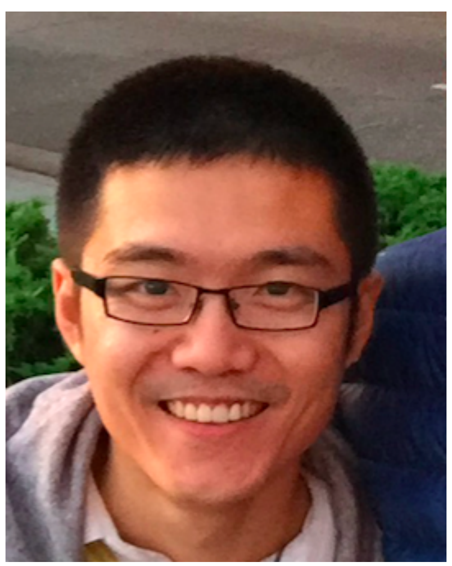
What's so Hard about Natural Language Understanding?

Alan Ritter
Computer Science and Engineering
The Ohio State University

Collaborators:
Jiwei Li, Dan Jurafsky (Stanford)
Bill Dolan, Michel Galley, Jianfeng Gao (MSR), Colin Cherry (Google)
Jeniya Tabassum (Ohio State), Alexander Konovalov (Ohio State), Wei Xu (Ohio State)
Brendan O'Connor (Umass)

THE OHIO STATE UNIVERSITY

What's so Hard about Natural Language Understanding?

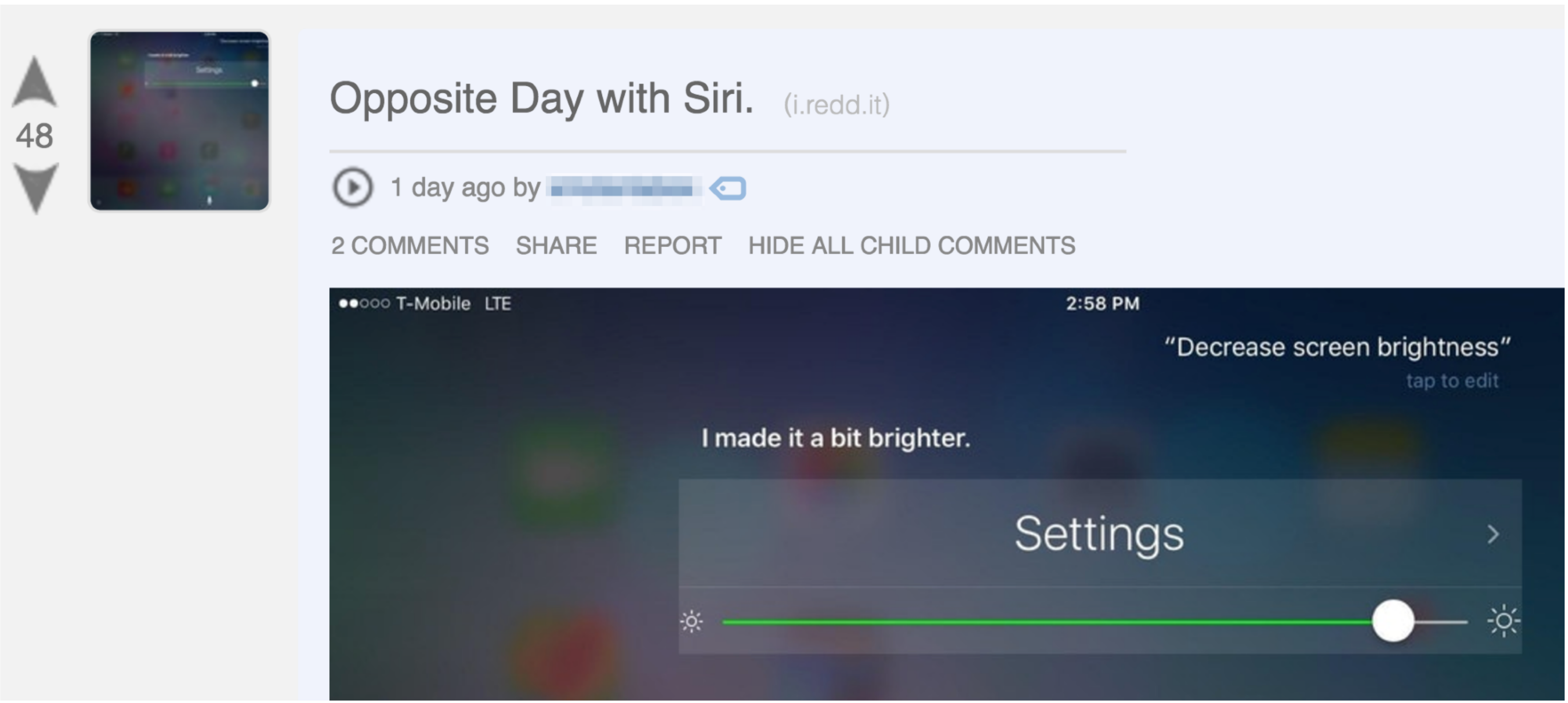


Alan Ritter
Computer Science and Engineering
The Ohio State University

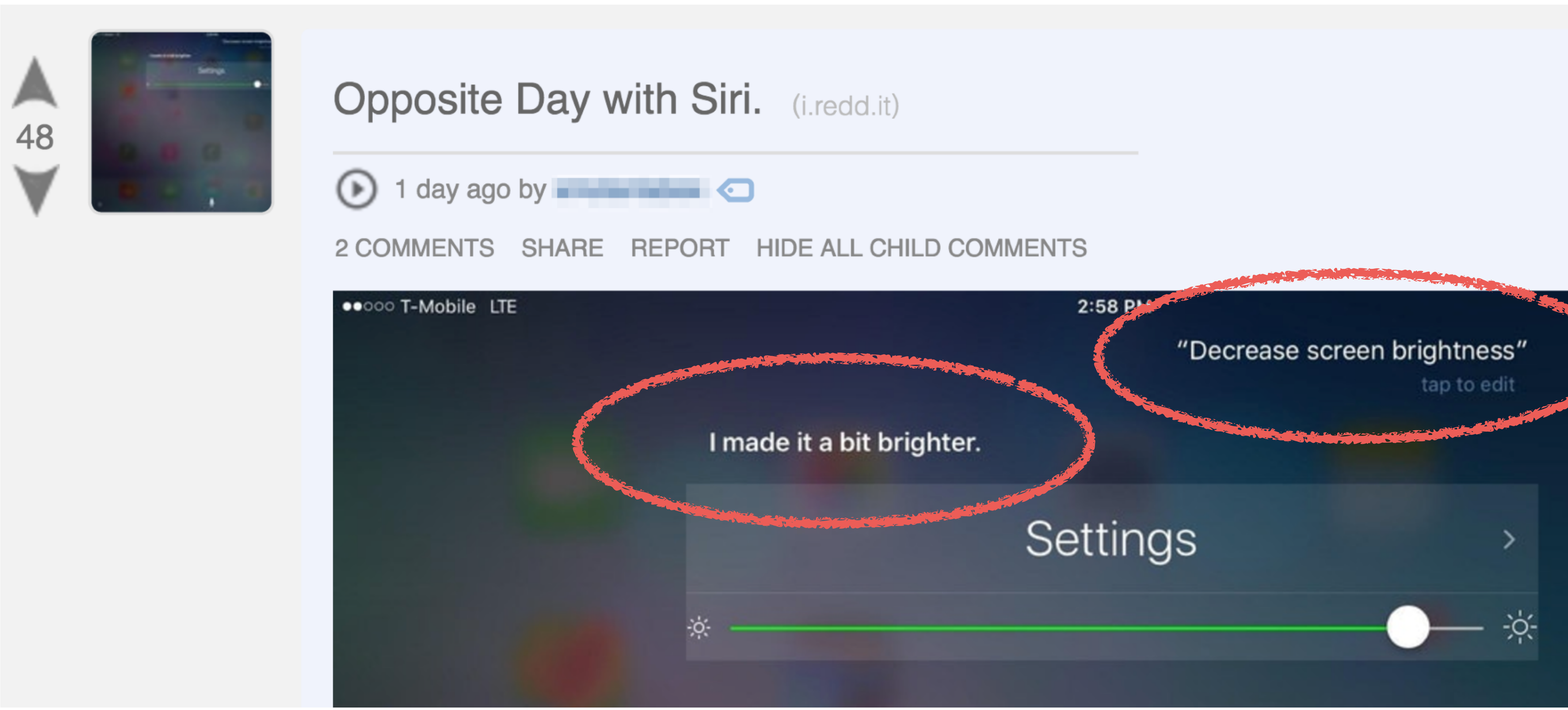
Collaborators:
Jiwei Li, Dan Jurafsky (Stanford)
Bill Dolan, Michel Galley, Jianfeng Gao (MSR), Colin Cherry (Google)
Jeniya Tabassum (Ohio State), Alexander Konovalov (Ohio State), Wei Xu (Ohio State)
Brendan O'Connor (Umass)

THE OHIO STATE UNIVERSITY

reddit /R/SIRIFAIL



reddit /R/SIRIFAIL



Q: Why are we so good at Speech, MT (but bad at NLU)?

People ***naturally*** translate and transcribe.

Q: Why are we so good at Speech, MT (but bad at NLU)?

People ***naturally*** translate and transcribe.

Q: Large, End-to-End Datasets for NLU?

- Web-scale Conversations?
- Web-scale Structured Data?

Q: Why are we so good at Speech, MT (but bad at NLU)?

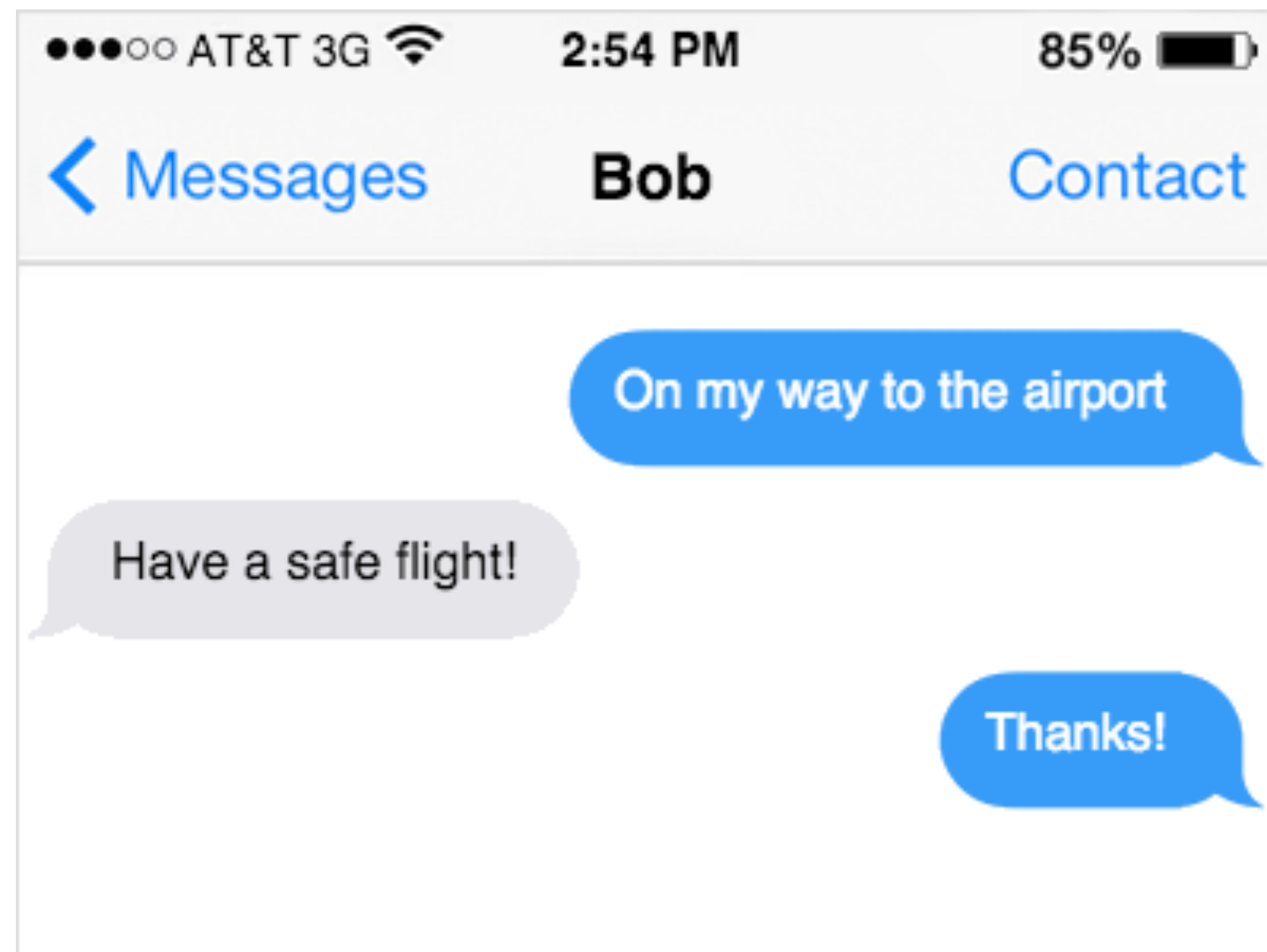
People ***naturally*** translate and transcribe.

Q: Large, End-to-End Datasets for NLU?

- Web-scale Conversations?
- Web-scale Structured Data?

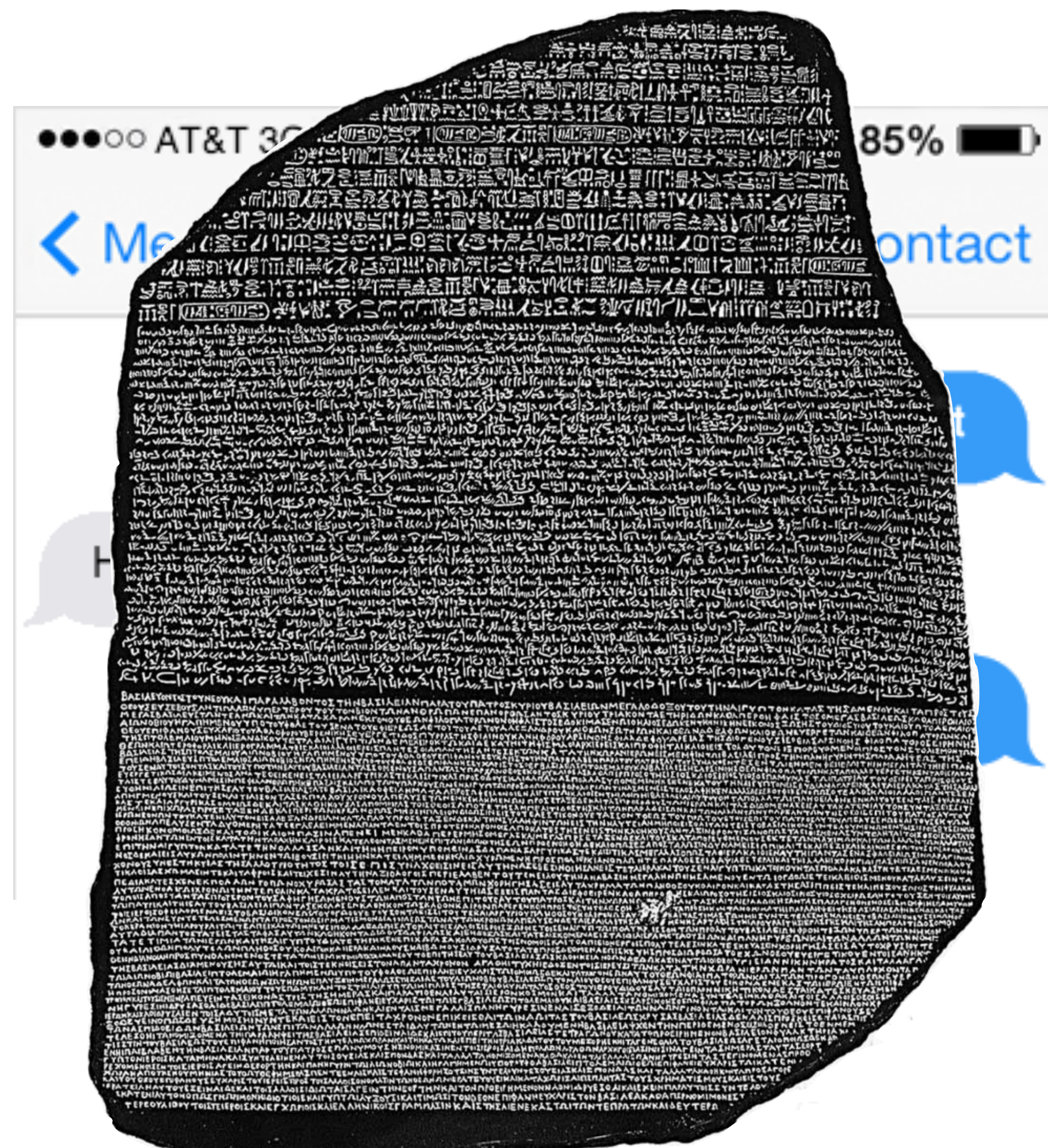
Data-Driven Conversation

- **Twitter:** ~ 500 Million Public SMS-Style Conversations ***per Month***
- **Goal:** Learn conversational agents directly from massive volumes of data.



