

# Automated Linguistic Personalization of Targeted Marketing Messages Mining User-generated Text on Social Media

**Rishiraj Saha Roy**

**Adobe Research Lab India**

**Aishwarya Padmakumar**

**IIT Madras**

**Guna Prasaad Jeganathan**

**IIT Bombay**

**Ponnurangam Kumaraguru**

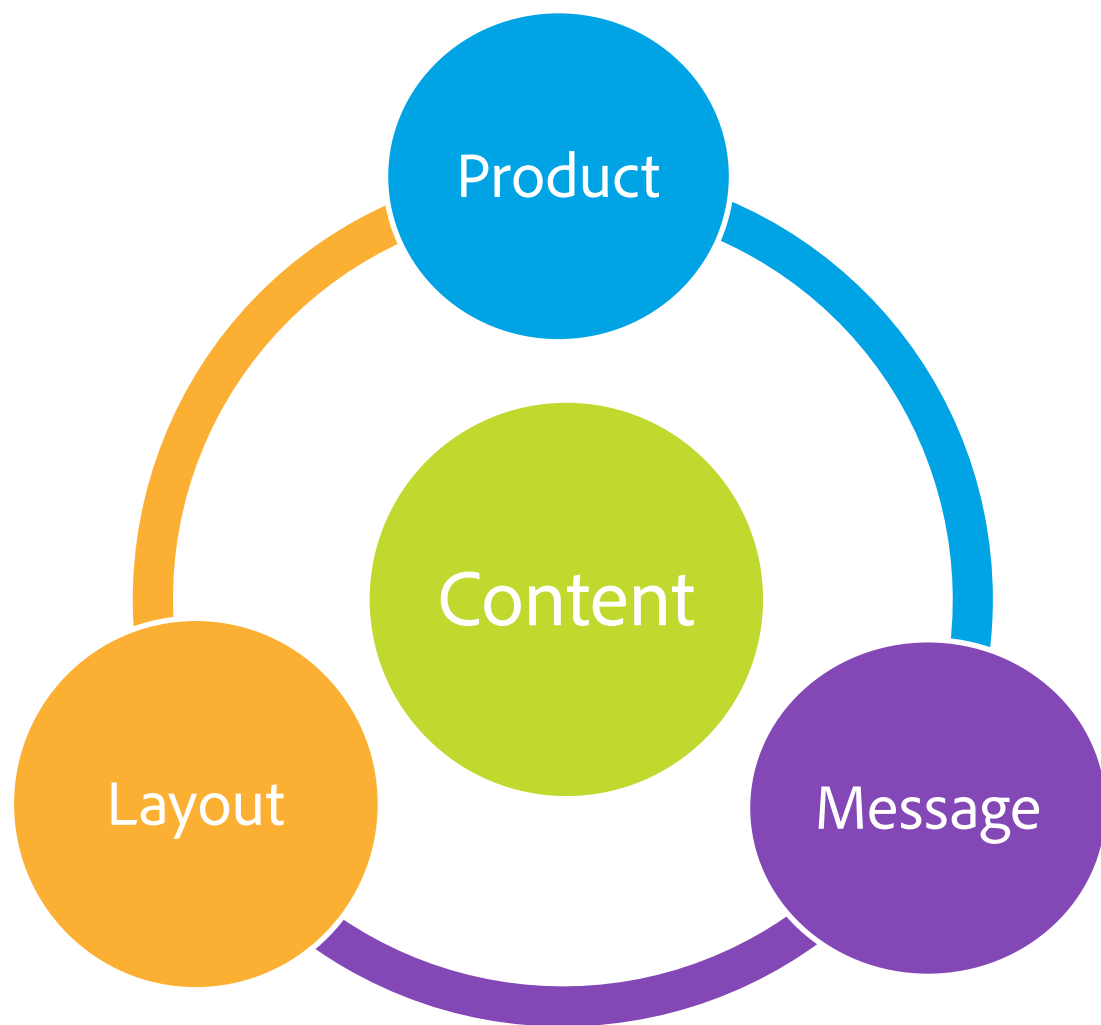
**IIIT Delhi**

# Agenda

- Motivation
- Data
- Method
- Results



# The need for personalization



<http://venturebeat.com/2013/05/07/mass-marketing-vs-personalization-infographic/>

# The power of personalization

## ALL POWERFUL

Check out the all-new powerful line of gaming laptops from Alienware.

