

# Constituent Tree Tags

## Phrase Level

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

Tag	Meaning	Example
NP	Noun Phrase	the woman
VP	Verb Phrase	calls the man
PP	Prepositional Phrase	from the store
ADVP	Adverb Phrase	quickly and quietly
ADJP	Adjective Phrase	blue and red
CONJP	Conjunctive Phrase	as well as
...	...	...

# Constituent Tree Tags

## Clause Level

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

Tag	Meaning	Example
S	Declarative Clause	the dog walks
SBAR	Conjunction + Clause	that the dog walks
...	...	...

# Knowledge Extraction

## Example

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

### Input Token

The man walks the dog.

# Knowledge Extraction

## Example

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

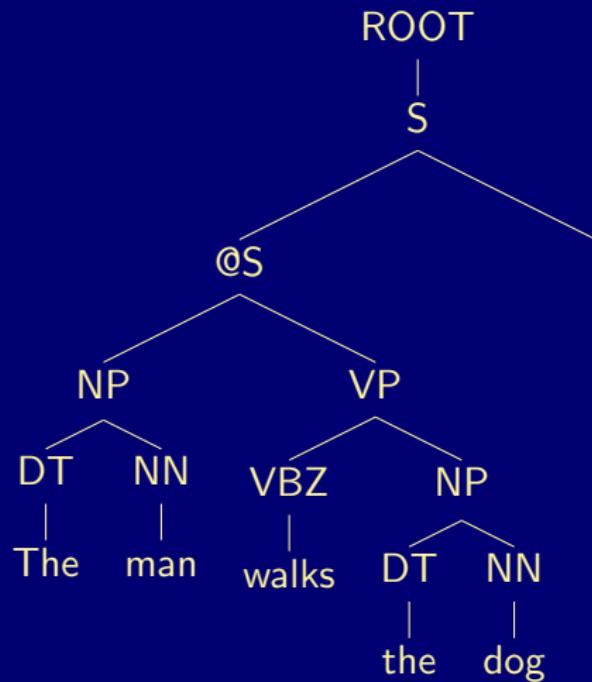
About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

### Parse Token



# Knowledge Extraction

## Example

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

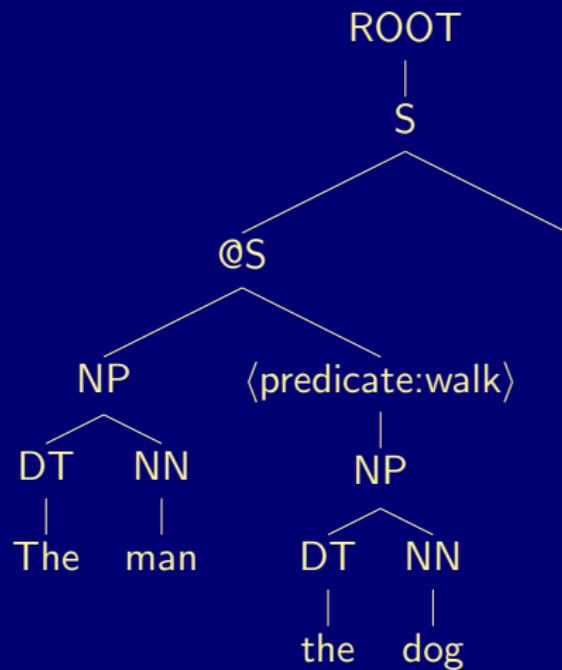
About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

### Compile Token



# Knowledge Extraction

## Example

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

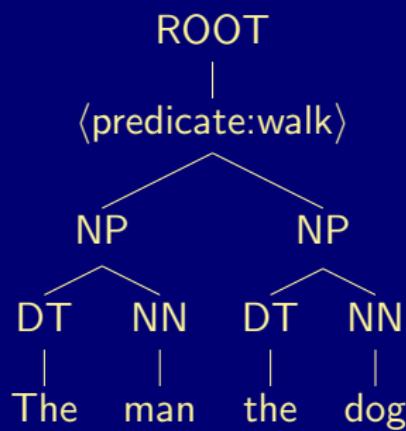
About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

### Compile Token



# Knowledge Extraction

## Example

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

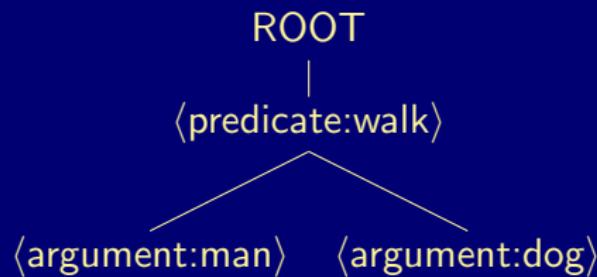
About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

### Compile Token



# Knowledge Extraction Example

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

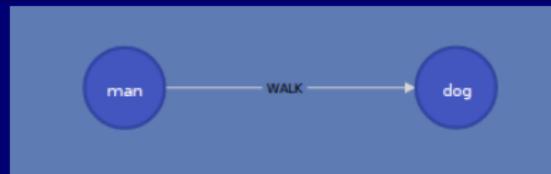
About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

## Output Fact



# Software Demonstration

A preview of CREA's knowledge base, compiled from PubMed abstracts.

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?



# High Performance

Knowledge can be extracted from many sentences at the same time.