Data-Driven Conversation

- **Twitter:** ~ 500 Million Public SMS-Style Conversations ***per Month***
- **Goal:** Learn conversational agents directly from massive volumes of data.



[Ritter, Cherry, Dolan EMNLP 2011]

Noisy Channel Model

Input:

Who wants to come over for dinner tomorrow?

[Ritter, Cherry, Dolan EMNLP 2011]

Noisy Channel Model

Input:

Who wants to come over $\underbrace{\text{for dinner}}$ tomorrow?

Output:

Yum ! I

[Ritter, Cherry, Dolan EMNLP 2011]

Noisy Channel Model

Input:

Who wants to come over for dinner tomorrow?

Output:

Yum ! I   want to

[Ritter, Cherry, Dolan EMNLP 2011]

Noisy Channel Model

Input:

Who wants to come over for dinner tomorrow?

Output:

Yum ! I   want to   be there

[Ritter, Cherry, Dolan EMNLP 2011]

Noisy Channel Model

Input:

Who wants to come over for dinner tomorrow?

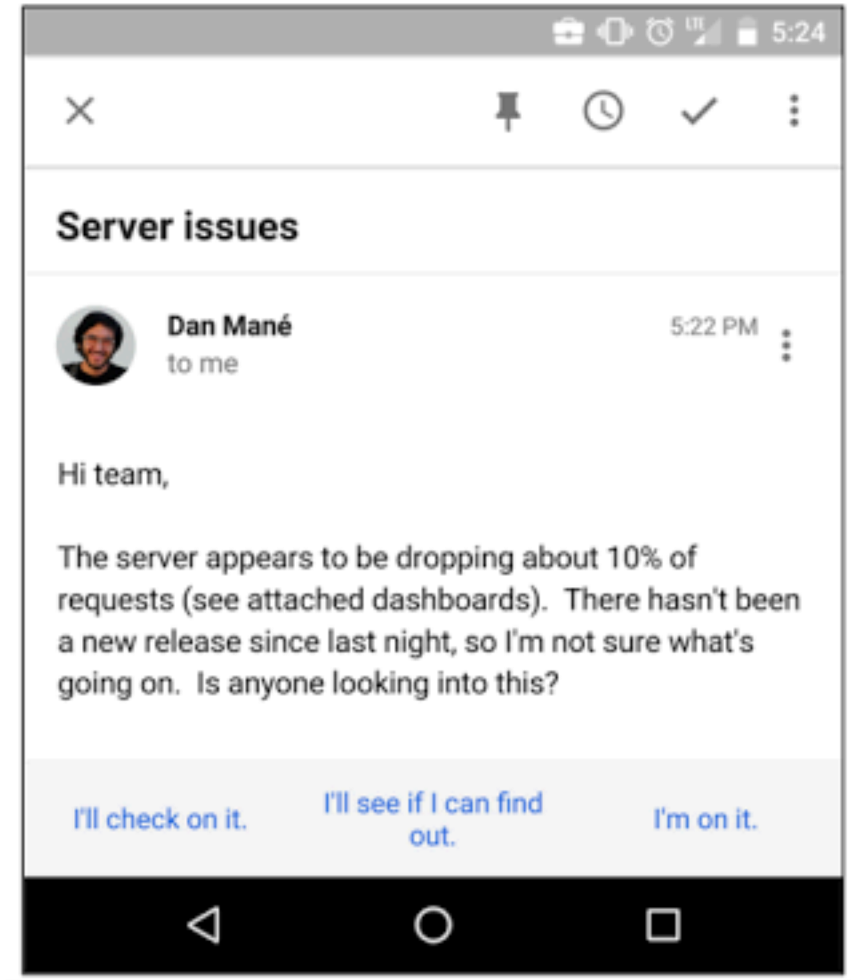
Output:

Yum ! I　want to　be there　tomorrow !

Neural Conversation

[Sordoni et. al. 2015] [Xu et. al. 2016] [Wen et. al. 2016]

[Li et. al. 2016] [Kannan et. al. 2016] [Serban et. al. 2016]



Google Research Blog

Computer, respond to this email.

Tuesday, November 03, 2015

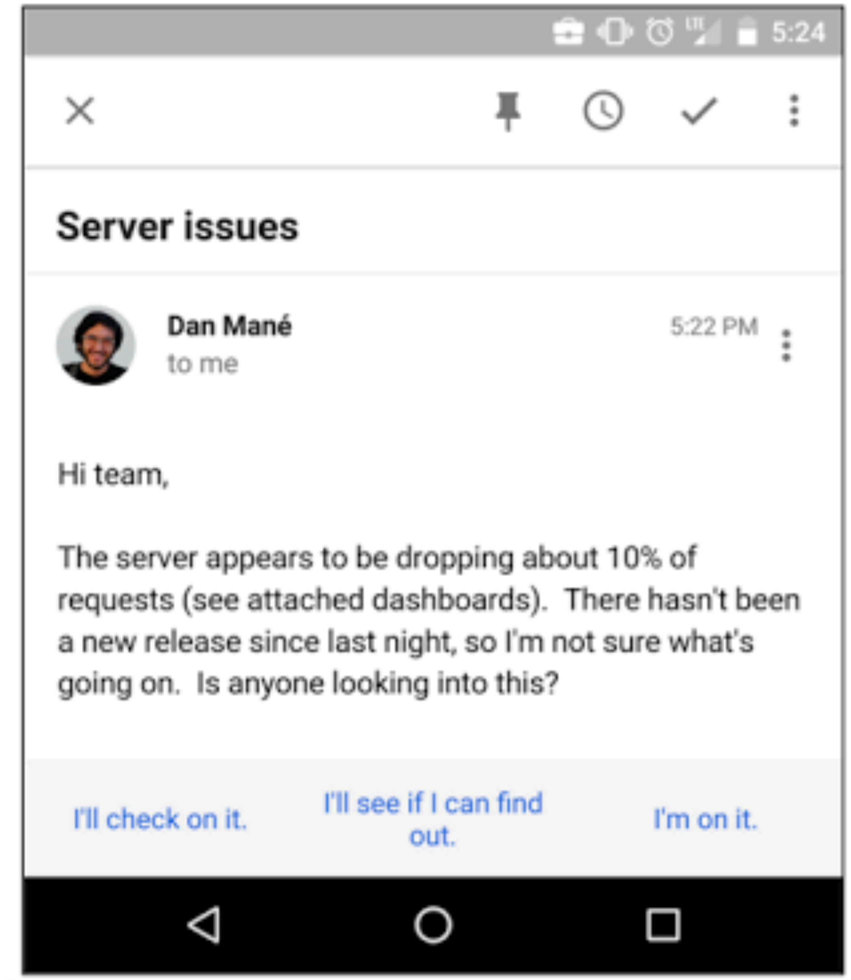
Posted by Greg Corrado*, Senior Research Scientist

Another bizarre feature of our early prototype was its propensity to respond with "I love you" to seemingly anything. As adorable as this sounds, it wasn't really what we were hoping for. Some analysis revealed that the system was doing exactly what we'd trained it to do, generate likely responses -- and it turns out that responses like "Thanks", "Sounds good", and "I love you" are super common -- so the system would lean on them as a safe bet if it was unsure. Normalizing the

Neural Conversation

[Sordoni et. al. 2015] [Xu et. al. 2016] [Wen et. al. 2016]

[Li et. al. 2016] [Kannan et. al. 2016] [Serban et. al. 2016]



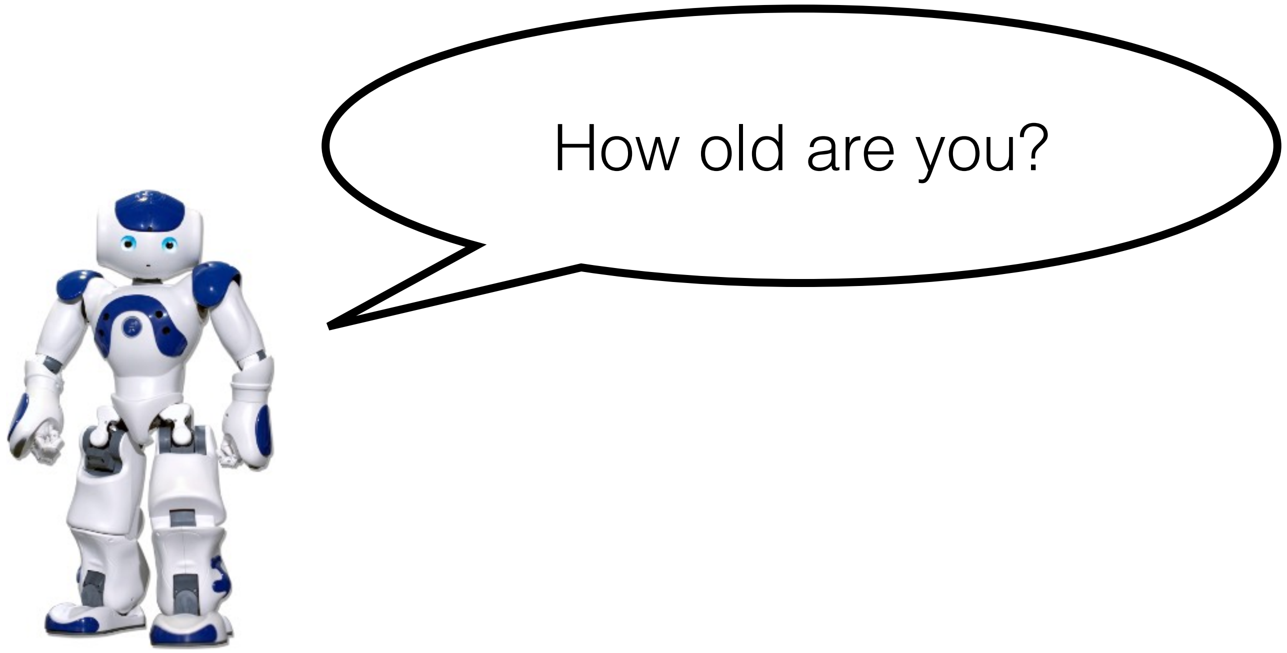
Google Research Blog

Computer, respond to this email.

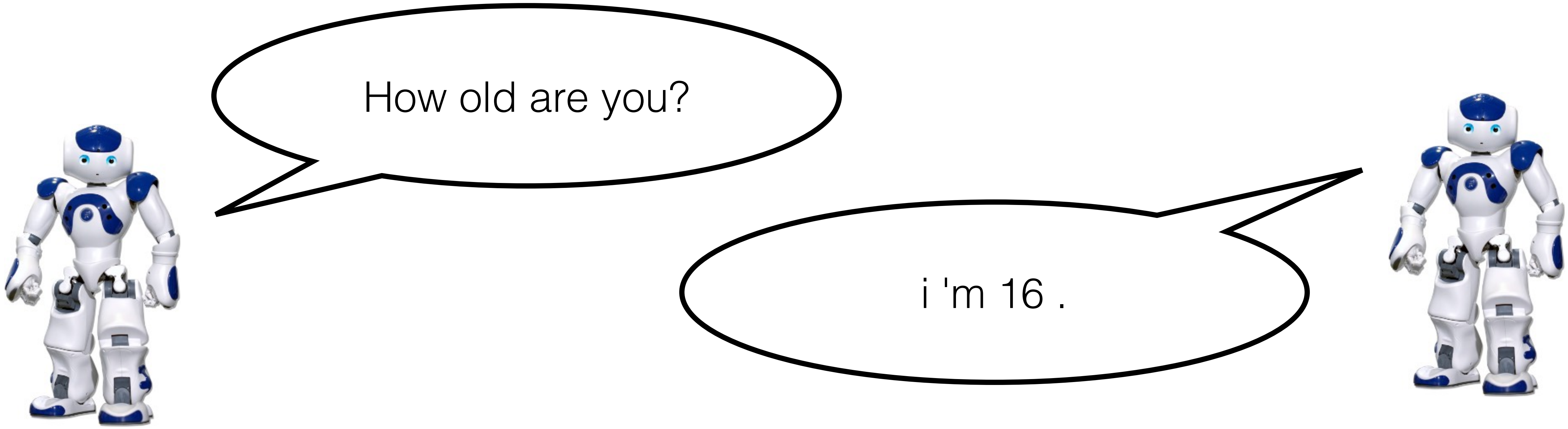
Tuesday, November 03, 2015

Posted by Greg Corrado*, Senior Research Scientist

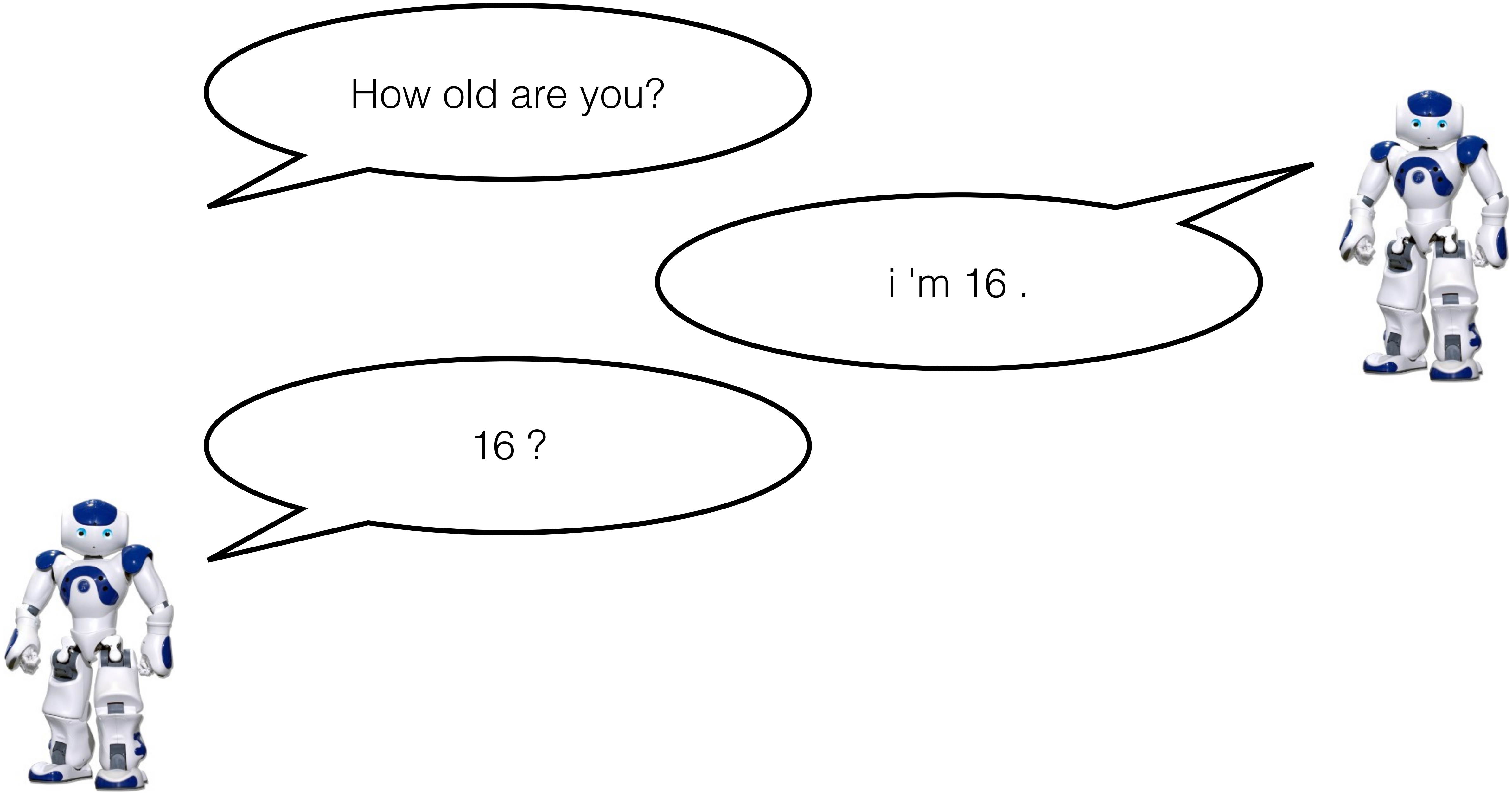
Another bizarre feature of our early prototype was its propensity to respond with “I love you” to seemingly anything. As adorable as this sounds, it wasn’t really what we were hoping for. Some analysis revealed that the system was doing exactly what we’d trained it to do, generate likely responses -- and it turns out that responses like “Thanks", "Sounds good", and “I love you” are super common -- so the system would lean on them as a safe bet if it was unsure. Normalizing the