Now pre-loaded with a copy of the award-winning game, Portal.

IMMEDIATE EXECUTE.  
ROGER THAT!

Place Order



PORTAL

## THE POWER TO DO MORE.

The Latitude Series makes high performance computing portable.

The best laptop in the market for high performance scientific computing.  
Improves the computing power by 20%.


Get the latest version of Mathematica completely free!

Place Order





# Importance of message text

-  If there are problems with how this message is displayed, click here to view it in a web browser.  
Click here to download pictures. To help protect your privacy, Outlook prevented automatic download of some pictures in this message.



[Unsubscribe](#)

----- Forwarded message -----

From: **Adobe** <[adobeindia@adobemarketing.co.in](mailto:adobeindia@adobemarketing.co.in)>

Date: 23 April 2014 12:16

Subject: Draw like never before using the all-new Pencil tool

To: [rishiraj.saharoy@live.in](mailto:rishiraj.saharoy@live.in)

[Read online](#)



## Adobe Illustrator CC

When designers have smart ideas for improving our features, we listen. Take a look at some of the features we redesigned to help you do more in less time.

# Why personalize ad message?

People from different demographics talk differently (Schwartz et al. 2013)

Linguistic style is an important component in advertising (Lowrey 1998)



**Communicating to specific audience segments in their own linguistic styles is expected to increase engagement with advertisements**

# Why automate message personalization?



Content authors or copywriters are creative individuals



Diversification and personalization key to user engagement



Several versions of same content required with limited resources





# What can we do as the first steps?

Word usage patterns representative of linguistic styles [Hu et al. 2013]

Adjectives convey emotion and make the advertisement sound more urgent and exciting [Content Marketing Tips, Econsultancy, June 2014]

**Draw with the Pencil tool.**

**Integrating platforms with flexibility.**

**Content management features and tools that get your work done.**



# What can we do as the first steps?

Word usage patterns representative of linguistic styles [Hu et al. 2013]

Adjectives convey emotion and make the advertisement sound more urgent and exciting [Content Marketing Tips, Econsultancy, June 2014]

Draw more accurately with the all-new Pencil tool.

Seamlessly integrating multiple platforms with amazing flexibility.

Must-have content management features and innovative tools that get your work done really fast.



# What can we do as the first steps?

- **Insertion** of **positive** adjectives and adverbs
- Preserves key **content** in the ad
- **Distributions** vary between **segments**
- **Automate** insertion **constrained** by usage patterns in segment