Constructing a  
Knowledge  
Base on Aging

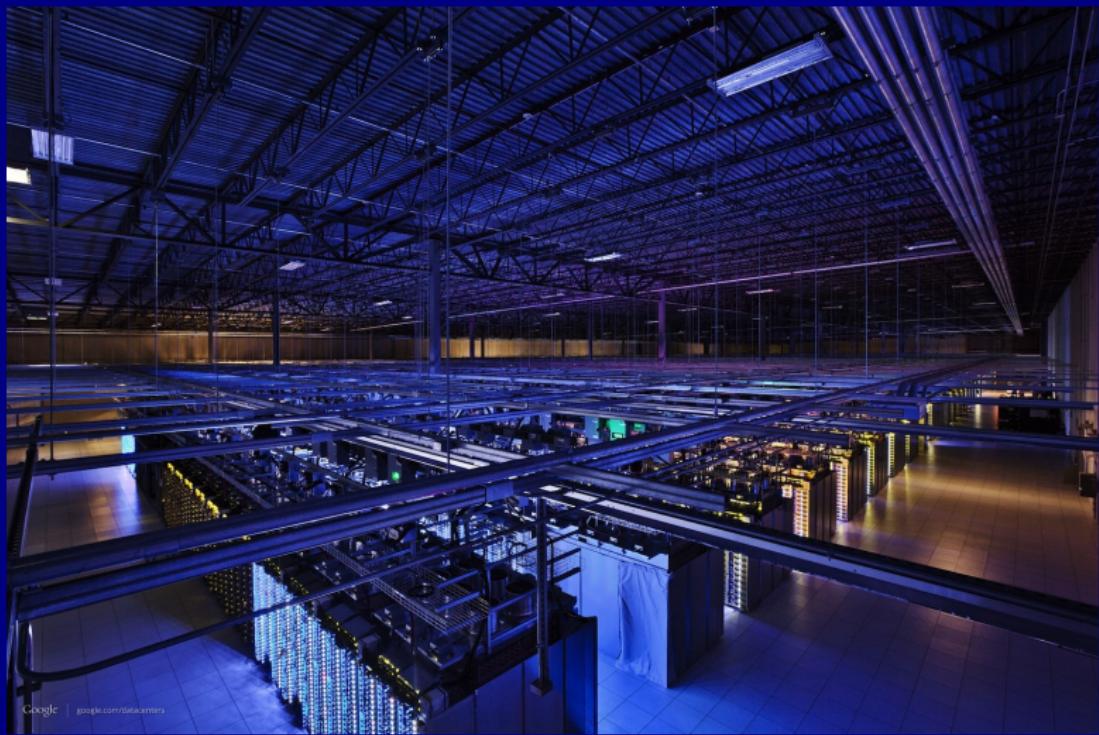
Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?



Google | [google.com/datacenters](http://google.com/datacenters)

# Outline

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

## 1 About Me

## 2 Understanding Aging

## 3 Knowledge Extraction

## 4 What Next?

# Accuracy

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

- Filter spam sentences from documents.

# Accuracy

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

- Filter spam sentences from documents.
- The accuracy of the parser could be optimized:

# Accuracy

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

- Filter spam sentences from documents.
- The accuracy of the parser could be optimized:
  - Should be trained to identify more nouns from the biomedical domain.

# Accuracy

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

- Filter spam sentences from documents.
- The accuracy of the parser could be optimized:
  - Should be trained to identify more nouns from the biomedical domain.
- Define more patterns for extracting facts:

# Accuracy

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

- Filter spam sentences from documents.
- The accuracy of the parser could be optimized:
  - Should be trained to identify more nouns from the biomedical domain.
- Define more patterns for extracting facts:
  - The software succeeds around 50% of the time.

# Missing Features

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

- Support negated clauses and conditional logic.

# Missing Features

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

- Support negated clauses and conditional logic.
- Facts can contradict each other:

# Missing Features

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

- Support negated clauses and conditional logic.
- Facts can contradict each other:
  - Store the probability that is true as the weight of its edge on the knowledge base's graph.

# Missing Features

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

- Support negated clauses and conditional logic.
- Facts can contradict each other:
  - Store the probability that is true as the weight of its edge on the knowledge base's graph.
- Scale and launch the software service.

# Conclusion

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

- Demonstrated a method for automatically constructing CREA's knowledge base on aging.

# Conclusion

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

- Demonstrated a method for automatically constructing CREA's knowledge base on aging.
- Showed how to extract facts from English text in the knowledge base's structured format.

# Conclusion

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

- Demonstrated a method for automatically constructing CREA's knowledge base on aging.
- Showed how to extract facts from English text in the knowledge base's structured format.
- Discussed the need to improve software accuracy by lensing in on the biomedical domain.

# Conclusion

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

- Demonstrated a method for automatically constructing CREA's knowledge base on aging.
- Showed how to extract facts from English text in the knowledge base's structured format.
- Discussed the need to improve software accuracy by lensing in on the biomedical domain.
- Suggested how the software implementation can be scaled for production usage.

# Acknowledgements

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

Steve Garan Bioinformatics Research Director  
Jeremy Wan Text Aggregation Development  
Grace Park Text Aggregation Development

# Get Involved

Contact Me for More Information

Constructing a  
Knowledge  
Base on Aging

Mark Farrell

About Me

Understanding  
Aging

Knowledge  
Extraction

What Next?

Email [m4farrel@csclub.uwaterloo.ca](mailto:m4farrel@csclub.uwaterloo.ca)

Website [markfarrell.ca](http://markfarrell.ca)