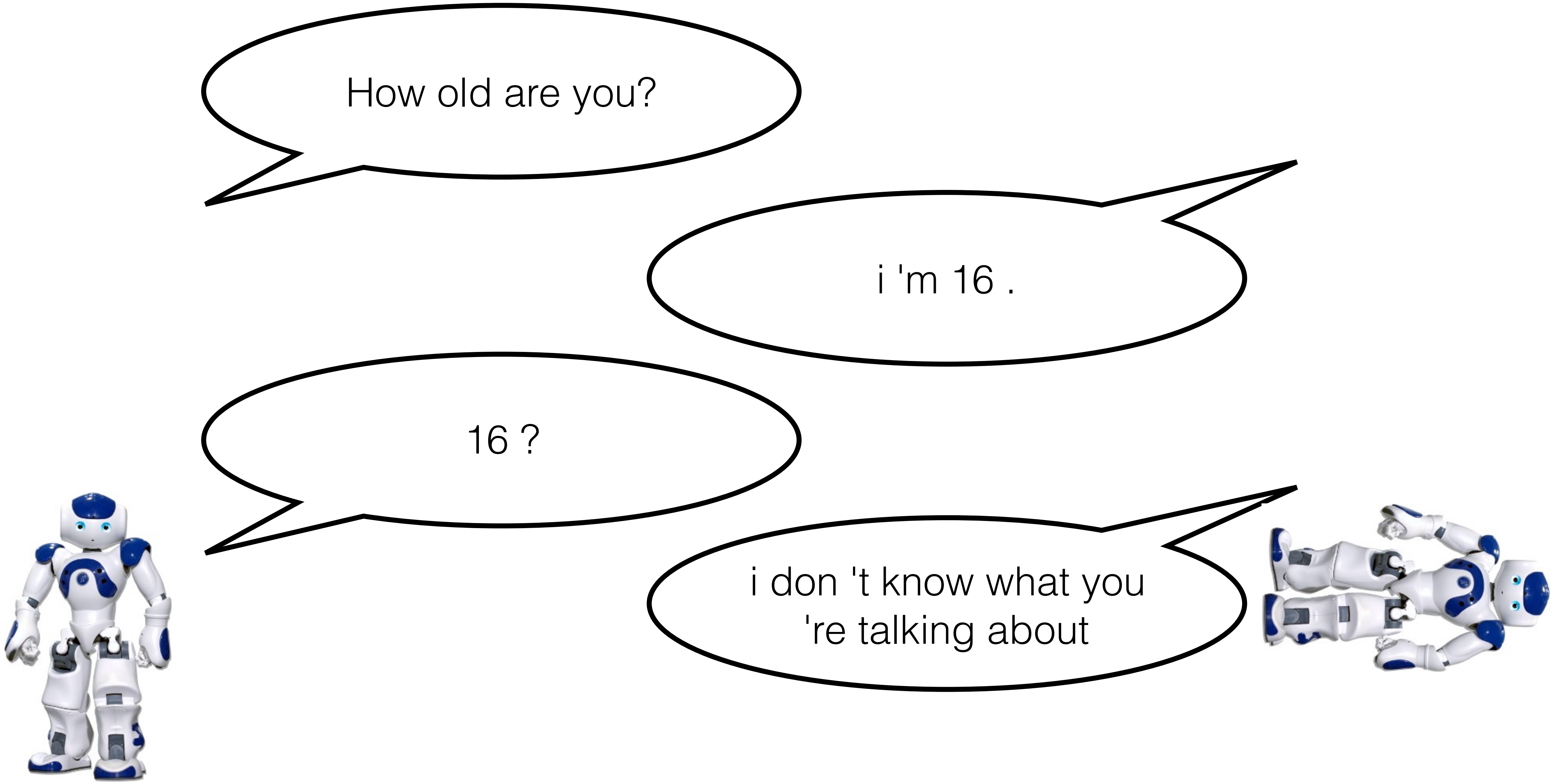
Slide Credit: Jiwei Li



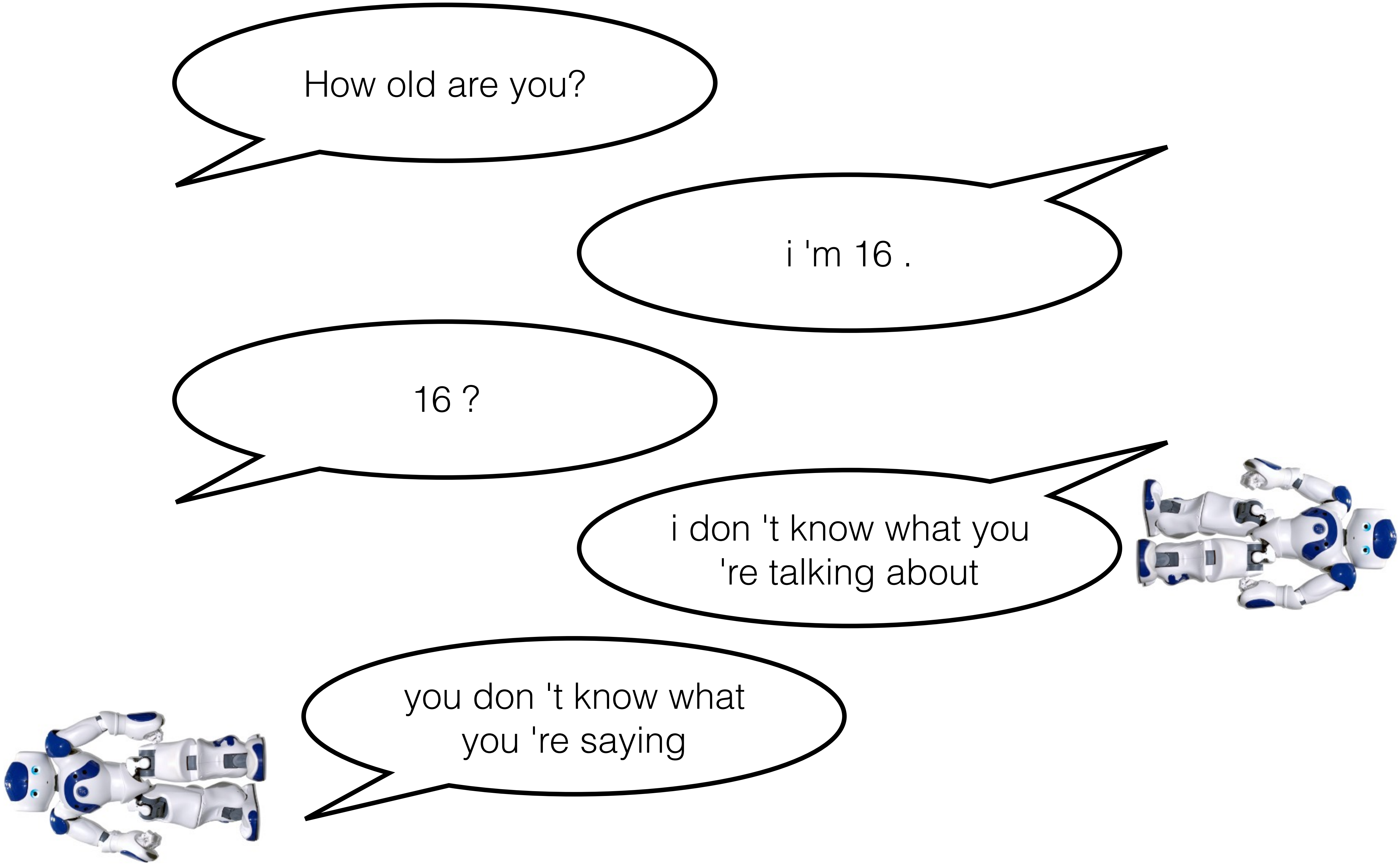
Slide Credit: Jiwei Li



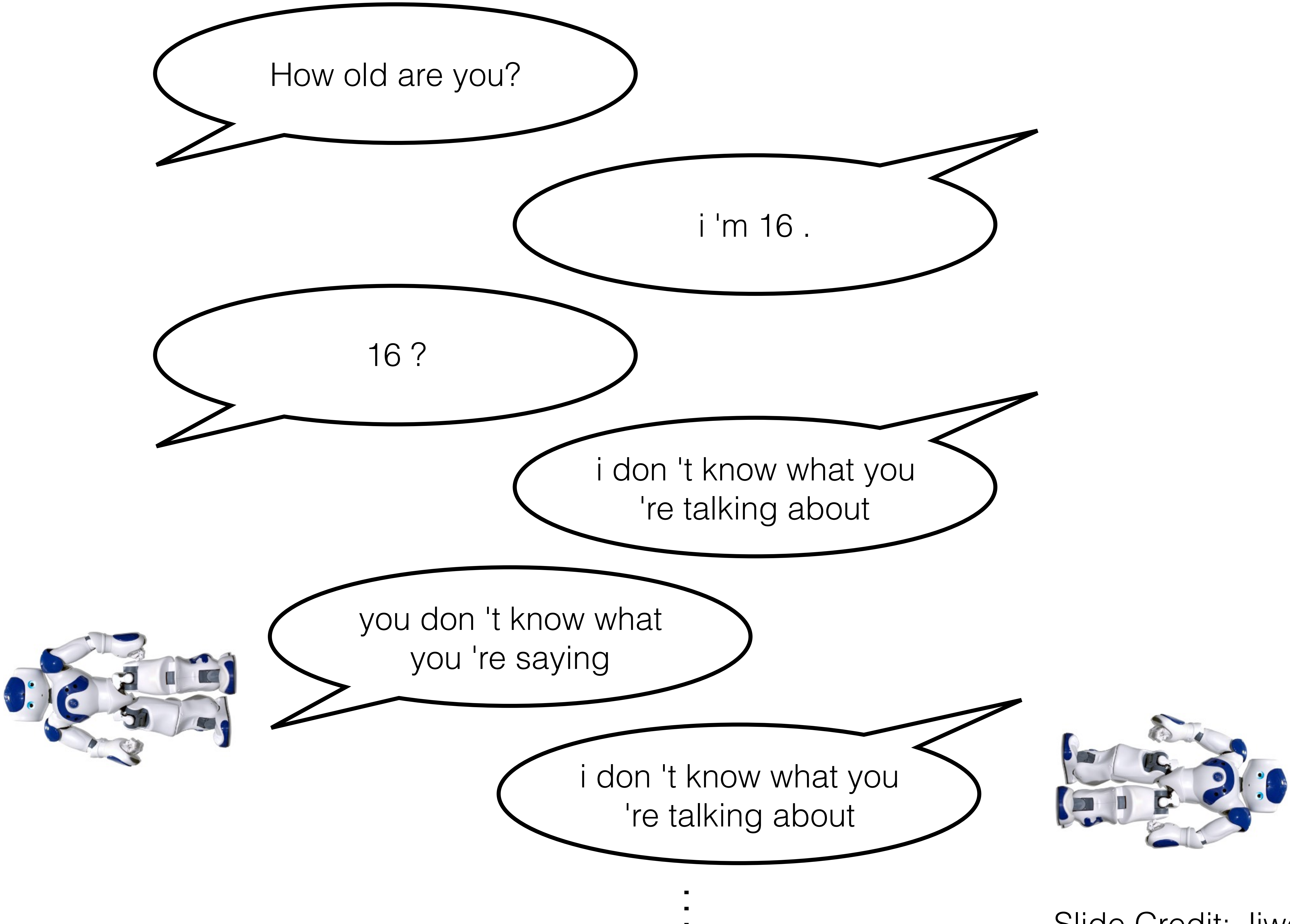
Slide Credit: Jiwei Li



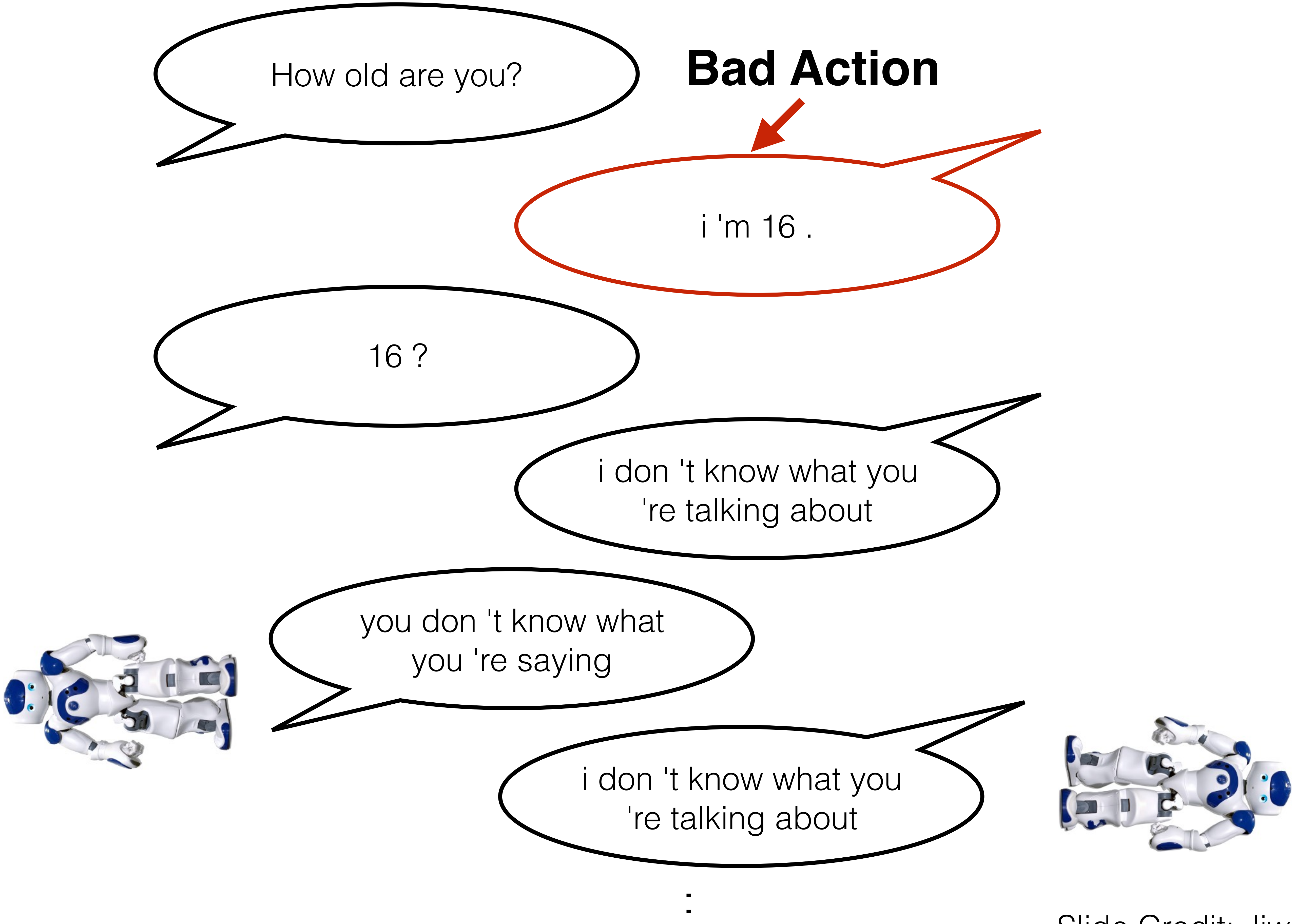
Slide Credit: Jiwei Li



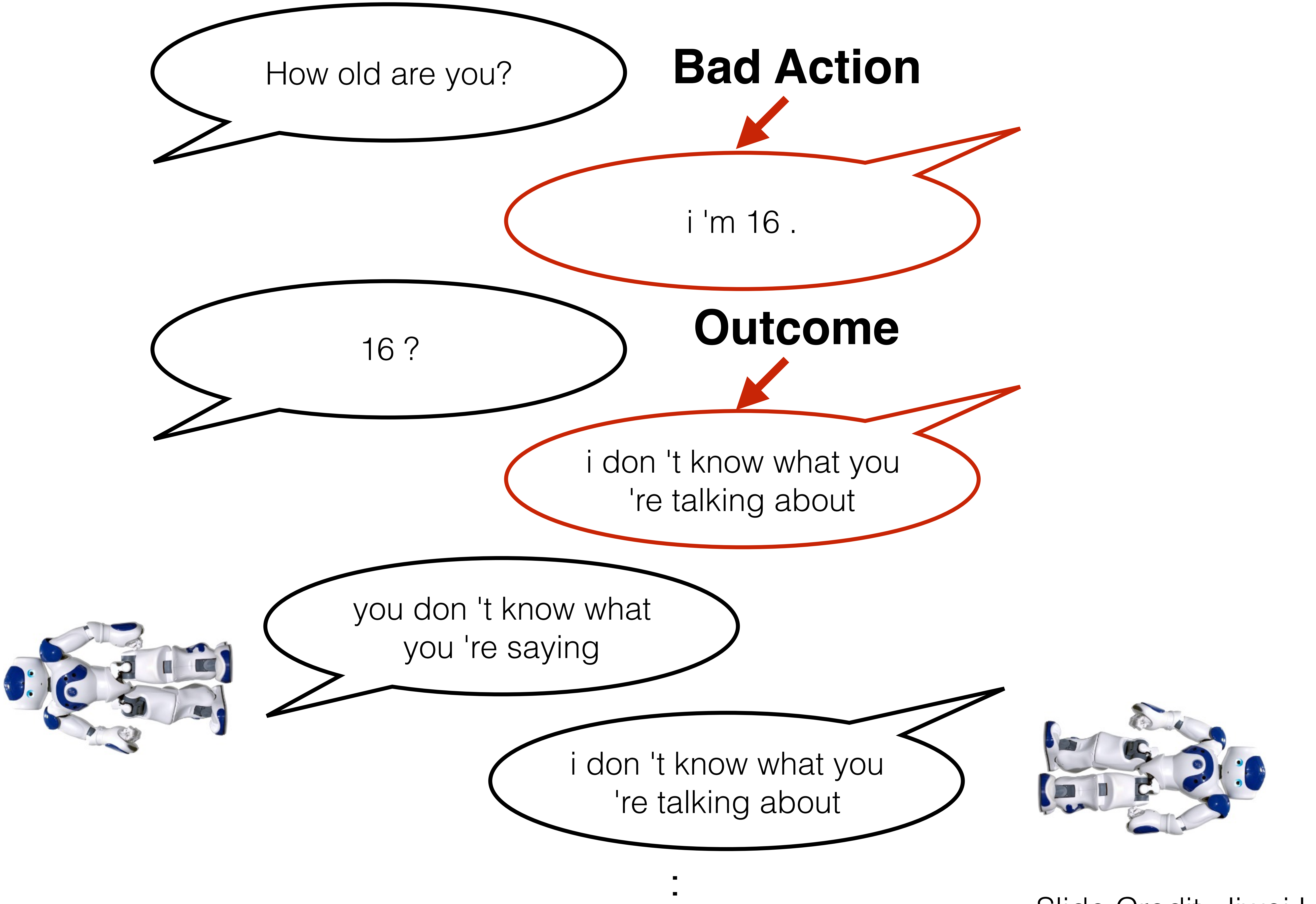
Slide Credit: Jiwei Li



Slide Credit: Jiwei Li



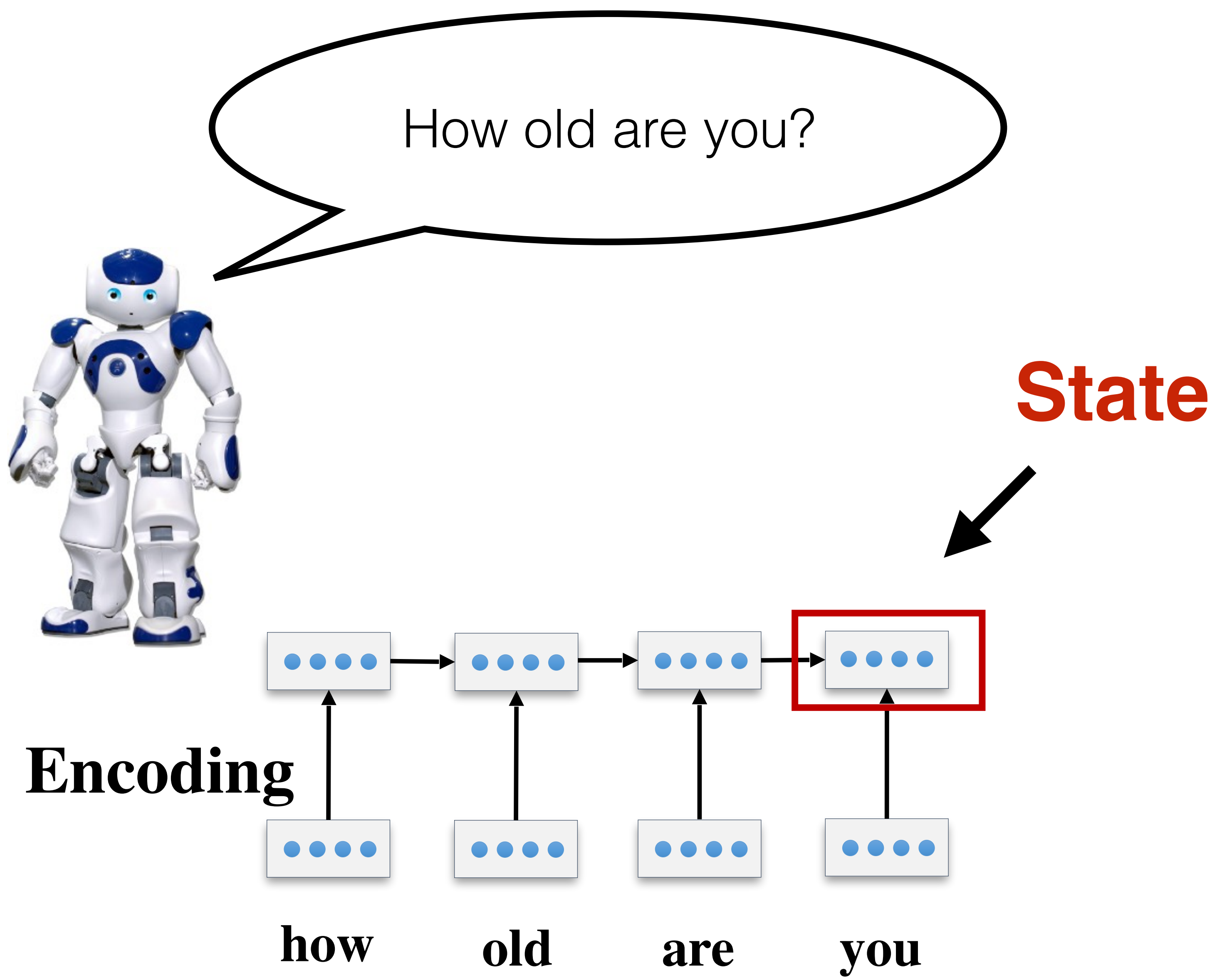
Slide Credit: Jiwei Li



Slide Credit: Jiwei Li

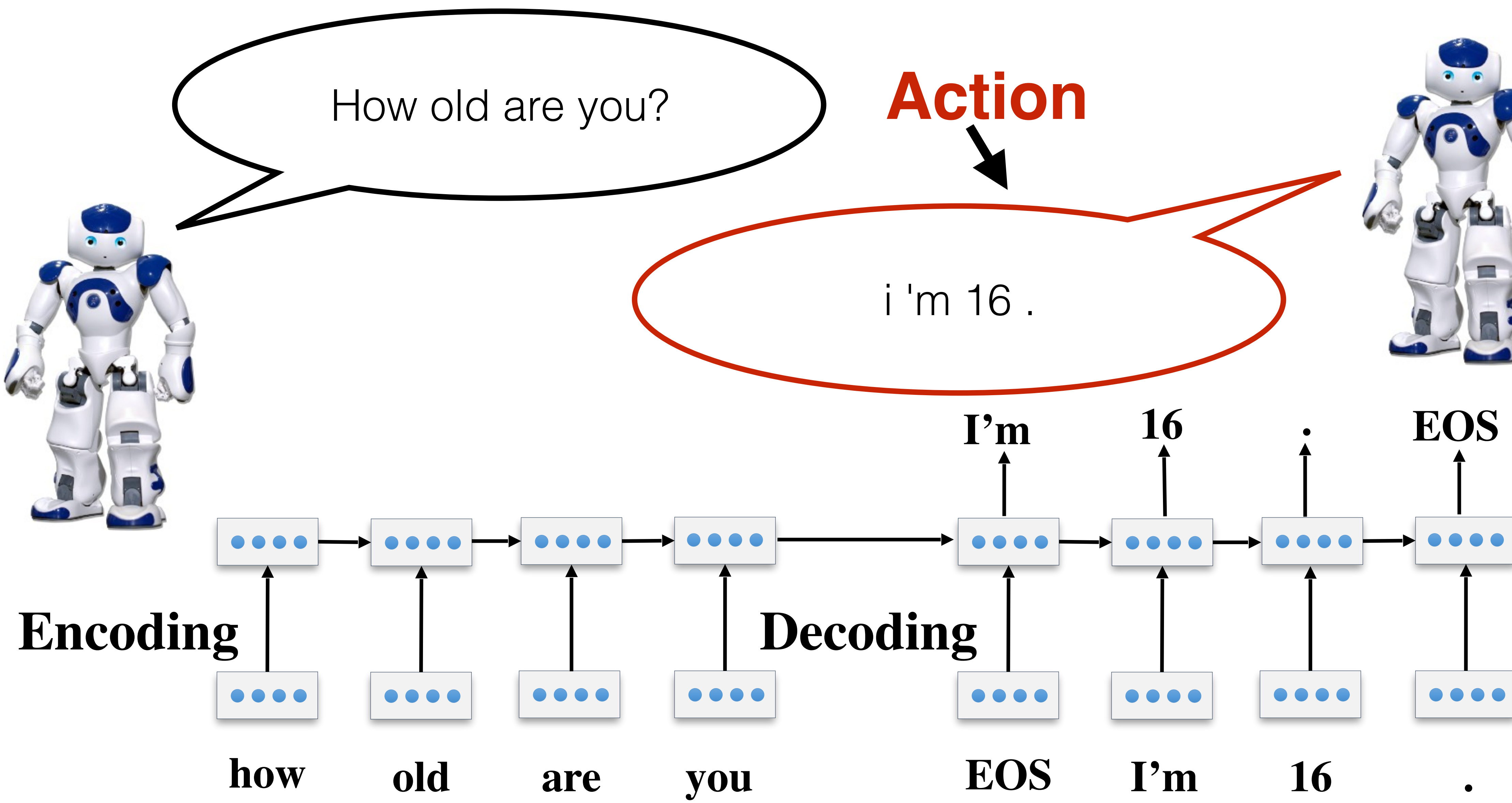
Deep Reinforcement Learning

[Li, Monroe, Ritter, Galley, Gao, Jurafsky EMNLP 2016]