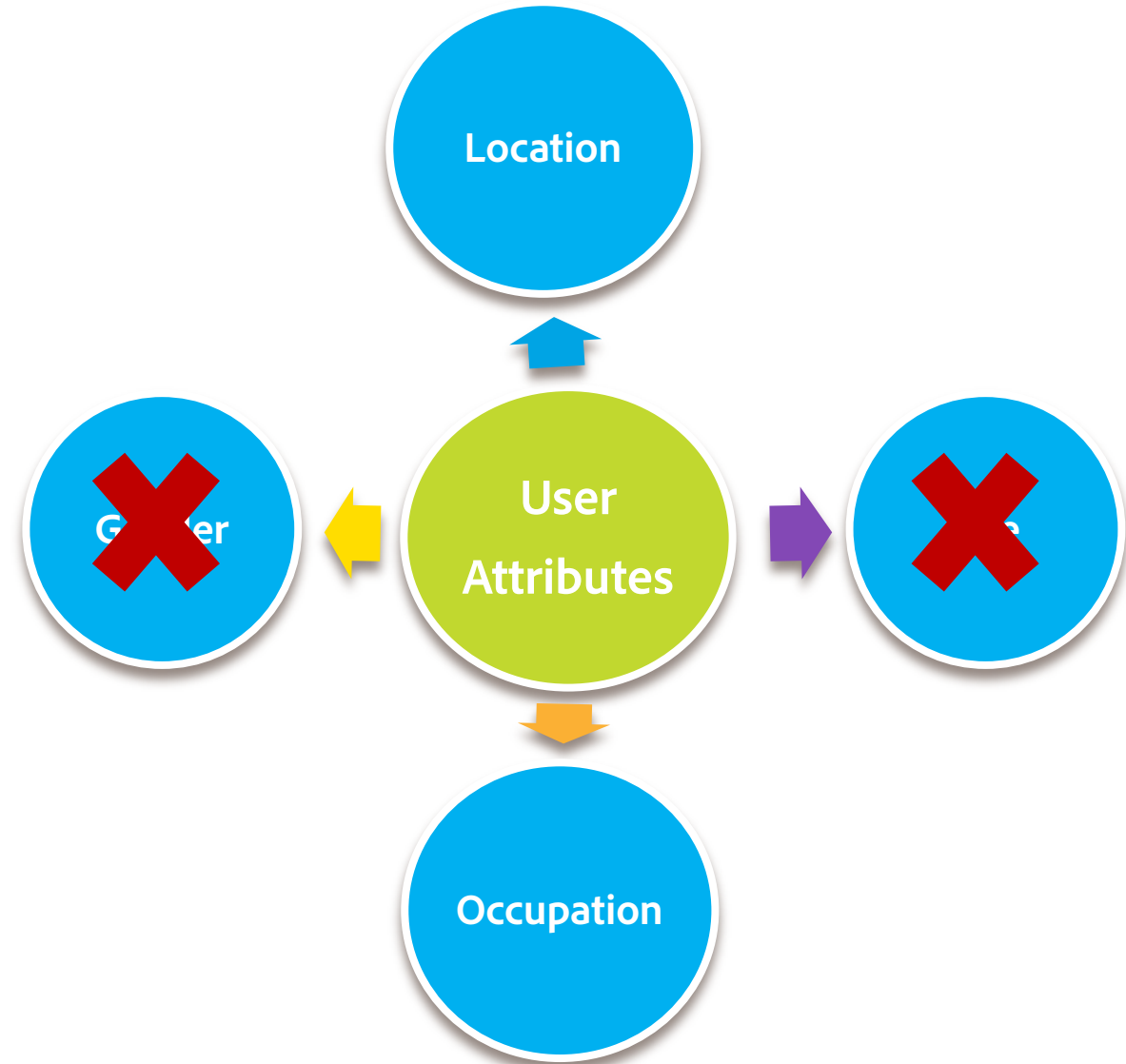
# Agenda

- Motivation
- **Data**
- Method
- Results

