



Ambient Search

just-in-time document recommendations for speech fragments

GSCL – 13th September 2017

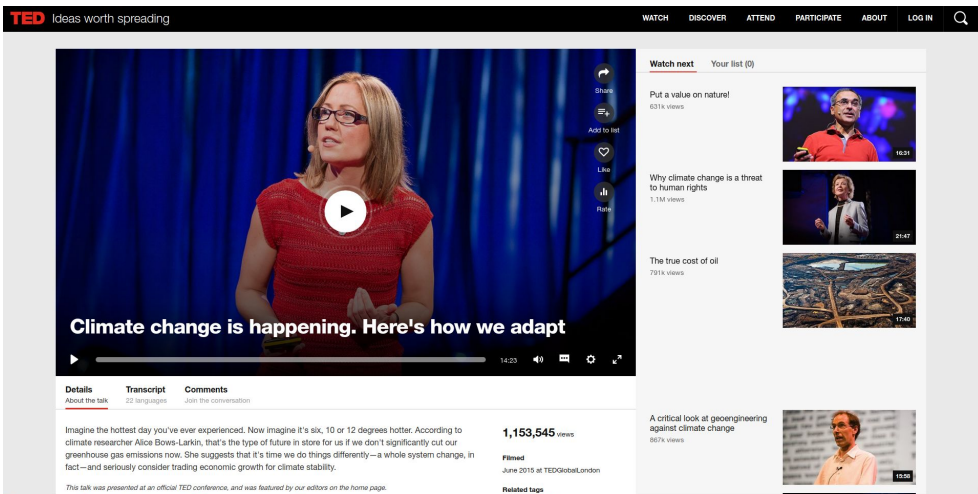
Bachelor thesis presentation by Jonas Wacker

Thesis supervision:

Prof. Dr. Chris Biemann and Benjamin Milde

Former Language Technology Department, TU Darmstadt

Ambient Search - Motivation



Screenshot from <https://www.ted.com/>

- Just-in-time document recommendations to complement audio input
- Knowledge retrieval **without interrupting** presentations / meetings / conversations
- Presented information can be inspected **during or after the talk**

System Overview



...over our lifetimes
we've all contributed
to climate change...

1) Transcription of
the audio signal

*Benjamin Milde, Jonas Wacker, Stefan Radomski,
Max Mühlhäuser and Chris Biemann:*

(1) Ambient Search: A Document Retrieval System for Speech Streams

*(2) Demonstrating Ambient Search: Implicit Document Retrieval for
Speech Streams*

*Proceedings of COLING 2016
Osaka, Japan*

System Overview



1) Transcription of
the audio signal



...over our **lifetimes**
we've all contributed
to **climate change**...

2) Keyphrase extraction

*Benjamin Milde, Jonas Wacker, Stefan Radomski,
Max Mühlhäuser and Chris Biemann:*

(1) Ambient Search: A Document Retrieval System for Speech Streams

*(2) Demonstrating Ambient Search: Implicit Document Retrieval for
Speech Streams*

*Proceedings of COLING 2016
Osaka, Japan*

System Overview

Benjamin Milde, Jonas Wacker, Stefan Radomski,
Max Mühlhäuser and Chris Biemann:

(1) *Ambient Search: A Document Retrieval System for Speech Streams*

(2) *Demonstrating Ambient Search: Implicit Document Retrieval for Speech Streams*

Proceedings of COLING 2016
Osaka, Japan



...over our **lifetimes**
we've all contributed
to **climate change**...

1) Transcription of
the audio signal

2) Keyphrase extraction



(I) climate change
(II) lifetimes
(III) ...

3) Keyphrase ranking

System Overview

Benjamin Milde, Jonas Wacker, Stefan Radomski,
Max Mühlhäuser and Chris Biemann:

(1) *Ambient Search: A Document Retrieval System for Speech Streams*

(2) *Demonstrating Ambient Search: Implicit Document Retrieval for Speech Streams*

Proceedings of COLING 2016
Osaka, Japan



...over our **lifetimes**
we've all contributed
to **climate change**...

1) Transcription of
the audio signal

2) Keyphrase extraction



(I) climate change
(II) lifetimes
(III) ...



"climate change"^2.3 OR
"lifetimes"^1.7 OR ...

3) Keyphrase ranking

4) Construction of a search query

System Overview

Benjamin Milde, Jonas Wacker, Stefan Radomski,
Max Mühlhäuser and Chris Biemann:

(1) Ambient Search: A Document Retrieval System for Speech Streams

(2) Demonstrating Ambient Search: Implicit Document Retrieval for Speech Streams

Proceedings of COLING 2016
Osaka, Japan



...over our **lifetimes**
we've all contributed
to **climate change**...