Deep Reinforcement Learning

[Li, Monroe, Ritter, Galley, Gao, Jurafsky EMNLP 2016]



Learning: Policy Gradient

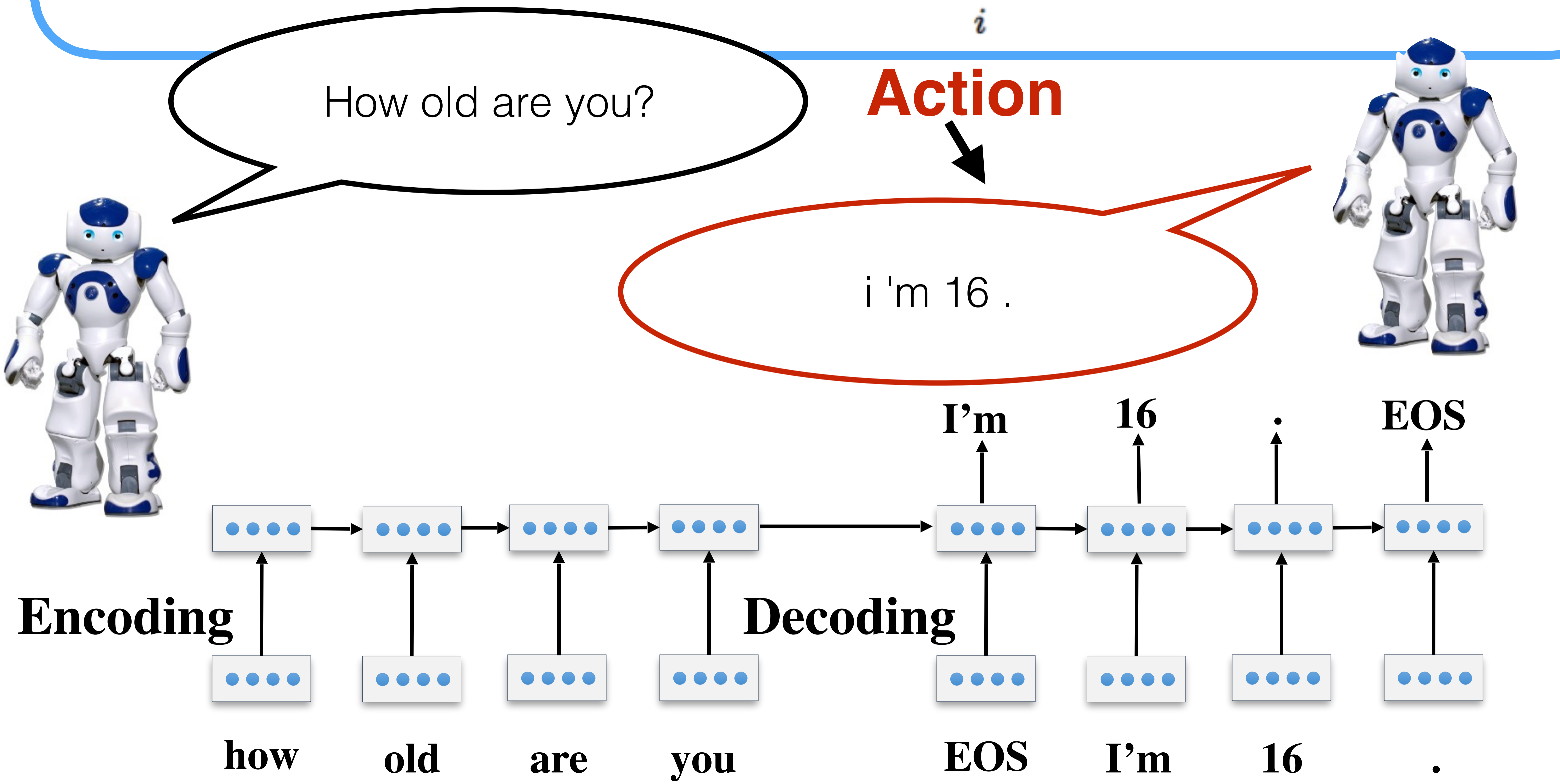
REINFORCE Algorithm (Williams,1992)

$$J(\theta) = \mathbb{E}[R(s_1, s_2, ..., s_N)]$$

$$\nabla J(\theta) = \nabla \log p(s_1, s_2, ..., s_N) R(s_1, s_2, ..., s_N)$$

What we want to learn

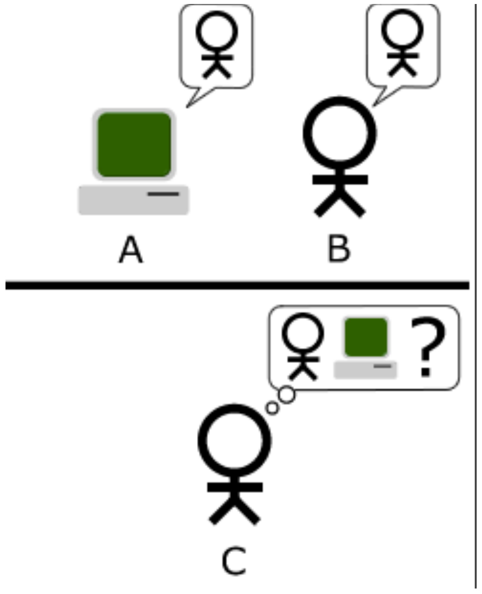
$$\nabla J(\theta) = \nabla \log \prod_i \boxed{p(s_i|s_{i-1})} R(s_1, s_2, ..., s_N)$$

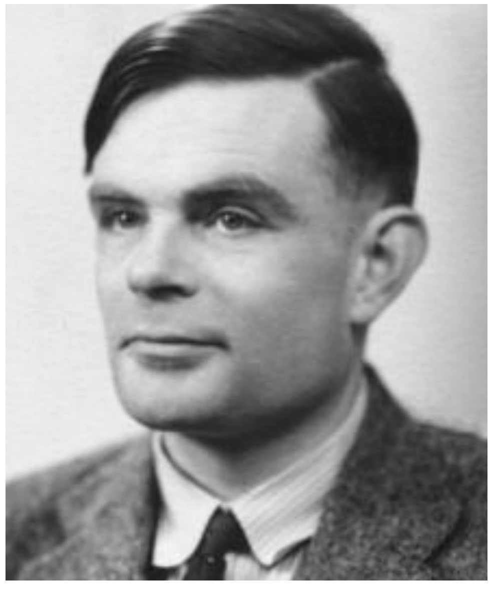


Q: Rewards?

Q: Rewards?