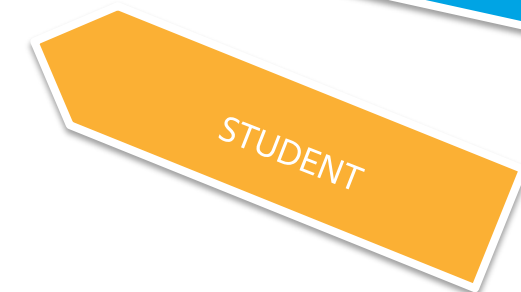
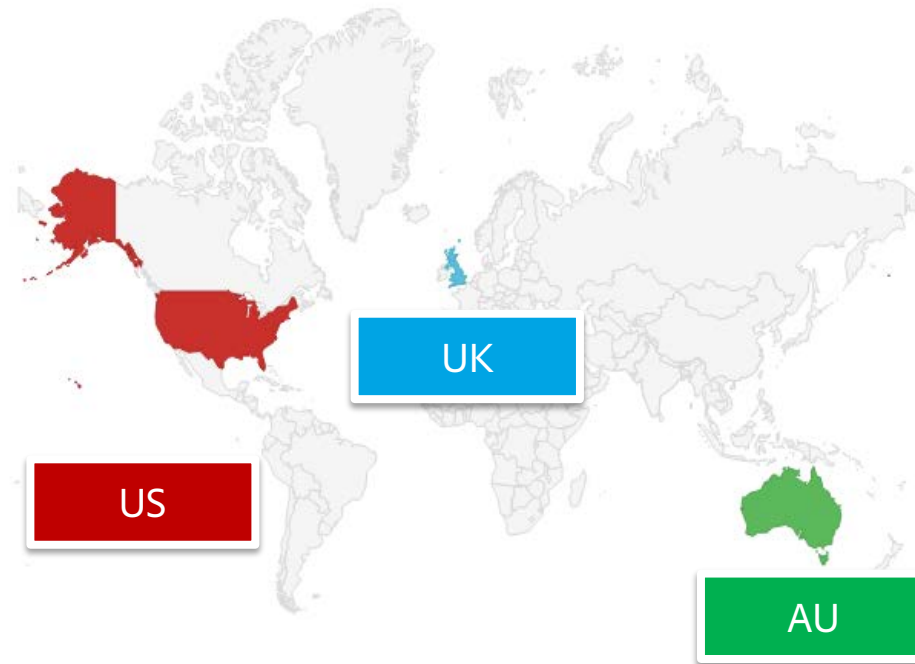
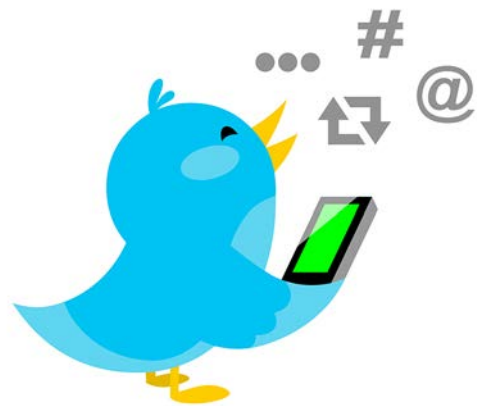