1) Transcription of
the audio signal

2) Keyphrase extraction



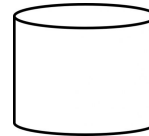
(I) climate change
(II) lifetimes
(III) ...



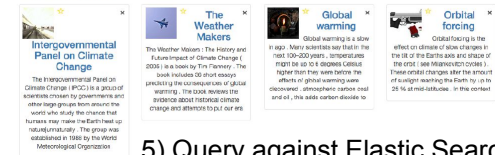
"climate change"^{2.3} OR
"lifetimes"^{1.7} OR ...

3) Keyphrase ranking

4) Construction of a search query



Relevant Entries



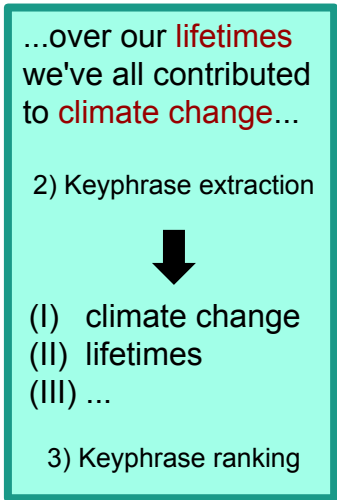
5) Query against Elastic Search index

Video of the system in action

System Overview



1) Transcription of the audio signal

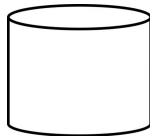


Focus of my work

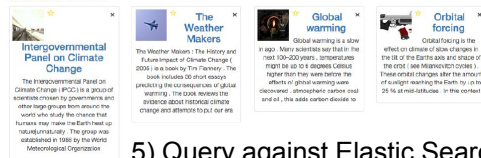


"climate change"^{2.3} OR
"lifetimes"^{1.7} OR ...

4) Construction of a search query



Relevant Entries



5) Query against Elastic Search index

Benjamin Milde, Jonas Wacker, Stefan Radomski,
Max Mühlhäuser and Chris Biemann:

(1) Ambient Search: A Document Retrieval System for Speech Streams

(2) Demonstrating Ambient Search: Implicit Document Retrieval for Speech Streams

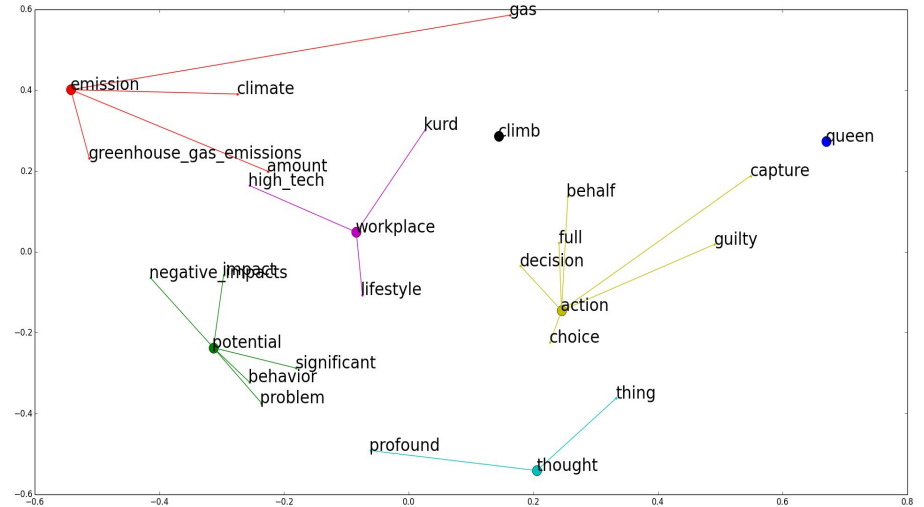
Proceedings of COLING 2016
Osaka, Japan

Keyphrase Extraction and Ranking

over our **lifetimes** we've all contributed to **climate change**. **actions choices** and the **behaviors** will have led us to have **increased** in **queen** has **gas emissions** and i think that's quite a **powerful full** but it does have the **potential** to make us feel **guilty**...



lifetimes
climate_change
actions
choices
behaviors
increased
queen
gas_emissions
powerful
full
potential
guilty
...