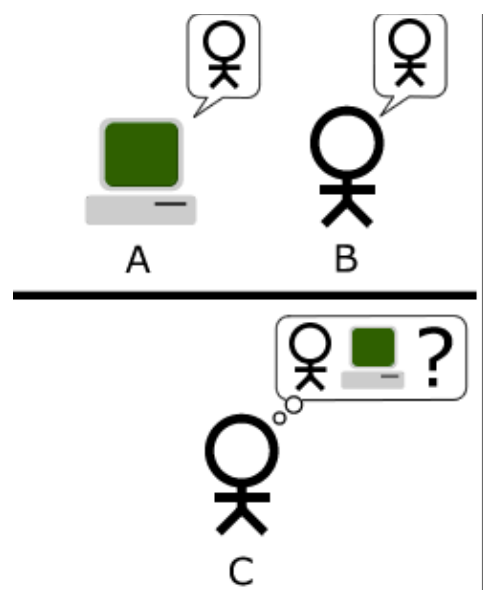
A: Turing Test

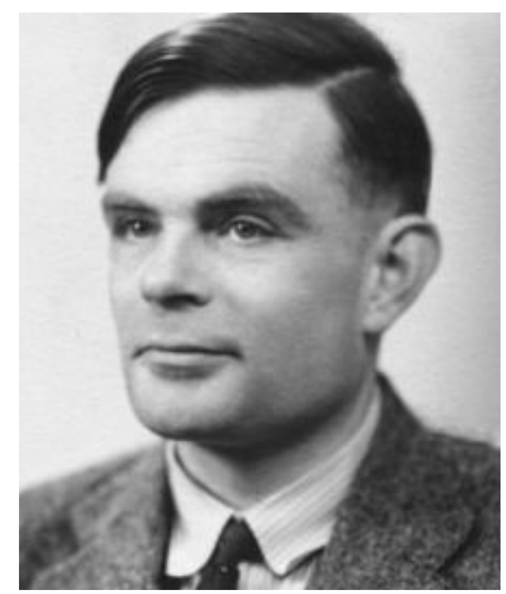




Q: Rewards?

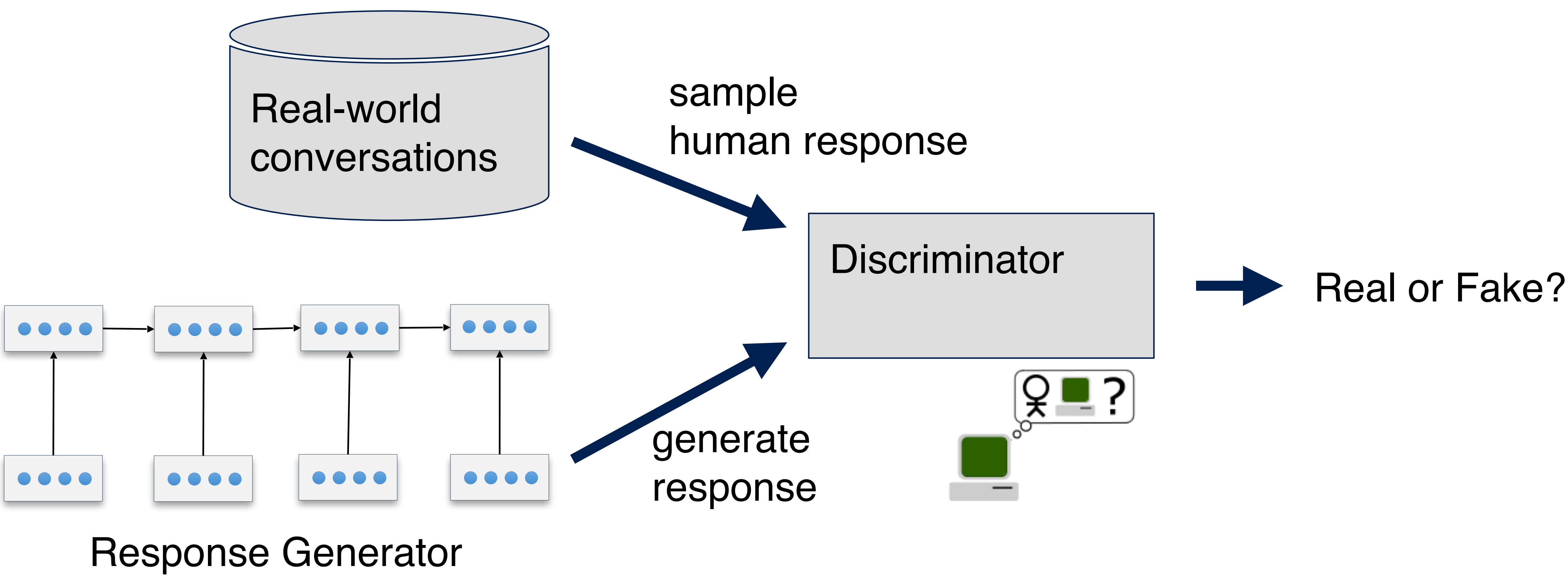
A: Turing Test





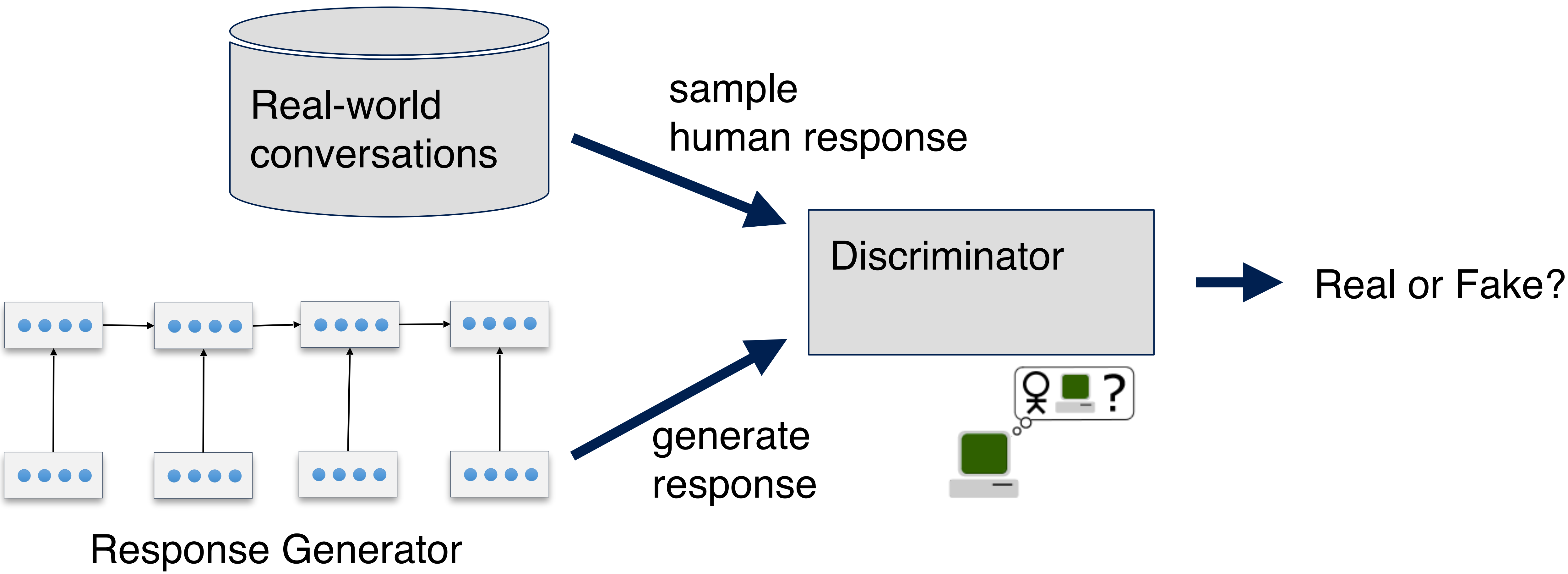
Adversarial Learning
(Goodfellow et al., 2014)

Adversarial Learning for Neural Dialogue [Li, Monroe, Shi, Jean, Ritter, Jurafsky EMNLP 2016]



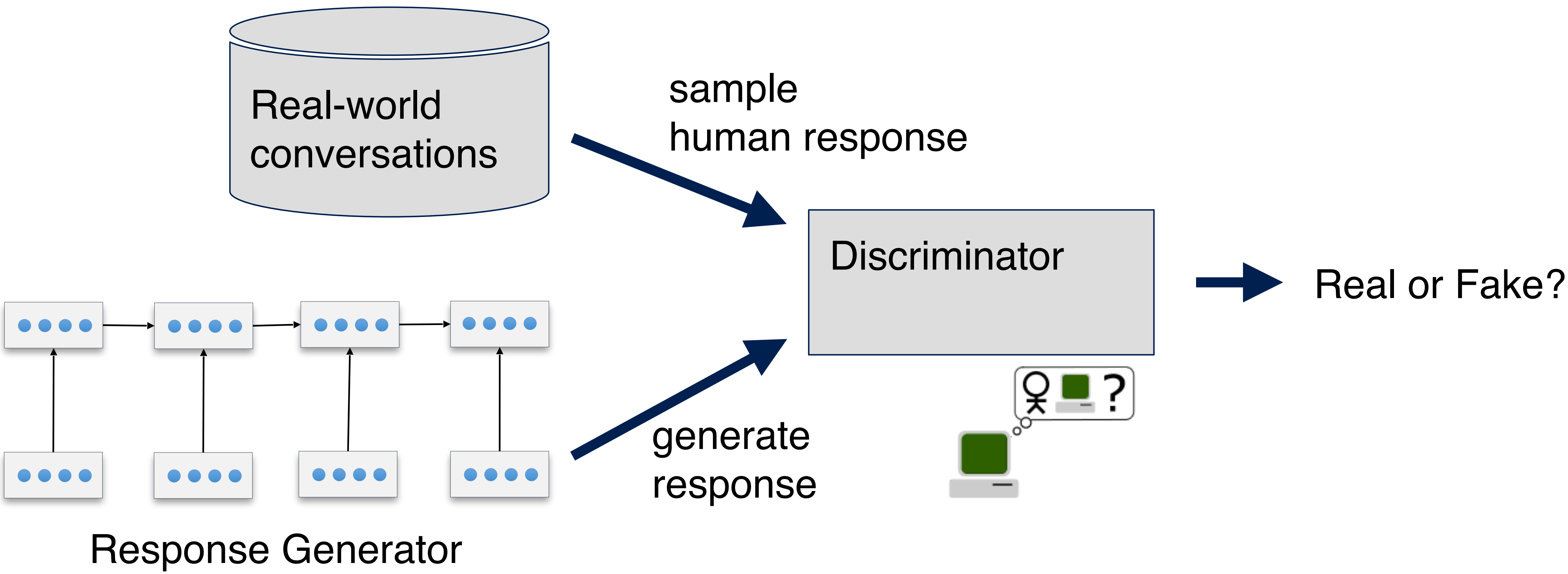
Adversarial Learning for Neural Dialogue [Li, Monroe, Shi, Jean, Ritter, Jurafsky EMNLP 2016]

(Alternate Between Training Generator and Discriminator)



Adversarial Learning for Neural Dialogue [Li, Monroe, Shi, Jean, Ritter, Jurafsky EMNLP 2016]

(Alternate Between Training Generator and Discriminator)



REINFORCE Algorithm (Williams,1992)

Adversarial Learning Improves Response Generation

Human Evaluator:

vs vanilla generation model

Adversarial Win	Adversarial Lose	Tie
62%	18%	20%

Machine Evaluator:
[Bowman et. al. 2016]

Adversarial Success (How often can you fool a machine)

Adversarial Learning	8.0%
Standard Seq2Seq model	4.9%

Slide Credit: Jiwei Li

Q: Why are we so good at Speech, MT (but bad at NLU)?

People ***naturally*** translate and transcribe.

Q: Large, End-to-End Datasets for NLU?

- Web-scale Conversations?
- Web-scale Structured Data?

Q: Why are we so good at Speech, MT (but bad at NLU)?

People ***naturally*** translate and transcribe.

Q: Large, End-to-End Datasets for NLU?

- Web-scale Conversations?

Generates fluent open domain replies

- Web-scale Structured Data?

Q: Why are we so good at Speech, MT (but bad at NLU)?

People ***naturally*** translate and transcribe.

Q: Large, End-to-End Datasets for NLU?

- Web-scale Conversations?
- Web-scale Structured Data?

Generates fluent open domain replies

Really Natural Language Understanding?

Q: Why are we so good at Speech, MT (but bad at NLU)?

People ***naturally*** translate and transcribe.

Q: Large, End-to-End Datasets for NLU?

- Web-scale Conversations?
- Web-scale Structured Data?

Generates fluent open domain replies

Really Natural Language Understanding?

Learning from **Distant Supervision** [Mintz et. al. 2009]

1) Named Entity Recognition

Challenge: highly ambiguous labels
[Ritter, et. al. EMNLP 2011]

2) Relation Extraction

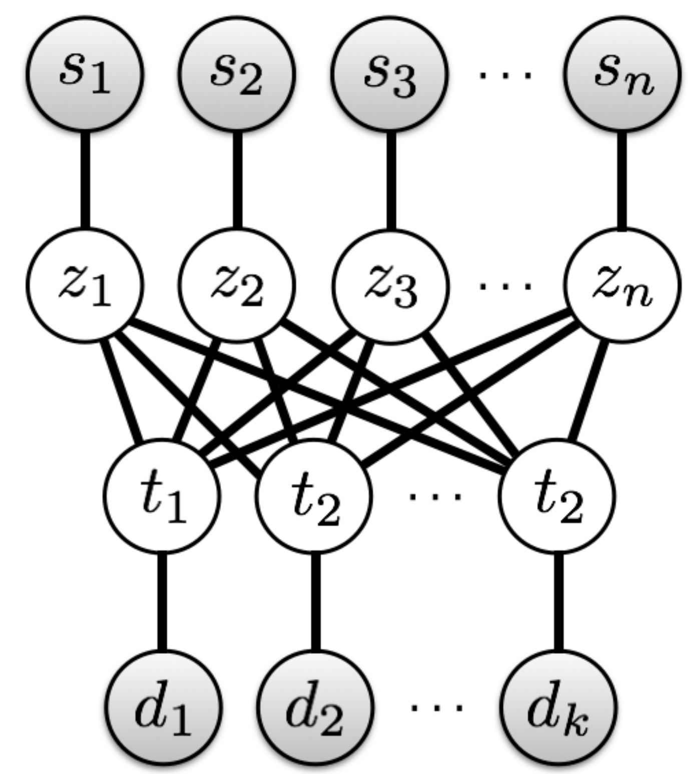
Challenge: missing data
[Ritter, et. al. TACL 2013]

3) Time Normalization

Challenge: diversity in noisy text
[Tabassum, Ritter, Xu, EMNLP 2016]

4) Event Extraction

Challenge: lack of negative examples
[Ritter, et. al. WWW 2015]
[Konovalov, et. al. WWW 2017]



$$O(\theta) = \underbrace{\sum_{i}^{N} \log p_\theta(y_i|x_i)}_{\text{Log Likelihood}} - \underbrace{\lambda^U D(\tilde{p}||\hat{p}_\theta^{\text{unlabeled}})}_{\text{Label regularization}}$$

Learning from **Distant Supervision**

[Mintz et. al. 2009]

1) Named Entity Recognition

Challenge: highly ambiguous labels
[Ritter, et. al. EMNLP 2011]

2) Relation Extraction

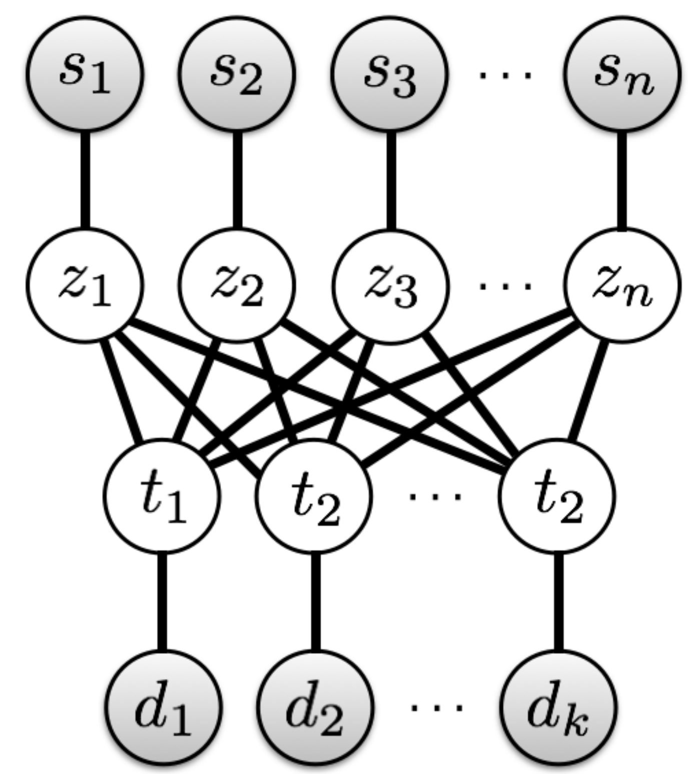
Challenge: missing data
[Ritter, et. al. TACL 2013]