# Characterizing audience segments

- Location, age, gender and occupation
- Personal information access restricted
- Data tagged by location and occupation



# Segment-specific data collection





# Product-specific data collection

- Different modifiers with different products
- **1621** Tweets collected for product
- **1364** Tweets with positive sentiment
- **562** modifiers extracted
- **60** ads collected for skeletons



Adobe Creative Cloud





# Agenda

- Motivation
- Data
- **Method**
- Results



# Statistical language models

- Formal characterization of word usage
- Probability distribution over all patterns
- Efficient way of storage and representation
- Model: Word pairs with parts-of-speech
- Estimated from Tweet corpora

Segment specific text corpus



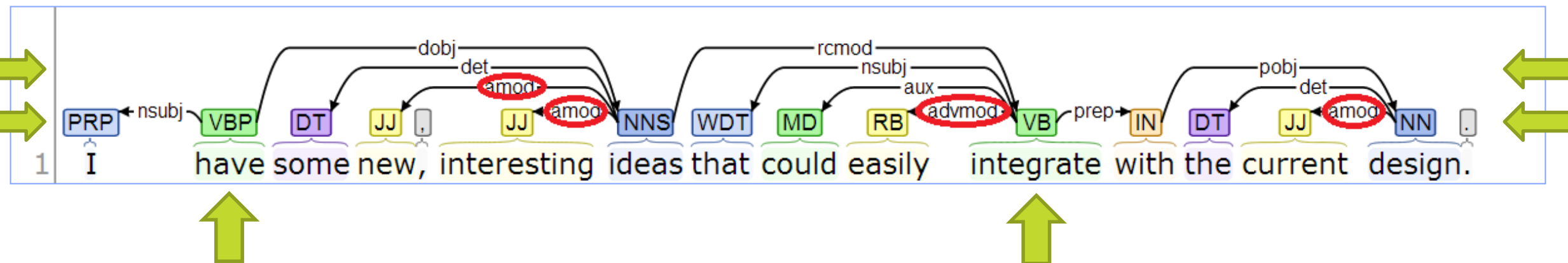
Extract patterns



Compute probability distributions

# Mining keyword-modifier pairs from segment-specific corpora

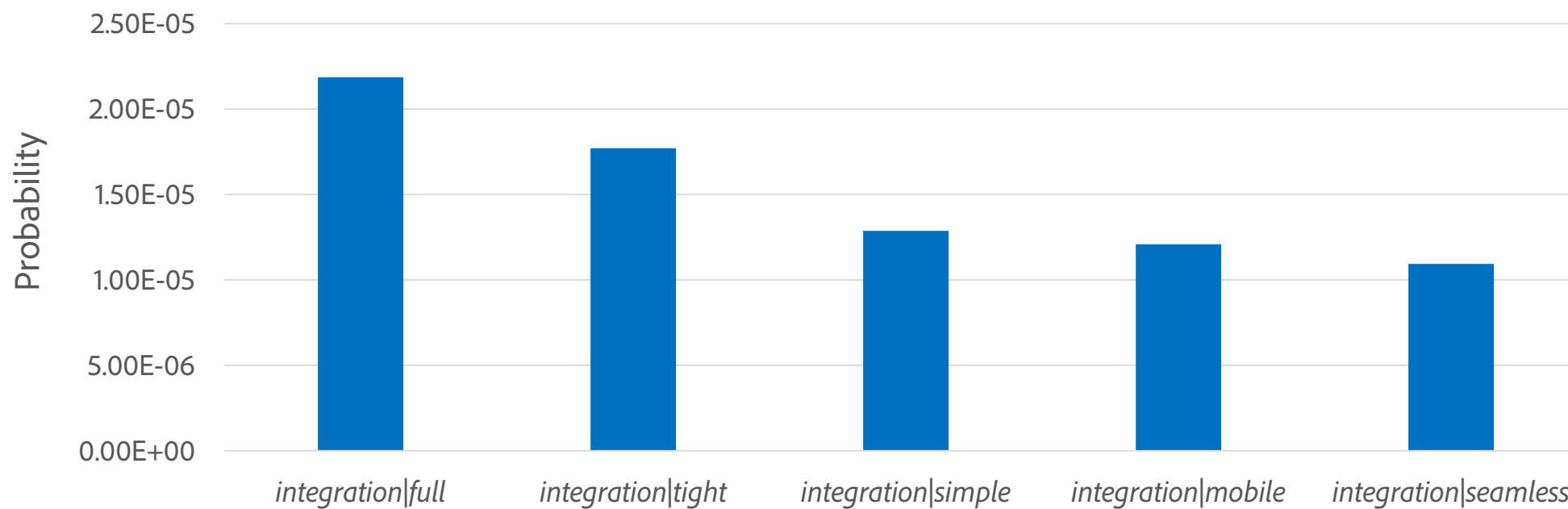
## Basic dependencies:



- Part-of-speech tagging
- Dependency parsing
- Lemmatization

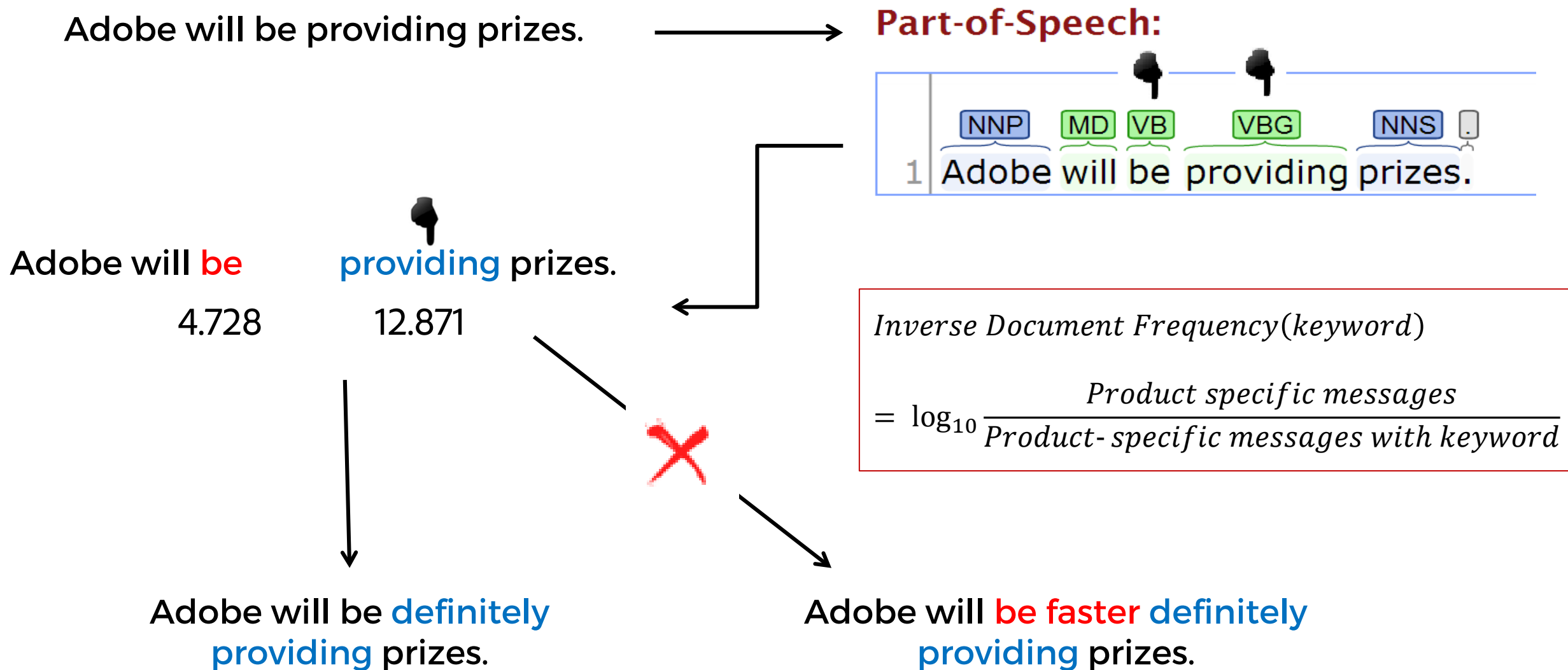
# Mining keyword-modifier pairs from segment-specific corpora