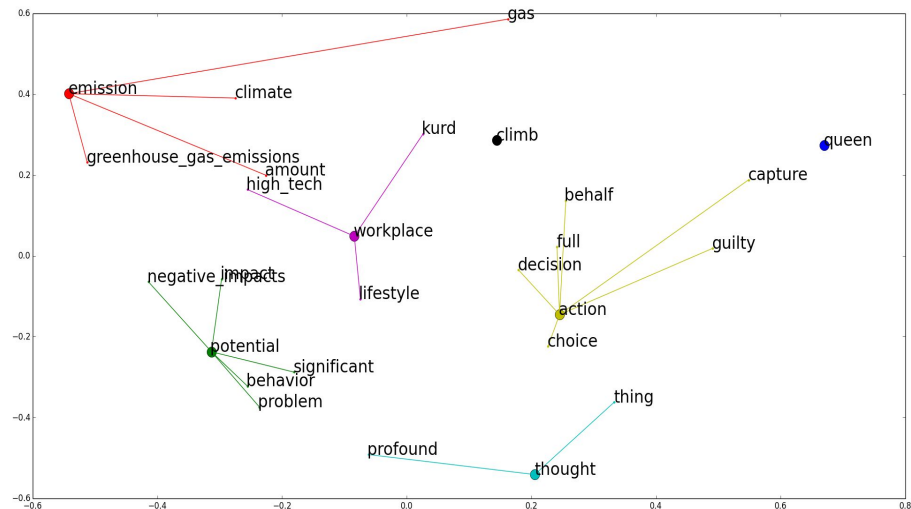


Keyphrase Extraction and Ranking

over our **lifetimes** we've all contributed to **climate change**. **actions choices** and the **behaviors** will have led us to have **increased** in **queen** has **gas emissions** and i think that's quite a **powerful full** but it does have the **potential** to make us feel **guilty**...



lifetimes
climate_change
actions
choices
behaviors
increased
queen
gas_emissions
powerful
full
potential
guilty
...



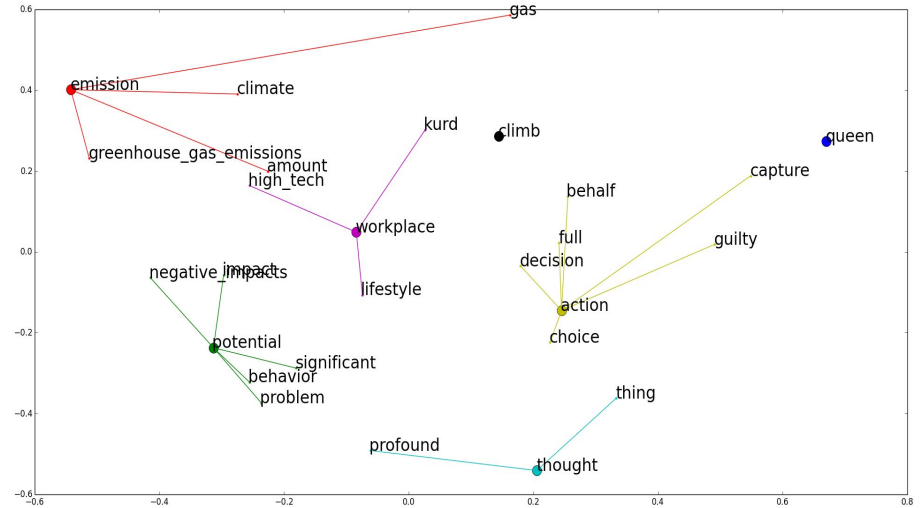
1. Filter adjectives and nouns
2. Threshold IDF scores
3. Construction of phrases (DRUID)

Keyphrase Extraction and Ranking

over our **lifetimes** we've all contributed to **climate change**. **actions choices** and the **behaviors** will have led us to have **increased** in **queen** has **gas emissions** and i think that's quite a **powerful full** but it does have the **potential** to make us feel **guilty**...



lifetimes
climate_change
actions
choices
behaviors
increased
queen
gas_emissions
powerful
full
potential
guilty
...



Map tokens into Word2Vec vector space

1. Filter adjectives and nouns
2. Threshold IDF scores
3. Construction of phrases (DRUID)

Ranking Formula

$$\text{score}(\text{term}_k, \text{trans}) = d_{\cos} \left(w_2 v(\text{term}_k), \frac{1}{|\text{terms}|} \sum_{i=1}^{|\text{terms}|} w_2 v(\text{term}_i) \right) \cdot \text{TFIDF}(\text{term}_k, \text{trans})$$

Ranking Formula

$$score(term_k, trans) = d_{cos} \left(w2v(term_k), \frac{1}{|terms|} \sum_{i=1}^{|terms|} w2v(term_i) \right) \cdot TFIDF(term_k, trans)$$

