

Human Goal Classification of Natural Language Text

Mark Kröll,

Knowledge Management Institute
Graz University of Technology

Reid Swanson and Andrew Gordon

Institute for Creative Technologies
University of Southern California

Excerpt from Barack Obama's Denver Speech:

I will stop giving the wealthiest Americans tax cuts that they don't need and didn't ask for, and restore fairness to our economy.

I'll give a tax cut to working people; provide relief to homeowners; and eliminate the income tax for seniors making under \$50,000 so they can retire with the dignity and security they have earned.

→ Intentional Profile of this speech

Abbreviation	Full label
Achieving salvation	Achieving salvation
Arts	Appreciating the arts
Aspirations	Achieving my aspirations
Attracting sexually	Being able to attract, please, sexually excite
Avoiding failure	Avoiding failure
<div style="border: 1px solid black; padding: 5px; display: inline-block;"> <p>Charity Helping the needy</p> </div>	
Being affectionate	Being affectionate toward others
Being ambitious	Being ambitious, hard-working
Being better than others	Being better than others, beating others
Being carefree	Being lighthearted, carefree, enjoying life
Being clean	Being clean, neat (personal care)

Taxonomy of Human Goals

(developed by Read et al. [Chulef01])

however, human goals are seldom mentioned explicitly in plain text ...

need a connection between text and the human goal taxonomy

→ actions that contribute to the achievement of a goal are expressed quite often

Profiles of People's Interests

- knowledge about a person's interests can be used to create an informative profile
- from knowing people's goals and interests one can infer
 - their opinions
 - their relationship with other people
 - their attitude towards life
- Acquiring the data represents the easy part
 - Weblogs
 - Transcripts of political speeches
- Creating an interest profile out of it, the more challenging part