- **Keywords** ( $|k|$ ) from ad messages and **modifiers** ( $|m|$ ) from product-specific corpora
- Fetch adjective-noun pair probabilities and adverb-verb pair probabilities
- From each target language model for all **( $|k| \times |m|$ ) pairs**





# Avoiding too many suggestions: Term weight thresholds



- Where to insert new words?

## General trends

- Adjective before noun
- Adverb after verb

But what about ...

If you have been waiting to join eventually the club, ...

This includes also the plans on marketing ...

Convert your artworks to useful something ...

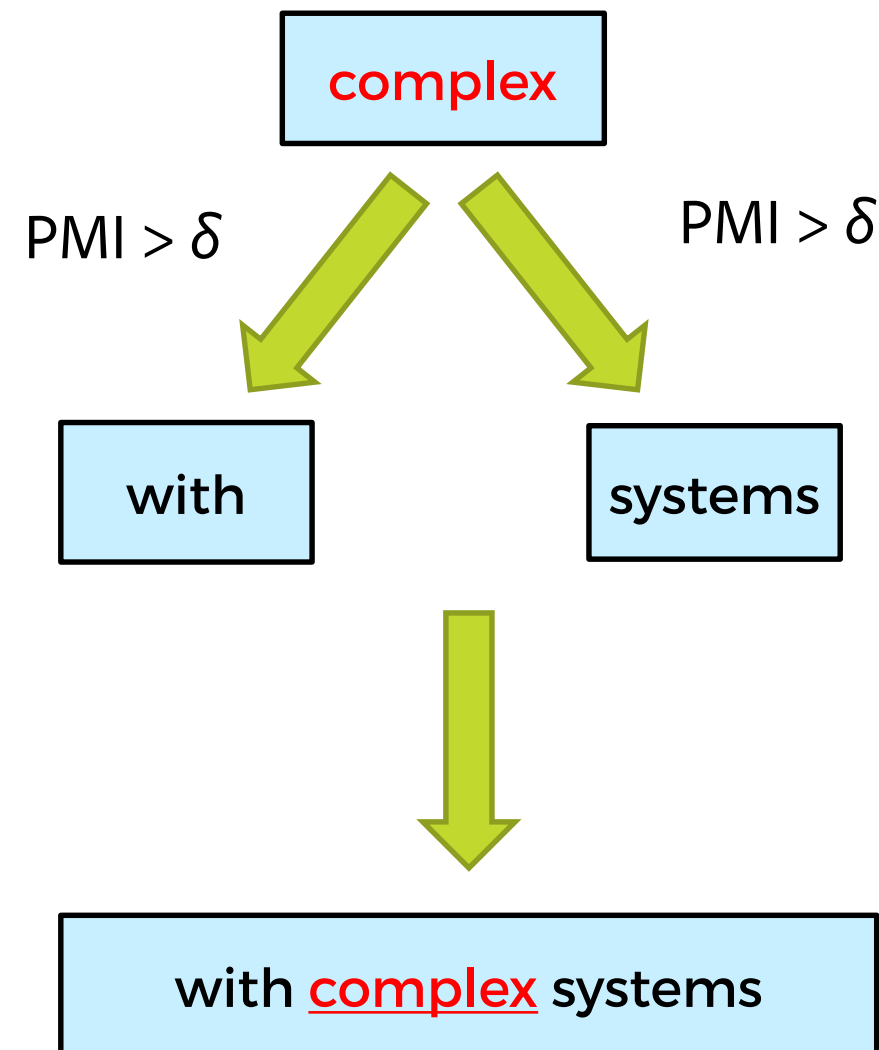
*Rules not good enough!*

# Pointwise mutual information

- Point-wise mutual information

$$PMI(x; y) = \log_2 \frac{p(x, y)}{p(x)p(y)}$$

- Probabilities from language models
- PMI thresholds control significance
- Left **and** right PMI thresholds must be satisfied





# Noun phrases are tricky

... license management tools ...



... license management tools ...

12.786    11.272    10.323



... **specific** license **new** management **free** tools ...

We provide license management tools.



Noun Phrase  
Chunker



We provide **license management tools**.

... license management tools ...



... **renewable** license management tools ...

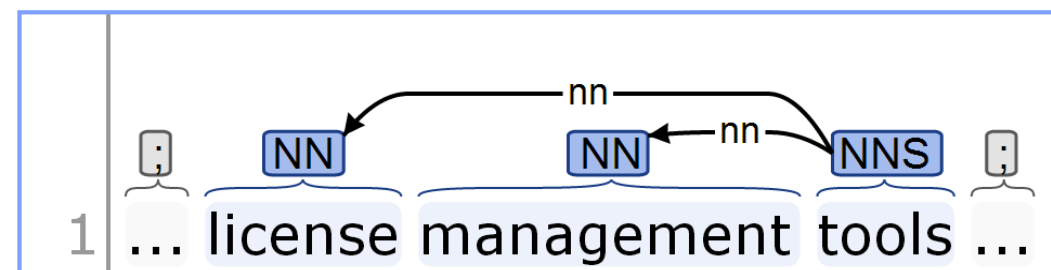


## Finding adjectives for nouns embedded inside noun phrases

Use dependencies to identify the **chunk-head** and find an adjective for it

## ... license management tools ...

## Basic dependencies:



... handy license management tools ...

... license management **tools** ...



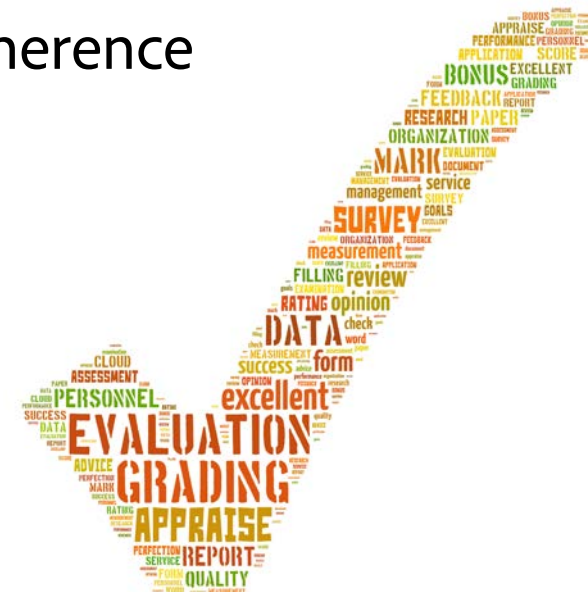
# Agenda

- Motivation
- Data
- Method
- **Results**



# Evaluation

- Statistical evaluation using cross-entropy
- Can approach target language models
- Crowdsourced evaluation through Amazon Mechanical Turk
- Machine transformed messages show syntactic and semantic coherence