Cosine similarity to average word vector

Ranking Formula

$$score(term_k, trans) = d_{cos} \left(w2v(term_k), \frac{1}{|terms|} \sum_{i=1}^{|terms|} w2v(term_i) \right) \cdot TFIDF(term_k, trans)$$

Cosine similarity to average word vector TF-IDF Score w.r.t. fragment

Evaluation Results

Dataset: 30 TED speech fragments (around 1.5-2.5 minutes each, 10 different talks)

Method	Avg. Recall (Std. Dev. in %)	Avg. HRR (Std. Dev. in %)	Keyphrase Relevance
(1) Our method with added TF-IDF and preprocessing	41.48% (11.45%)	24.78% (16.4%)	16.70%
(2) Habibi and Popescu-Belis with stopwords removal only, no TF-IDF	38.88% (16.42%)	32.59% (16.47%)	6.29%
(3) Our method with stopwords removal only, no TF-IDF	34.44% (15.8%)	30.78% (13.05%)	3.66%

Reference Paper for approach (2):

Maryam Habibi and Andrei Popescu-Belis: *Keyword Extraction and Clustering for Document Recommendation in Conversations*
IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2015

Thank you!

Link to the Github Repository: <https://github.com/bmilde/ambientsearch>

Timeline



Relevant Entries

Speech Input

You: over our lifetimes we've all contributed to climate change.

You: actions choices and the behaviors will have less to have increased in crane has gas emissions.

0:03

Relevant Entries



☆

Intergovernmental Panel on Climate Change ✕

The Intergovernmental Panel on Climate Change is a group of scientists chosen by governments and other large groups from around the world who study the chance that humans may make the Earth heat up nature



☆

Global warming ✕

Climate change has happened many times over the history of the Earth, including the coming and going of ice ages.



☆

Greenhouse effect ✕

The greenhouse effect occurs when certain gases in the Earth's atmosphere entraps infrared radiation.



☆

Stern Review ✕

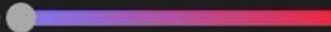
The British government. It discusses the effect of climate change and global warming on the world economy.

Speech Input

You: Over our lifetimes we've all contributed to climate change.

You: actions choices and the behaviors will have less to have increased in crane has gas emissions and i think that's quite a powerful full but it does have the potential to make us feel guilty when we think about decisions we might have made around where to travel to.

You: how often and how about.



0:33



Greenhouse effect



The greenhouse effect occurs when certain gases in the Earth's atmosphere entrap infrared radiation.

0:49



Global warming



Climate change has happened many times over the history of the Earth, including the coming and going of ice ages.

0:49



Intergovernmental Panel on Climate Change



The Intergovernmental Panel on Climate Change is a group of scientists chosen by governments and other large groups from around the world who study the chance that humans may make the Earth heat up nature

0:49



Storms of My Grandchildren



Storms of My Grandchildren is the first book written by climate scientist James Hansen.

Intergovernmental Panel on Climate Change



Intergovernmental Panel on Climate Change

From Wikipedia, the free encyclopedia

The **Intergovernmental Panel on Climate Change (IPCC)** is a group of scientists chosen by governments and other large groups from around the world who study the chance that humans may make the Earth heat up unnaturally. The group was established in 1988 by the World Meteorological Organization (WMO) and the United Nations Environment Programme (UNEP), two organizations of the United Nations.

The IPCC shared the 2007 Nobel Peace Prize with former Vice President of the United States Al Gore who won for working on the same problems.^[1]

A lot of IPCC work is publishing reports about the United Nations Framework Convention on Climate Change (UNFCCC),^[2] an international agreement that human inventions and chemistry may make the Earth too hot to live on. The UNFCCC was the beginning of the Kyoto Protocol. Members of the IPCC read, write, and calculate as much as they can. Only member states of the WMO and UNEP may be members of IPCC. A lot of professors trust the IPCC work.^{[3][4]}

Contents

- 1 IPCC Reports
 - 1.1 IPCC Fourth Assessment Report: Climate Change 2007
 - 1.2 IPCC Third Assessment Report: Climate Change 2001
 - 1.2.1 Changing scientists reports

Close

Relevant



Intergovernmental Panel on Climate Change

The Intergovernmental Panel on Climate Change (IPCC) is a group of scientists chosen by governments and other large groups from around the world who study the chance that humans may make the Earth heat up unnaturally.

Speech

You: actions of powerful full but it to travel to.

You: how often

Stern Review

British government. Effect of climate warming on the economy.



"I think I speak for all of us when I say what in God's name are you talking about?"