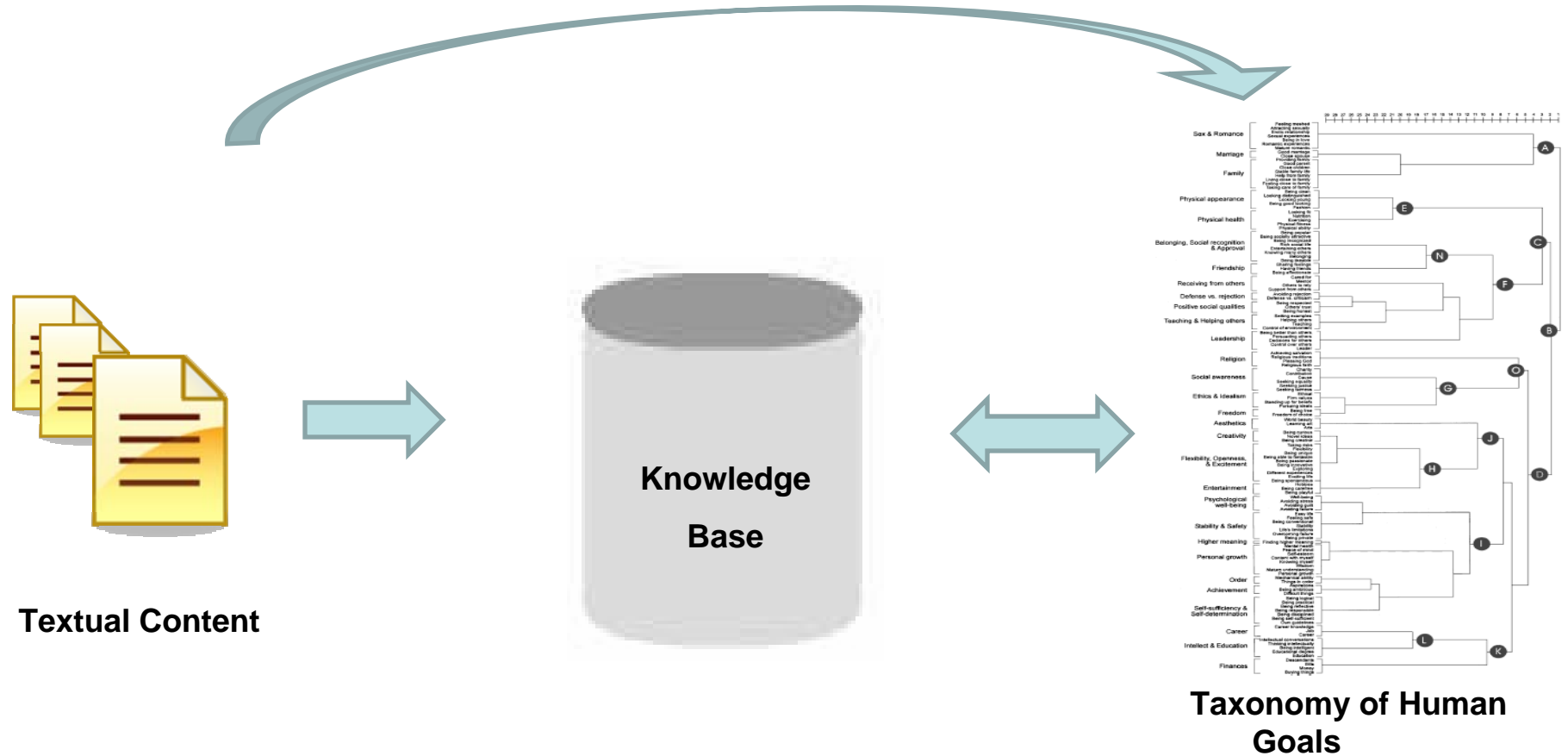


Textual data



??

2008



The idea is to:

- 1.) collect a list of representative actions that hint towards goal categories (→ Knowledge Base)
- 2.) based on the identification of actions, goal categories are assigned

Abbreviation	Full label
Achieving salvation	Achieving salvation
Arts	Appreciating the arts
Aspirations	Achieving my aspirations
Attracting sexually	Being able to attract, please, sexually excit
Avoiding failure	Avoiding failure
Avoiding guilt	Avoiding feelings of guilt
Avoiding rejection	Avoiding rejection by others
Category: Looking Young	
Being better than others	Being better than others, beating others
Being carefree	Being lighthearted, carefree, enjoying life
Being clean	Being clean, neat (personal care)

Taxonomy of Human Goals

Brainstorming

Phrases:

- Avoid wrinkles
- Age well
- Be vibrant with Energy
- Looking Vital
- ...

Causal Relations

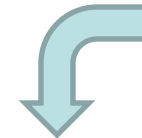
Phrase Search Queries

- In order to avoid wrinkles
- Essential for aging well
- Necessary for looking vital
- ...



[Yahoo! BOSS API](#)

Processing of textual content



Data preparation and searching the index

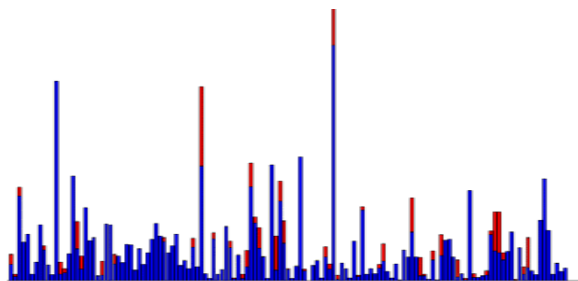


Looking Young ← you need **to moisturize inside and out**

Looking Young ← but the biggest reason women have such high risk of vitamin D deficit according to Holick, "women are encouraged **to avoid all sunlight** and skin cancer.