# Representative transformation

Ad Skeleton	Transformation 1	Transformation 2
<p>Even as the landscape continues to change, MAX will remain the place to learn about generating graphics content for devices, and discovering about tools, development approaches, and formats. Learn about license management tools and all the things you wanted to know!</p>	<p>Even as the landscape continues to change <u>dramatically</u>, MAX will <u>always</u> remain the <u>first</u> place to learn about generating <u>original</u> graphics content for <u>mobile</u> devices, and discovering about tools, <u>unique</u> development approaches, and formats. Learn about <u>valuable</u> license management tools and all the <u>greatest</u> things you wanted to know!</p> <p>(Segment: US-student)</p>	<p>Even as the landscape continues to change <u>daily</u>, MAX will remain the <u>first</u> place to <u>quickly</u> learn about generating <u>adaptive</u> graphics content for devices, and discovering about tools, <u>critical</u> development approaches, and formats. Learn about <u>handy</u> license management tools and all the <u>best</u> things you wanted to know <u>right</u>!</p> <p>(Segment: AU-designer)</p>



# Experimental results on cross-entropy

Model	Unigram (Sentence)		Unigram (Frequency)		Bigram (Sentence)		Bigram (Frequency)	
Segment	Ads with drop	Mean % drop in CE	Ads with drop	Mean % drop in CE	Ads with drop	Mean % drop in CE	Ads with drop	Mean % drop in CE
AU-designer	60/60	11.65	60/60	10.82	60/60	10.17	60/60	46.08
AU-developer	60/60	11.58	60/60	10.74	60/60	10.09	60/60	46.04
AU-manager	60/60	11.10	60/60	10.34	60/60	9.64	60/60	45.81
AU-student	60/60	11.02	60/60	10.28	60/60	9.68	60/60	45.76
UK-designer	60/60	13.71	60/60	12.89	60/60	12.00	60/60	47.34
UK-developer	60/60	12.65	60/60	11.77	60/60	11.10	60/60	46.66
UK-manager	60/60	12.39	60/60	11.64	60/60	10.90	60/60	46.59
UK-student	60/60	10.63	60/60	9.86	60/60	9.27	60/60	45.51
US-designer	60/60	11.98	60/60	11.21	60/60	10.57	60/60	46.32
US-developer	60/60	12.27	60/60	11.43	60/60	10.71	60/60	46.44
US-manager	60/60	12.18	60/60	11.27	60/60	10.58	60/60	46.36
US-student	60/60	11.05	60/60	10.23	60/60	9.61	60/60	45.73



# Experimental results through crowdsourcing

- Presented human assessors with triplets of ad messages
- Human composed message hidden among two computer modified ones
- Evaluator asked to identify human message and score the other two
- Computer modified messages get comparable ratings as human compositions

Segment	Average Rating
Human	3.91
AU-designer	3.54
AU-developer	3.71
AU-manager	3.50
AU-student	3.57
UK-designer	3.50
UK-developer	3.69
UK-manager	3.55
UK-student	3.44
US-designer	3.50
US-developer	3.56
US-manager	3.58
US-student	3.54





## Next steps

- Improved techniques like synonym, hypernym, hyponym analysis
- Controlled insertions and substitutions
- Beyond parts of speech – Automatic transformations for:
  - Formal and informal styles
  - Statements, questions, interjections
  - Sentence lengths



# Not too far away?



Dear Ms. Navarro:

Enclosed please find information that was requested during our telephone communication of 10/15/14. As mentioned at that time, our facility has significantly more doctors of exceptional quality than any other facility in the state.

As you were also informed, our organization has an extensive network of doctors and other health-care professionals with offices located throughout the state. Additional information you should need a specialist, our professionals will be able to make an appropriate recommendation.

In the event that you have questions or would like additional information, you may certainly contact me during business hours.

Most sincerely yours,

Samuel G. Berenz

Hi Gabriella:

Hope all is well. Just sending along the information you asked for. As I said on Friday, our hospital has more and better doctors than any other hospital in the state.

We also have a large group of doctors and other health professionals with offices close to you at work or at home. Need to know more? They'll refer you to the right one.

Just give me a ring if you want to know more. Any time from 9:00 to 5:00 should be fine.

Take care,

Sam





# Questions?