3) Time Normalization

Challenge: diversity in noisy text
[Tabassum, Ritter, Xu, EMNLP 2016]

4) Event Extraction

Challenge: lack of negative examples
[Ritter, et. al. WWW 2015]
[Konovalov, et. al. WWW 2017]



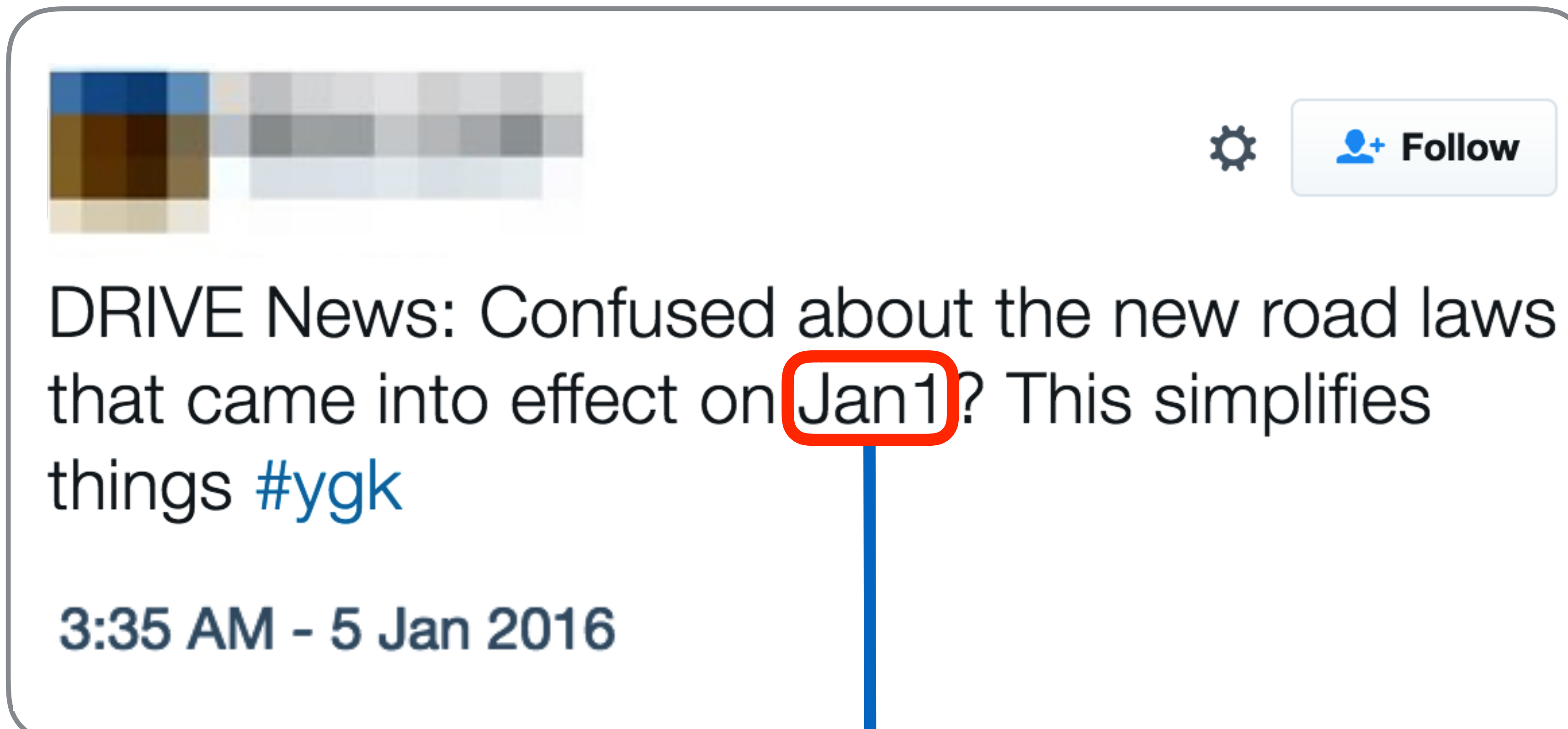
$$O(\theta) = \underbrace{\sum_{i}^{N} \log p_\theta(y_i|x_i)}_{\text{Log Likelihood}} - \underbrace{\lambda^U D(\tilde{p}||\hat{p}_\theta^{\text{unlabeled}})}_{\text{Label regularization}}$$

Time Normalization

[Tabassum, Ritter, Xu EMNLP 2016]

State-of-the-art time resolvers

- TempEX
- HeidelTime
- SUTime
- UWTime



1 Jan 2016

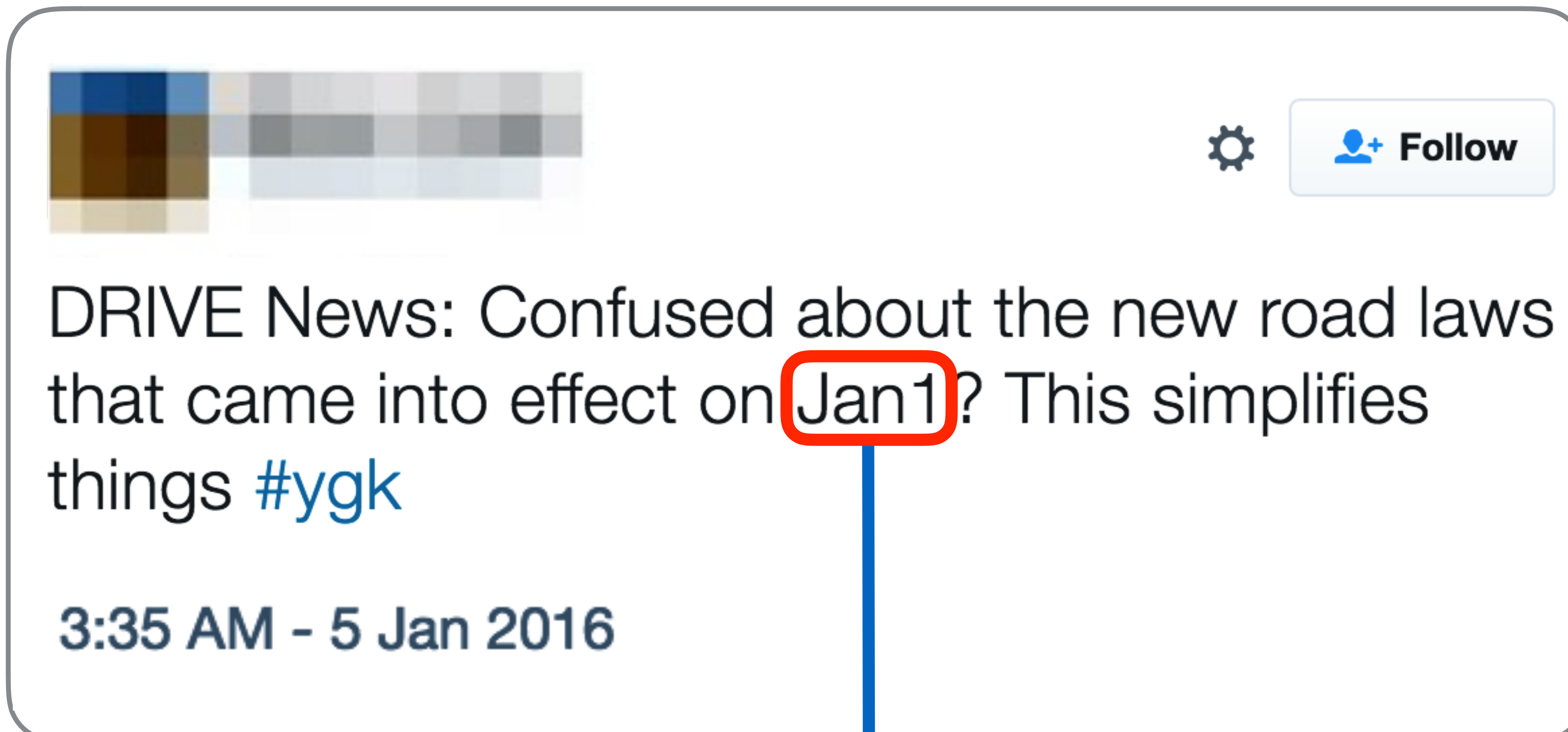
Time Normalization

[Tabassum, Ritter, Xu EMNLP 2016]

Distant Supervision
(no human labels or rules!)

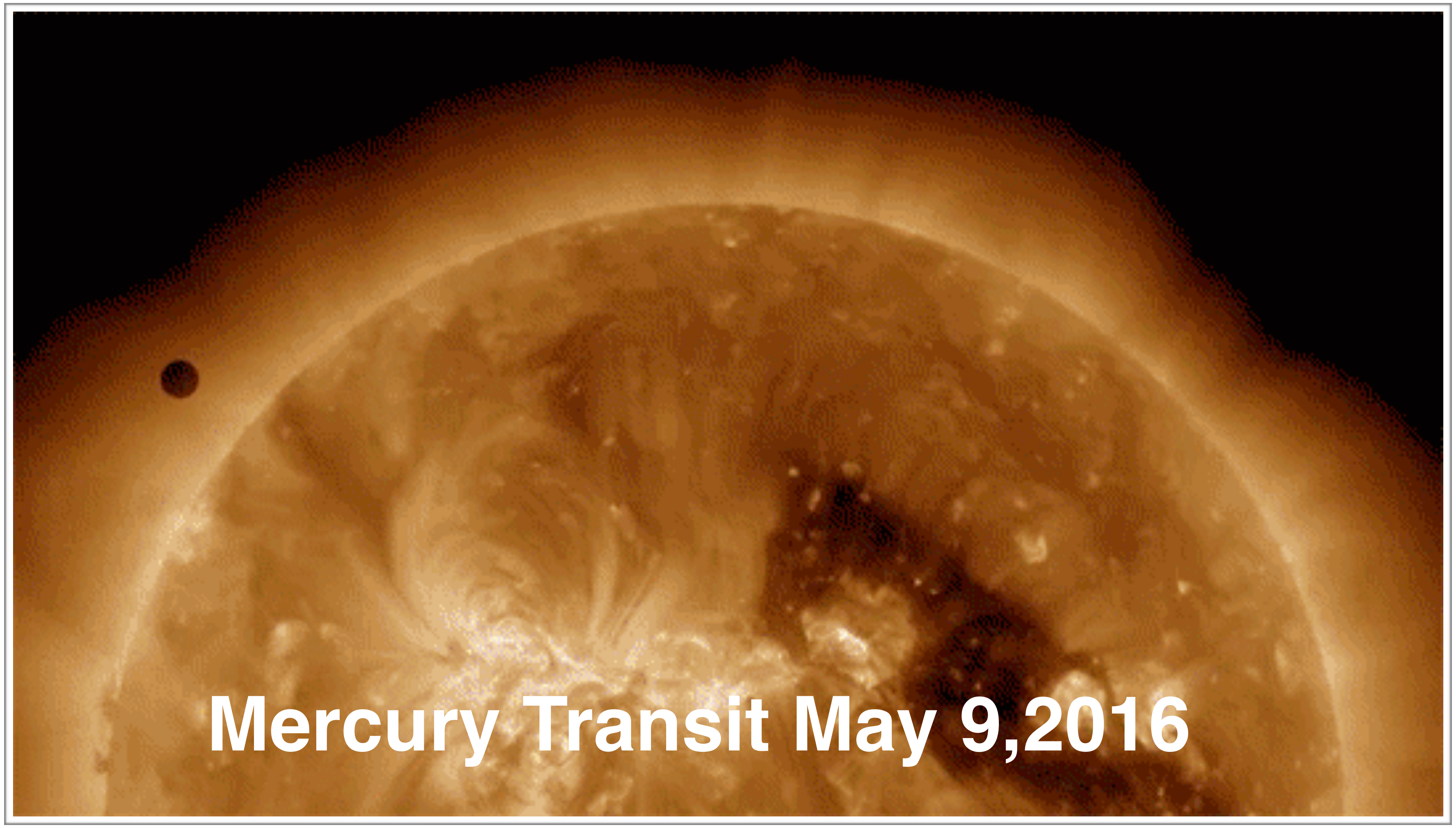
State-of-the-art time resolvers

- TempEX
- HeidelTime
- SUTime
- UWTime

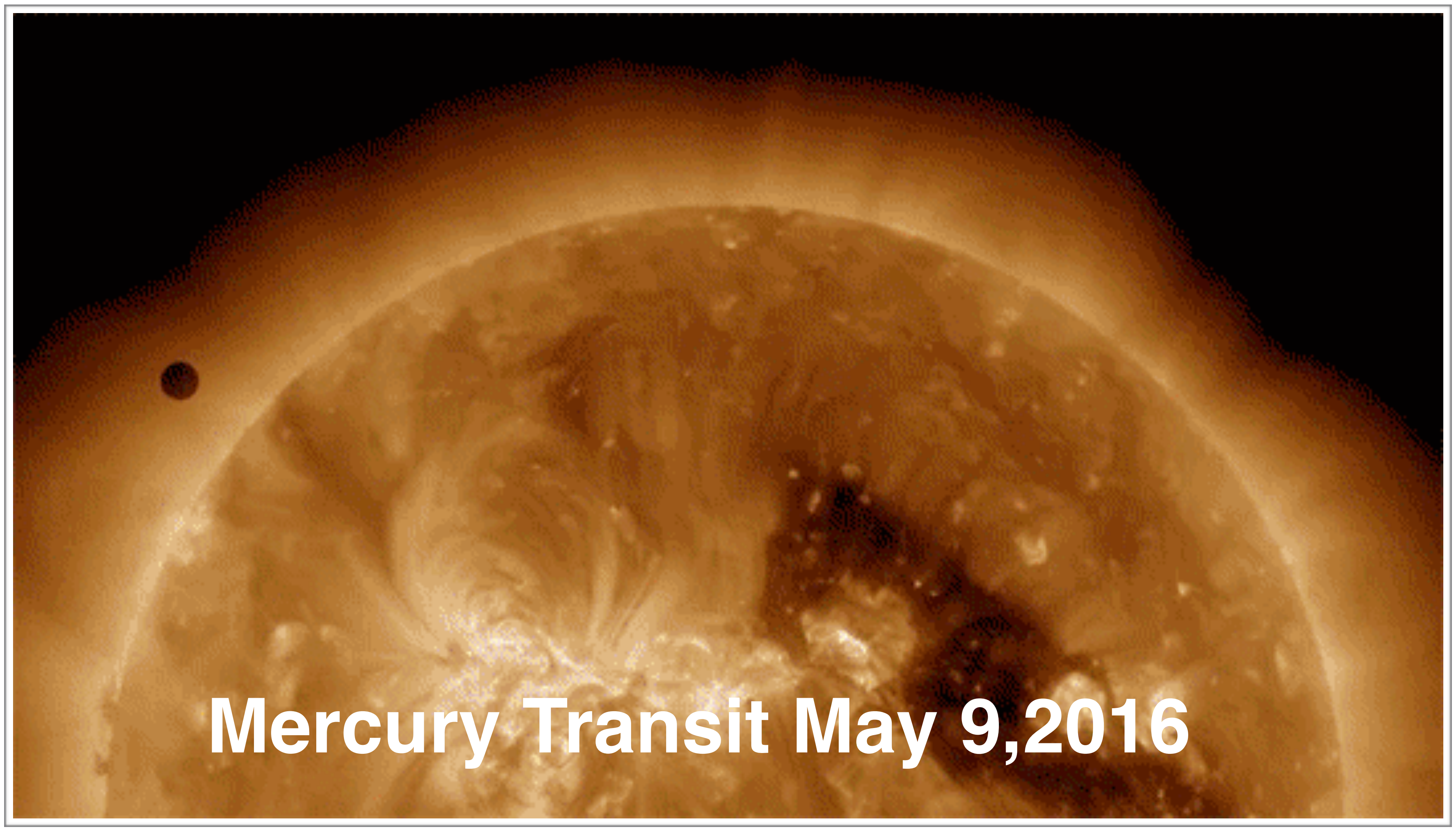


1 Jan 2016

Distant Supervision Assumption

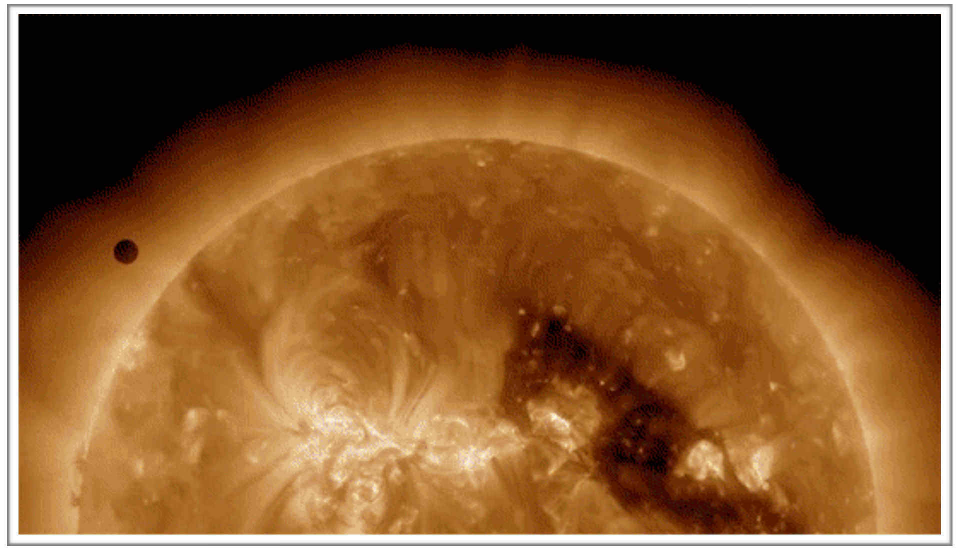


Distant Supervision Assumption



Distant Supervision Assumption

Mercury Transit
May 9,2016



8 May

Follow

Mercury will make a rare transit across the sun tmrw morning (Mon). If you're able to catch it, don't miss out -- and use a solar filter!