Political Speeches



Profile

Profile Creation by Action Identification

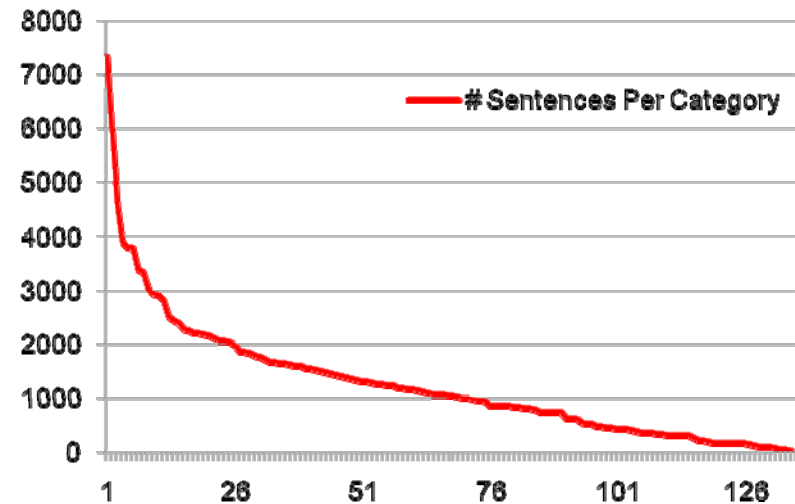


Knowledge Base/ Index

2008

Quality of the Knowledge Base

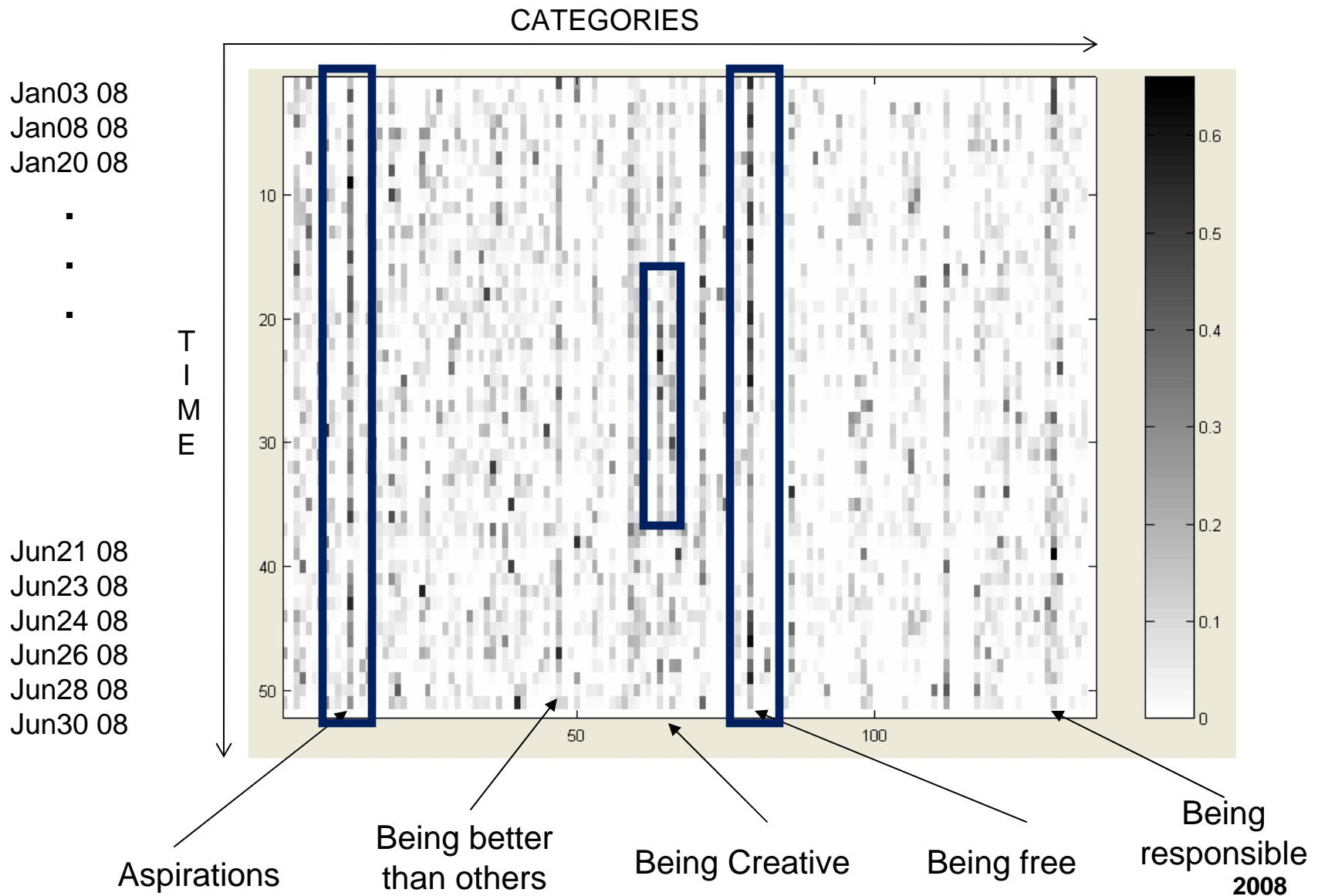
- Some facts:
 - contains 168.657 sentences
 - min: 12 (Category: Firm Values)
 - max: 7323 (Category: Helping Others)
 - yielding a skewed distribution