10:28 PM - 8 May 2016

Follow

Mercury Transit 2morrow starting at 6:00 AM Mercury will pass in front of Sun @14News @14FirstAlert #mercurytransit

7:30 PM - 8 May 2016

9 May

Follow

Paul from Creators Hand Photography captured a shot of today's Mercury transit, along with a larger sunspot that... fb.me/7jaxf4rfC

Follow

i didn't get to see mercury transit today because of this horrible weather 😤

3:54 PM - 9 May 2016

10 May

Follow

I watched this event yesterday by a small telescope with all the precautions, but this transit of Mercury is great!

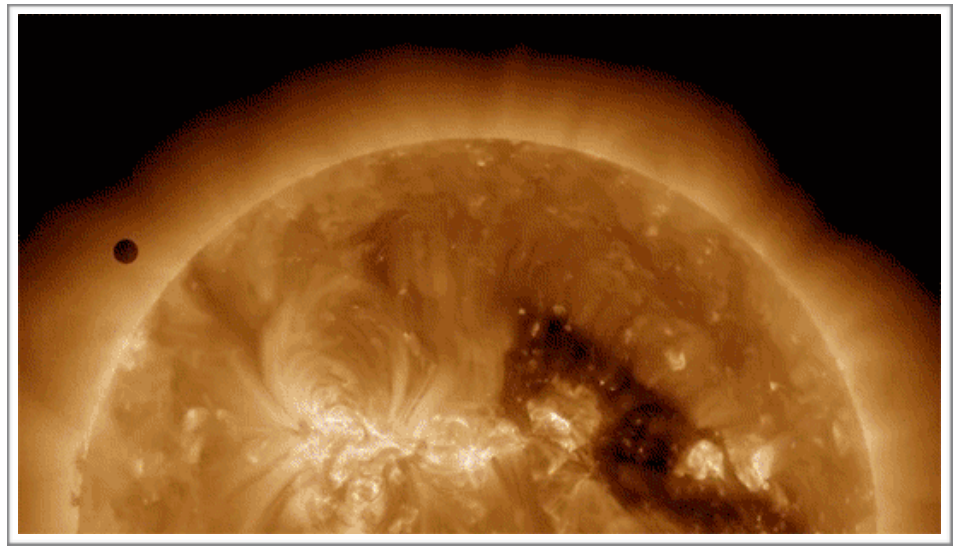
Follow

Mercury passes between Earth and the sun only about 13 times a century.
It was yesterday! May 9th #lagalaxiaensmira

3:17 PM - 9 May 2016

Distant Supervision Assumption

Mercury Transit
May 9,2016



8 May

Mercury will make a rare transit across the sun tmrw morning (Mon). If you're able to catch it, don't miss out -- and use a solar filter!

Mercury Transit 2morrow starting at 6:00 AM Mercury will pass in front of Sun @14News @14FirstAlert #mercurytransit

7:30 PM - 8 May 2016

9 May

Paul from Creators Hand Photography captured a shot of today's Mercury transit, along with a larger sunspot that... fb.me/7jaxf4rfC

i didn't get to see mercury transit today because of this horrible weather 😤

3:54 PM - 9 May 2016

10 May

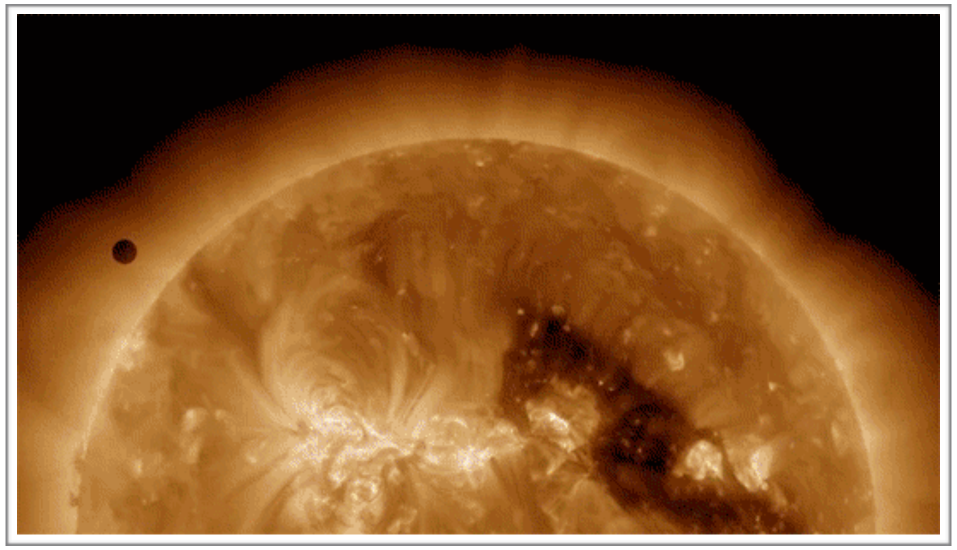
I watched this event yesterday by a small telescope with all the precautions, but this transit of Mercury is great!

Mercury passes between Earth and the sun only about 13 times a century.
It was yesterday! May 9th #lagalaxiaensmira

3:17 PM - 9 May 2016

Distant Supervision Assumption

Mercury Transit
May 9,2016



8 May

Mercury will make a rare transit across the sun tmrw morning Mon! If you're able to catch it, don't miss out -- and use a solar filter!

Mercury Transit 2morrow starting at 6:00 AM Mercury will pass in front of Sun @14News @14FirstAlert #mercurytransit

7:30 PM - 8 May 2016

9 May

Paul from Creators Hand Photography captured a shot of today's Mercury transit, along with a larger sunspot that... fb.me/7jaxf4rfC

i didn't get to see mercury transit today because of this horrible weather

3:54 PM - 9 May 2016

10 May

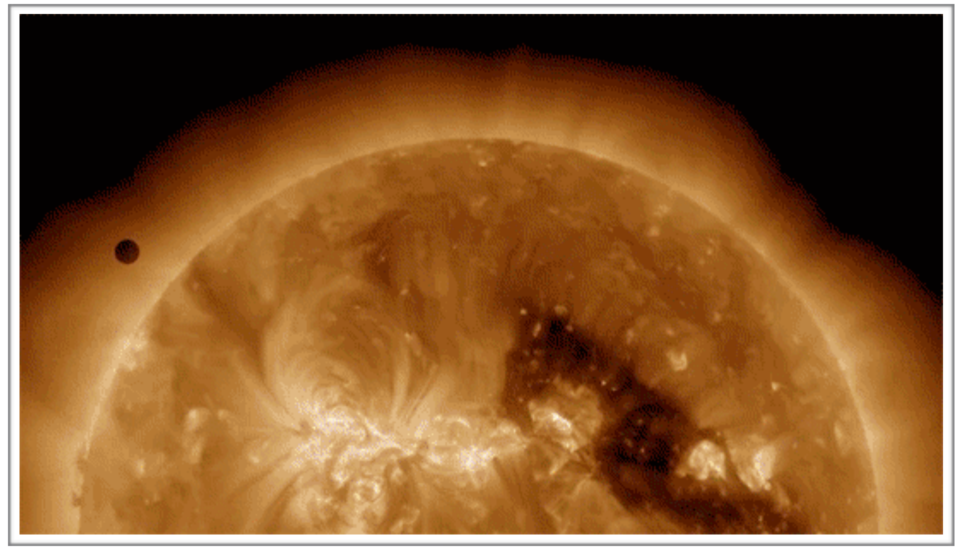
I watched this event yesterday by a small telescope with all the precautions, but this transit of Mercury is great!

Mercury passes between Earth and the sun only about 13 times a century.
It was yesterday May 9th #lagalaxiaensmira

3:17 PM - 9 May 2016

Distant Supervision Assumption

Mercury Transit
May 9,2016



8 May | 9 May | 10 May

8 May

Follow

Mercury will make a rare transit across the sun tmrw morning Mon). If you're able to catch it, don't miss out -- and use a solar filter!

10:28 PM - 8 May 2016

Follow

Mercury Transit 2morrow starting at 6:00 AM Mercury will pass in front of Sun @14News @14FirstAlert #mercurytransit

7:30 PM - 8 May 2016

9 May

Follow

Paul from Creators Hand Photography captured a shot of today's Mercury transit, along with a larger sunspot that... fb.me/7jaxf4rfC

Follow

i didn't get to see mercury transit today because of this horrible weather 😤

3:54 PM - 9 May 2016

10 May

Follow

I watched this event yesterday by a small telescope with all the precautions, but this transit of Mercury is great!

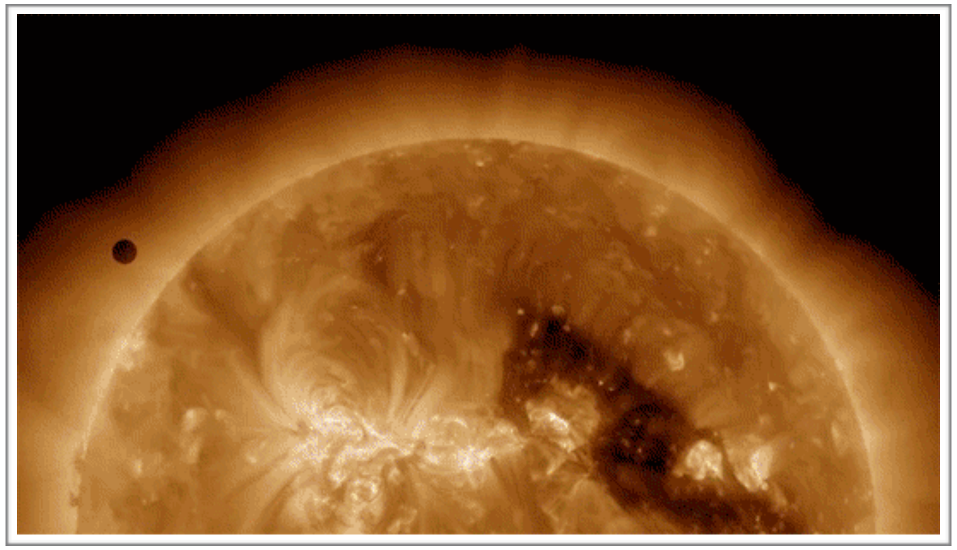
Follow

Mercury passes between Earth and the sun only about 13 times a century.
It was yesterday May 9th #lagalaxiaensmira

3:17 PM - 9 May 2016

Distant Supervision Assumption

Mercury Transit
May 9,2016



8 May

Mercury will make a rare transit across the sun tmrw morning (Mon). If you're able to catch it, don't miss out -- and use a solar filter!

Mercury Transit 2morrow starting at 6:00 AM Mercury will pass in front of Sun @14News @14FirstAlert #mercurytransit

7:30 PM - 8 May 2016

9 May

Paul from Creators Hand Photography captured a shot of today's Mercury transit, along with a larger sunspot that... fb.me/7jaxf4rfC

i didn't get to see mercury transit today because of this horrible weather

3:54 PM - 9 May 2016

10 May

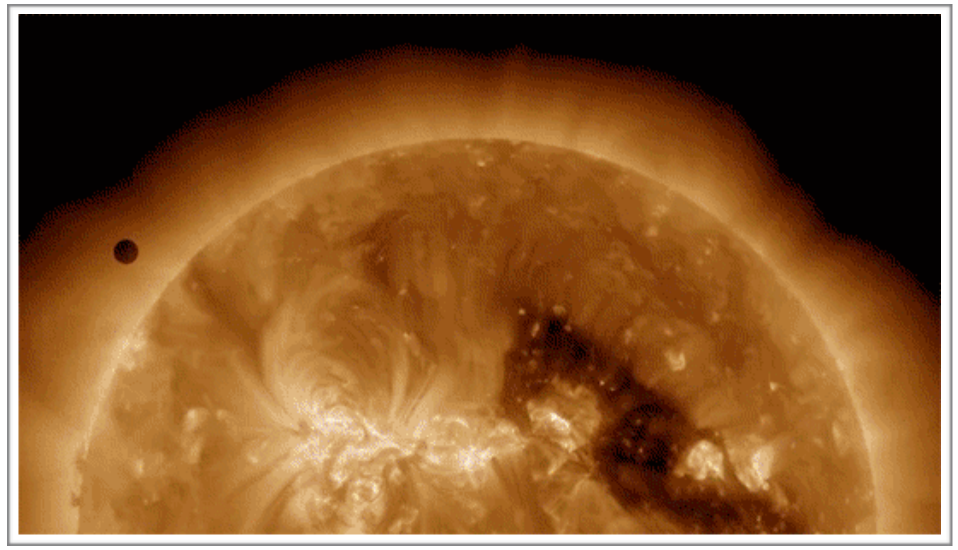
I watched this event yesterday by a small telescope with all the precautions, but this transit of Mercury is great!

Mercury passes between Earth and the sun only about 13 times a century.
It was yesterday May 9th #lagalaxiaensmira

3:17 PM - 9 May 2016

Distant Supervision Assumption

Mercury Transit
May 9,2016



8 May

Follow

Mercury will make a rare transit across the sun tmrw morning (Mon). If you're able to catch it, don't miss out -- and use a solar filter!

10:28 PM - 8 May 2016

Follow

Mercury Transit 2morrow starting at 6:00 AM Mercury will pass in front of Sun @14News @14FirstAlert #mercurytransit

7:30 PM - 8 May 2016

⋮

9 May

Follow

Paul from Creators Hand Photography captured a shot of today's Mercury transit, along with a larger sunspot that... fb.me/7jaxf4rfC

Follow

i didn't get to see mercury transit today because of this horrible weather 😤

3:54 PM - 9 May 2016

⋮

10 May

Follow

I watched this event yesterday by a small telescope with all the precautions, but this transit of Mercury is great!