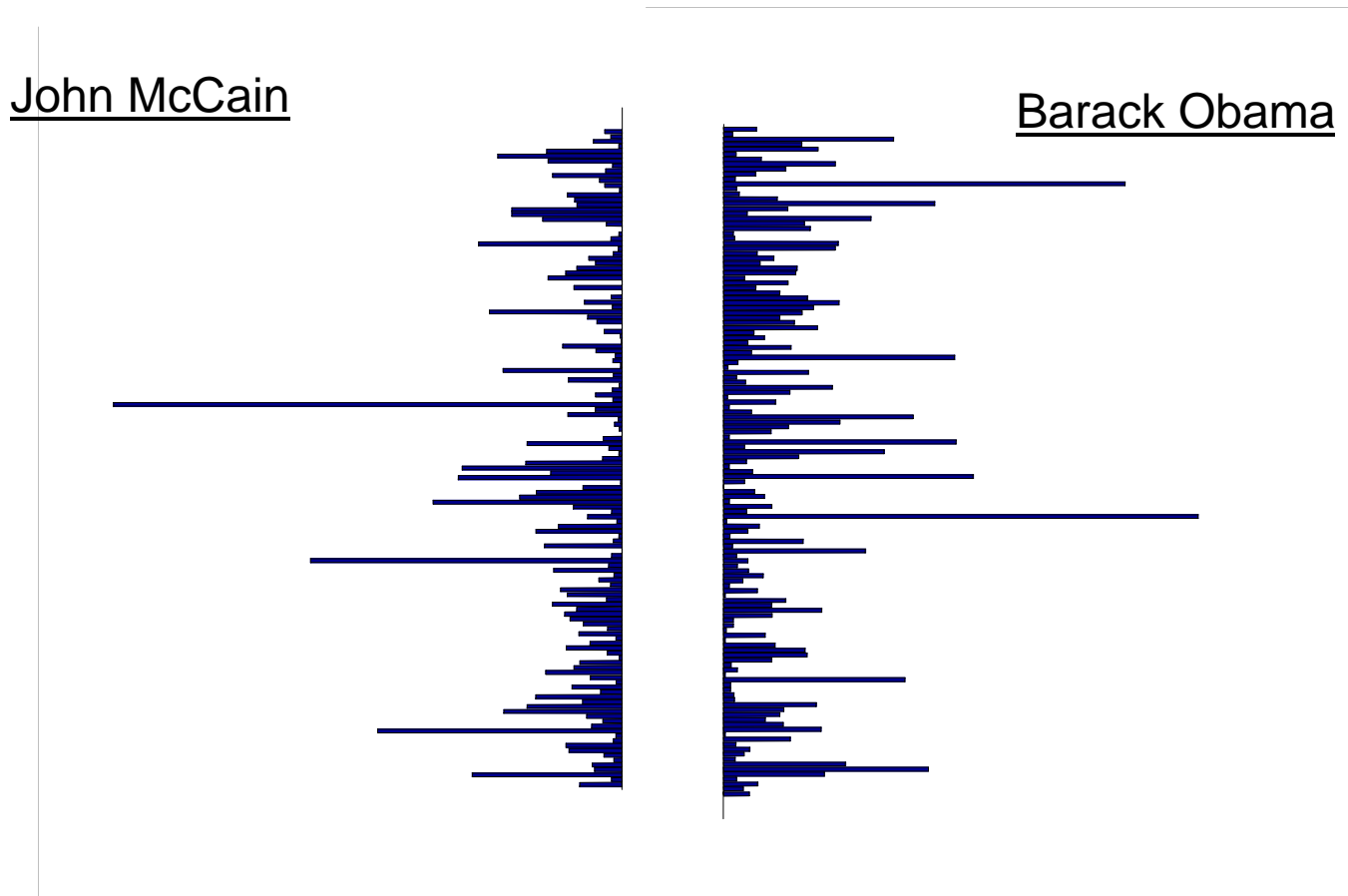
- Annotation Task
 - to approximate the precision of the entries
 - not relevant to the category
 - not containing an action that can be performed to achieve the goal
 - random sample consisting of 674 entries

57% correct entries vs. 43% incorrect entries

Barack Obama – 51 Speeches (135 Categories)



Comparing Average Profiles



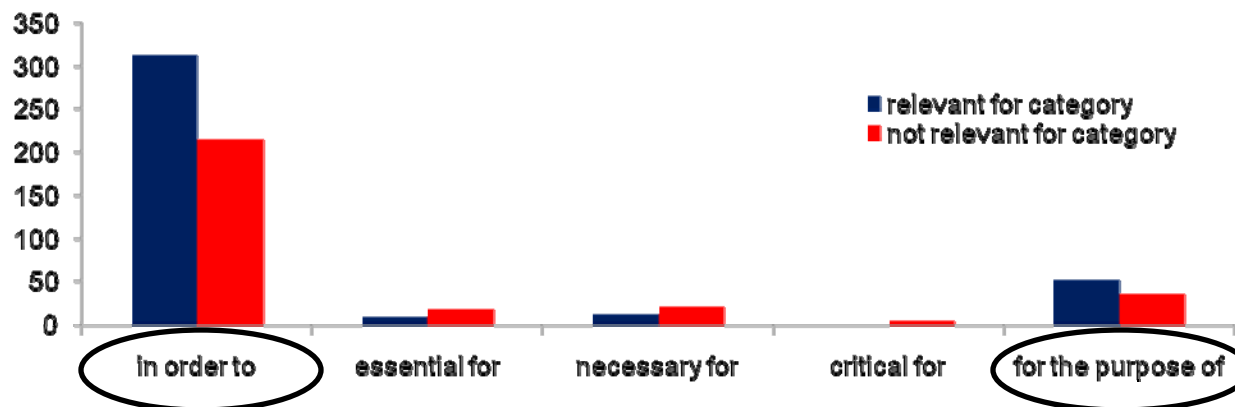
Average Profiles based on 51 speeches of Obama and 43 speeches of McCain given between January and June 2008

Evaluation

<u>Sentences out of speech:</u>	<u>Assigned Category:</u>	<u>Score:</u>
I'll give a tax cut to working people; provide relief to homeowners; and eliminate the income tax for seniors making under \$50,000 so they can retire with the dignity and security they have earned.	Charity	0.59
We need to widely reform the way we do business in Washington; to end wasteful spending that does little if anything to meet government's obligations to the American people.	Ethical	0.62
I am running for President because I believe that we need fundamental change in America.	Bills	0.92

Improving the Quality

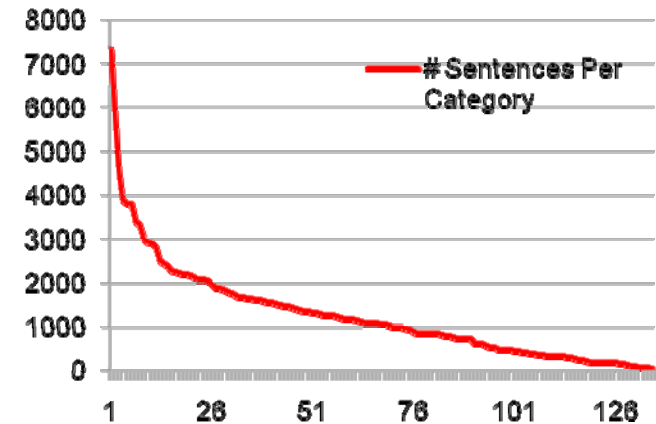
- by a more sophisticated pre-processing
 - using bigrams
 - using verb/noun bigrams (need part-of-speech tagging)
- by applying a pre-classification
 - where sentences are pre-classified to ensure presence of an action
 - using for instance verb phrases out of parse trees as features
- by using only advantageous causal relation according to the annotation task



2008

Size of the Knowledge Base

- Weak points
 - skewed distribution of sentences
 - number of sentences per category too low
- Means to increase the amount of sentences
 - Revising the search phrases
 - adding further phrases
 - expansion of present phrases (word net)
 - Use Yahoo! BOSS API to retrieve more results per submitted query
 - Now restricted to 500



Discussion

- How could we identify actions that are relevant for a certain category?
 - Example for the search phrase: “in order to age well”
Cork has been used for over 400 years, and many winemakers today still believe that in order to age well, wine needs gradual exposure to oxygen
 - Heuristics vs. automatic approach
- How important is the corpus where we acquire the actions from?
 - Are other corpora (Yahoo! Answers, Wikipedia) better suited?
- To what extent does the difference in vocabulary (web vs. Political speeches) influence the profile generation?

Thank you for your attention!

References

- [Chulef01] Chulef, A. S.; Read, S. J. & Walsh, D. A. (2001), 'A Hierarchical Taxonomy of Human Goals', *Motivation and Emotion* **25**(3), 191--232.
- [Quirk85] Quirk, R.; Greenbaum, S.; Leech, G. & Svartvik, J. (1985), *A Comprehensive Grammar of the English Language*, Longman, London.

- verb/noun bigram example

The sentence:

“In order to look young, people are willing to undergo surgeries and enhancement procedures that cost a lot of time and money.”

would produce following bigrams:

“undergo surgeries”

“undergo enhancement”

“undergo procedures”

“cost time”

“cost money”

Finding Actions - Examples

Search phrase: “In order to avoid wrinkles”

Extracted Sentences out of Web Content:

- You need **to moisturize inside and out**, in order to avoid wrinkles.
- But the biggest reason women have such high risk of vitamin D deficit according to Holick, “is that women are encouraged **to avoid all sunlight** in order to avoid wrinkles and skin cancer.

[back](#)