Follow

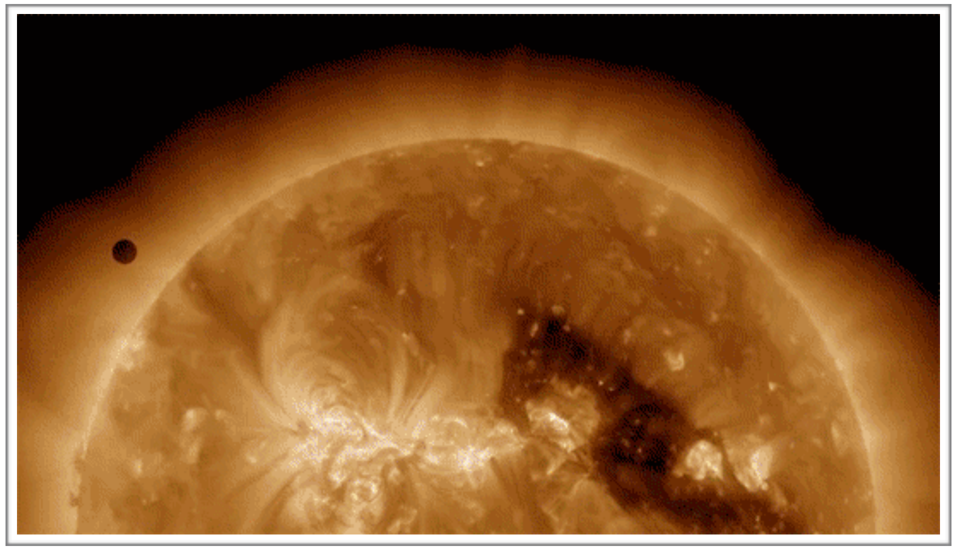
Mercury passes between Earth and the sun only about 13 times a century.
It was yesterday May 9th #lagalaxiaensmira

3:17 PM - 9 May 2016

⋮

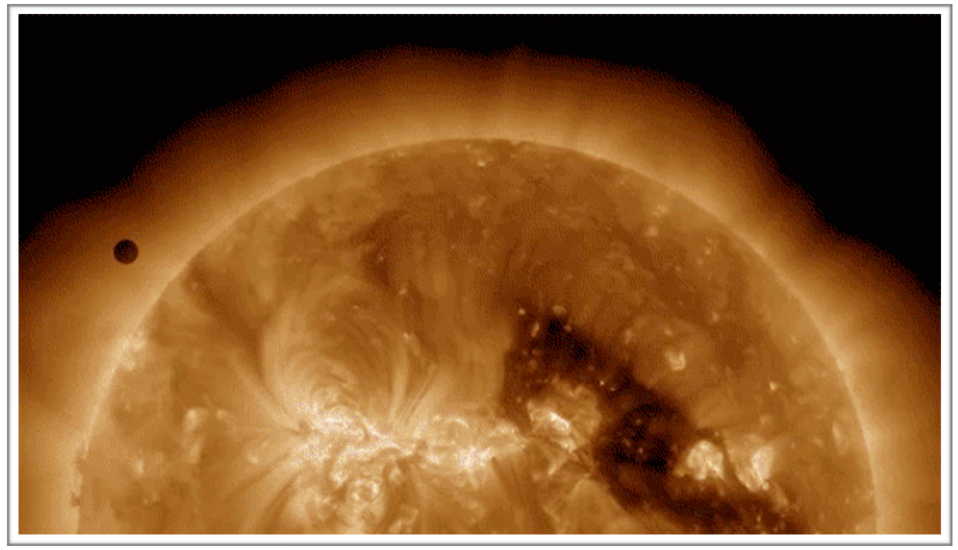
Distant Supervision Assumption

Mercury Transit
May 9,2016

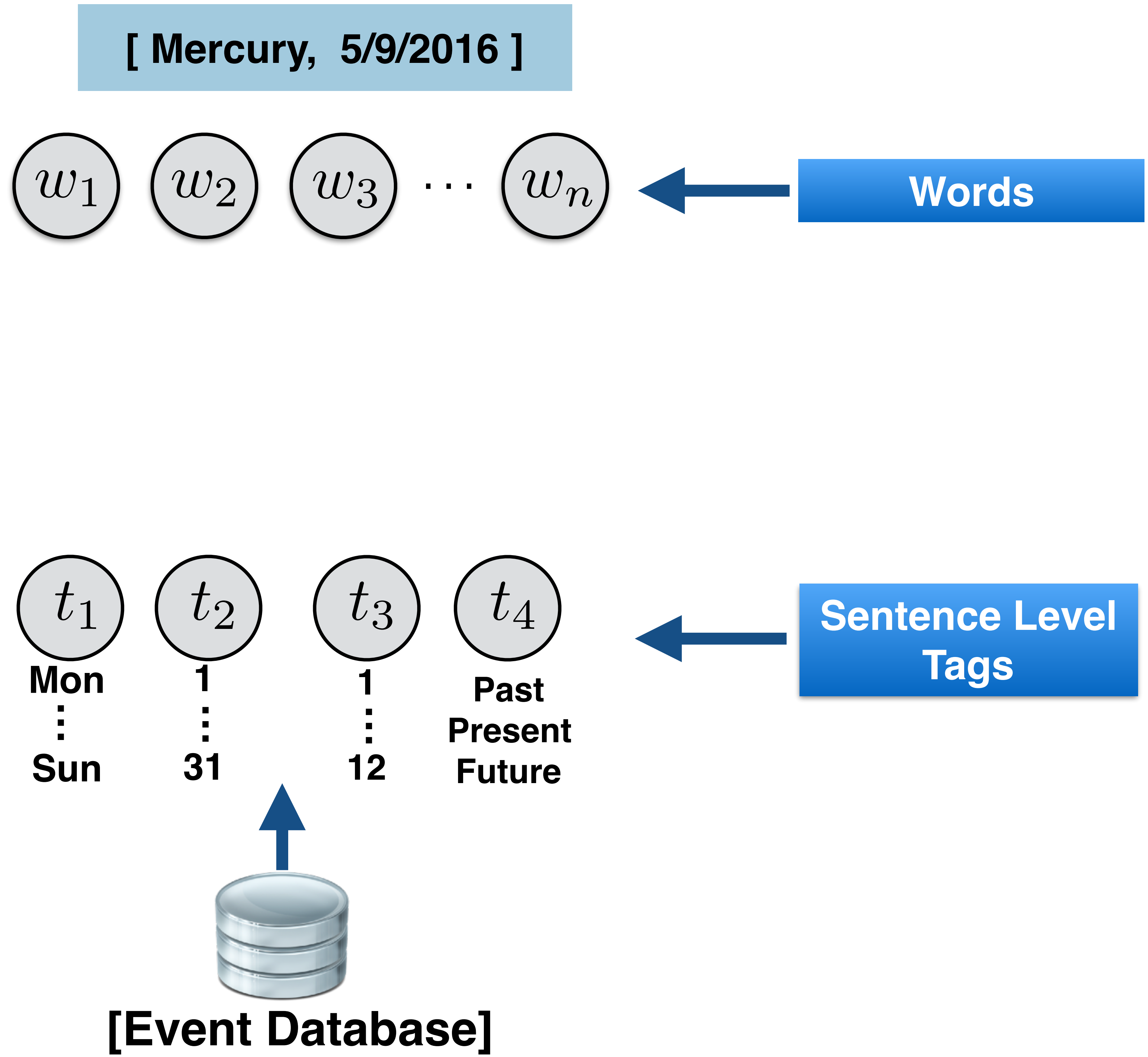


Distant Supervision Assumption

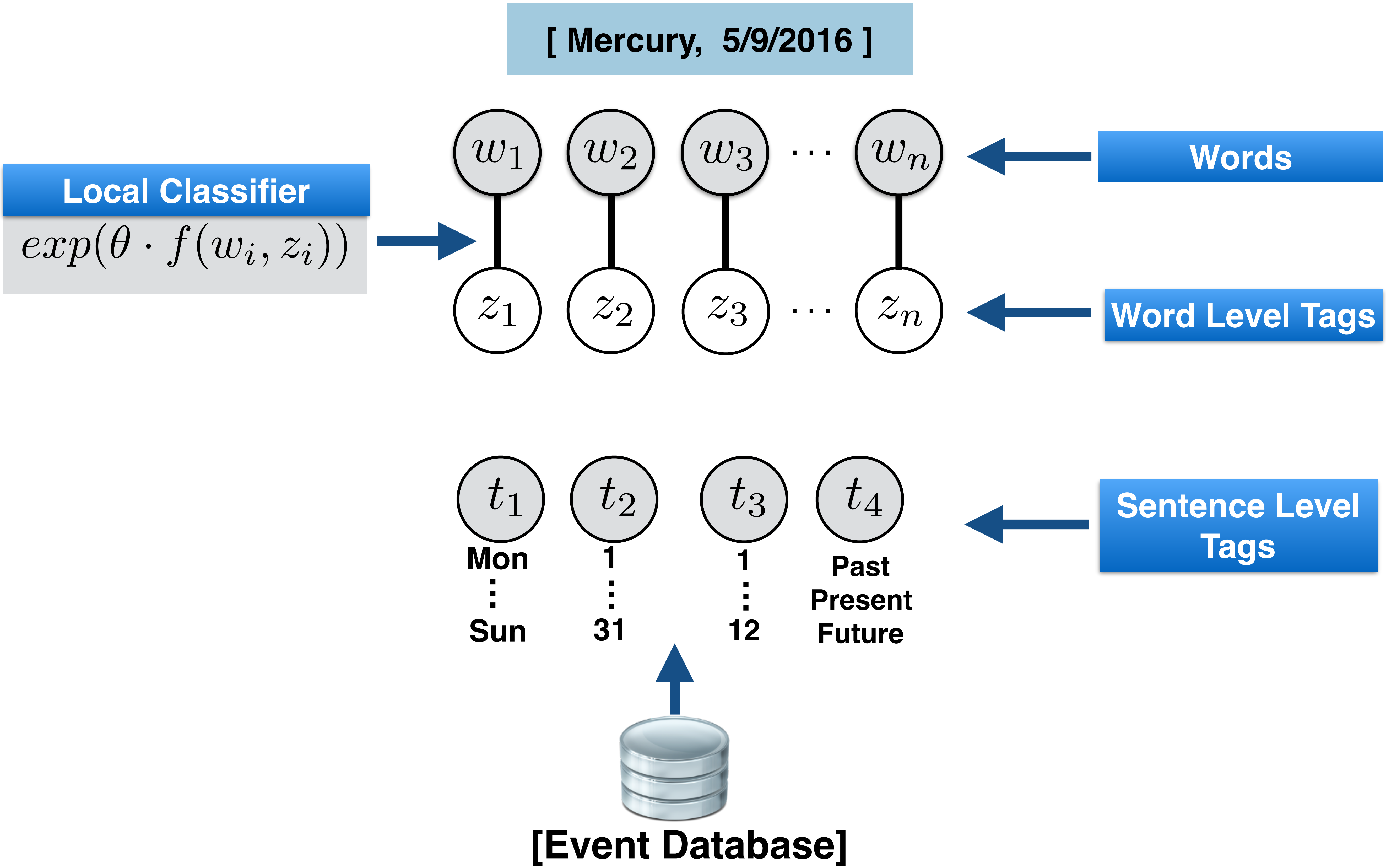
Mercury Transit
May 9,2016



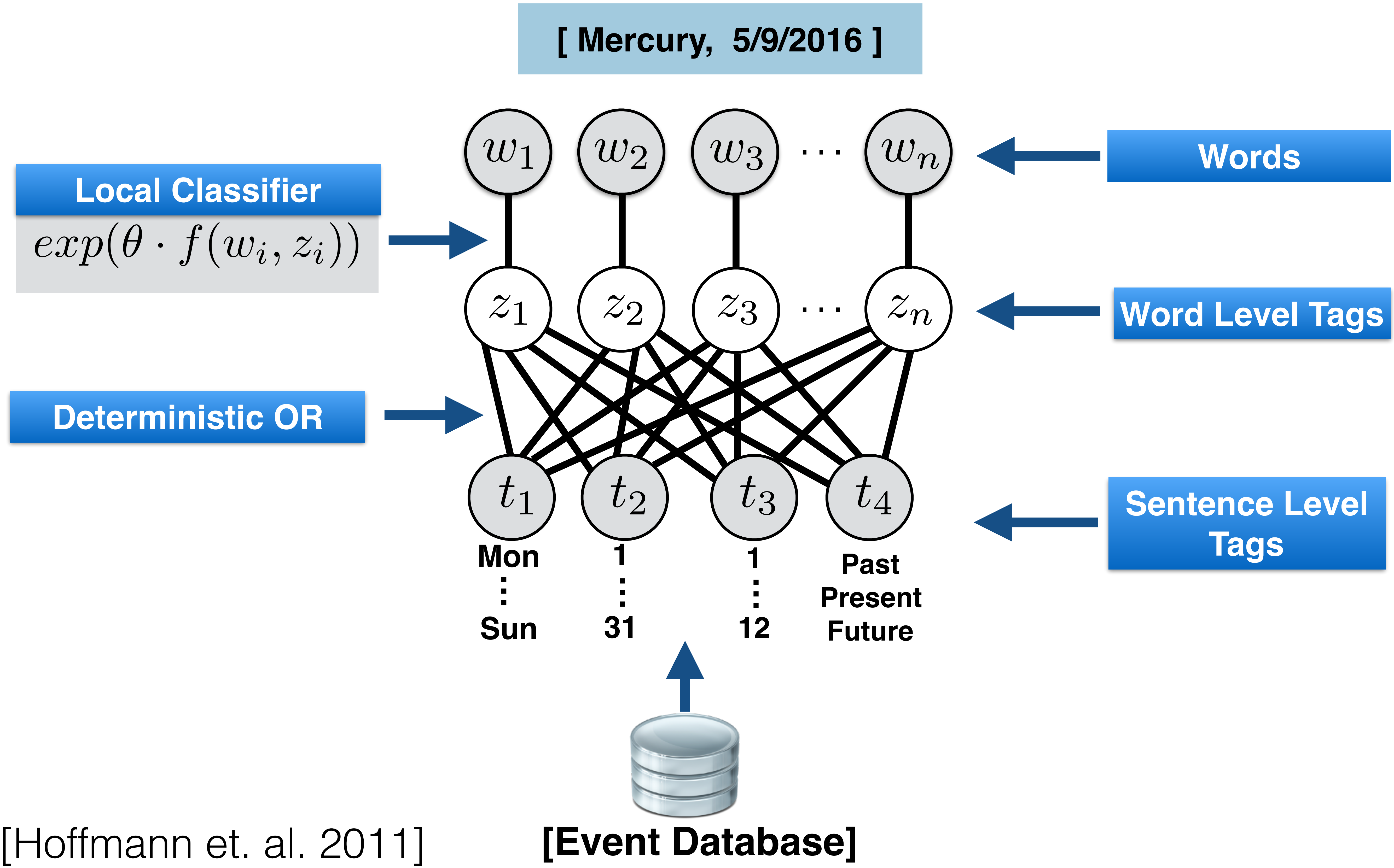
Multiple Instance Learning Tagger



Multiple Instance Learning Tagger

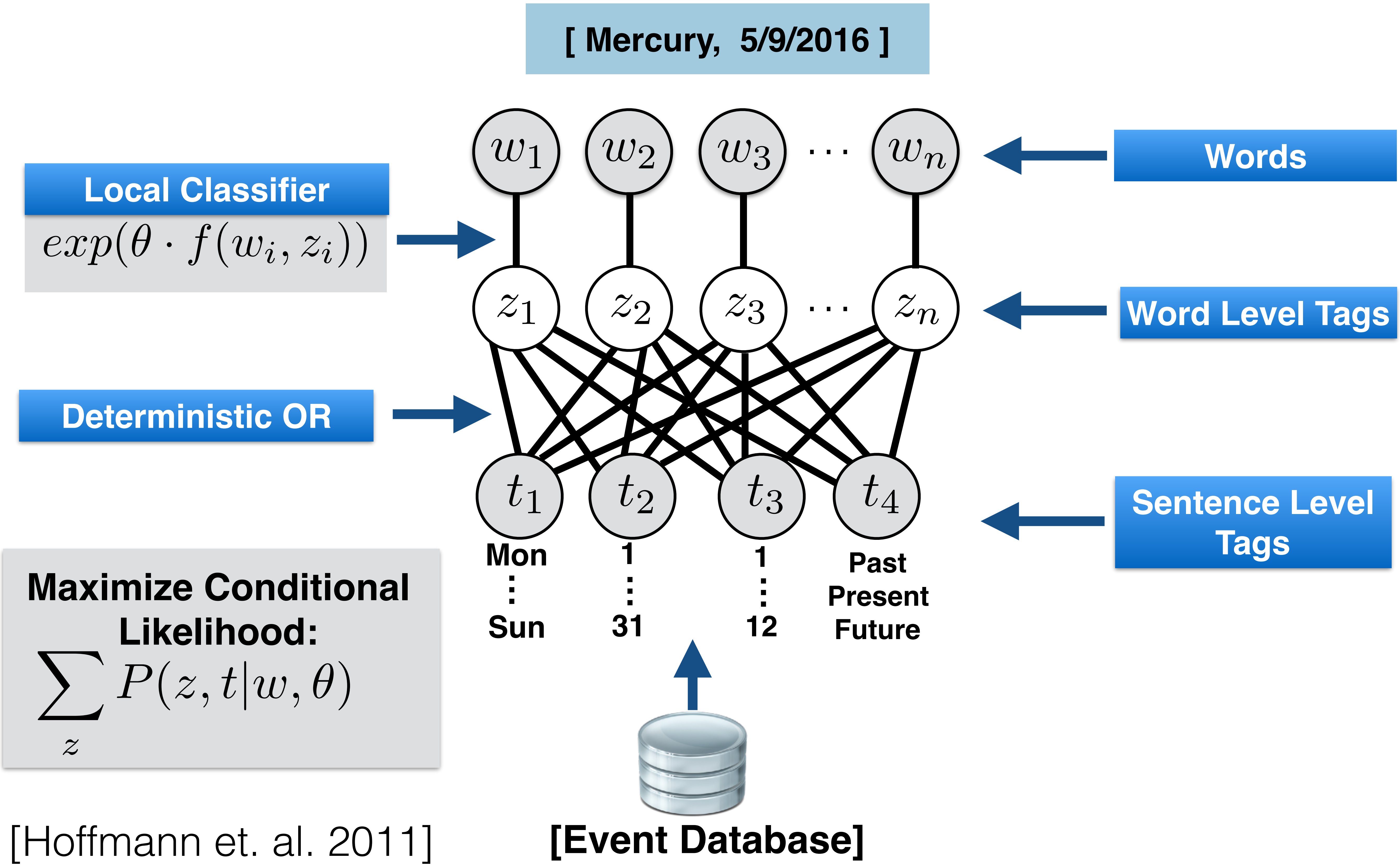


Multiple Instance Learning Tagger



[Hoffmann et. al. 2011]

Multiple Instance Learning Tagger



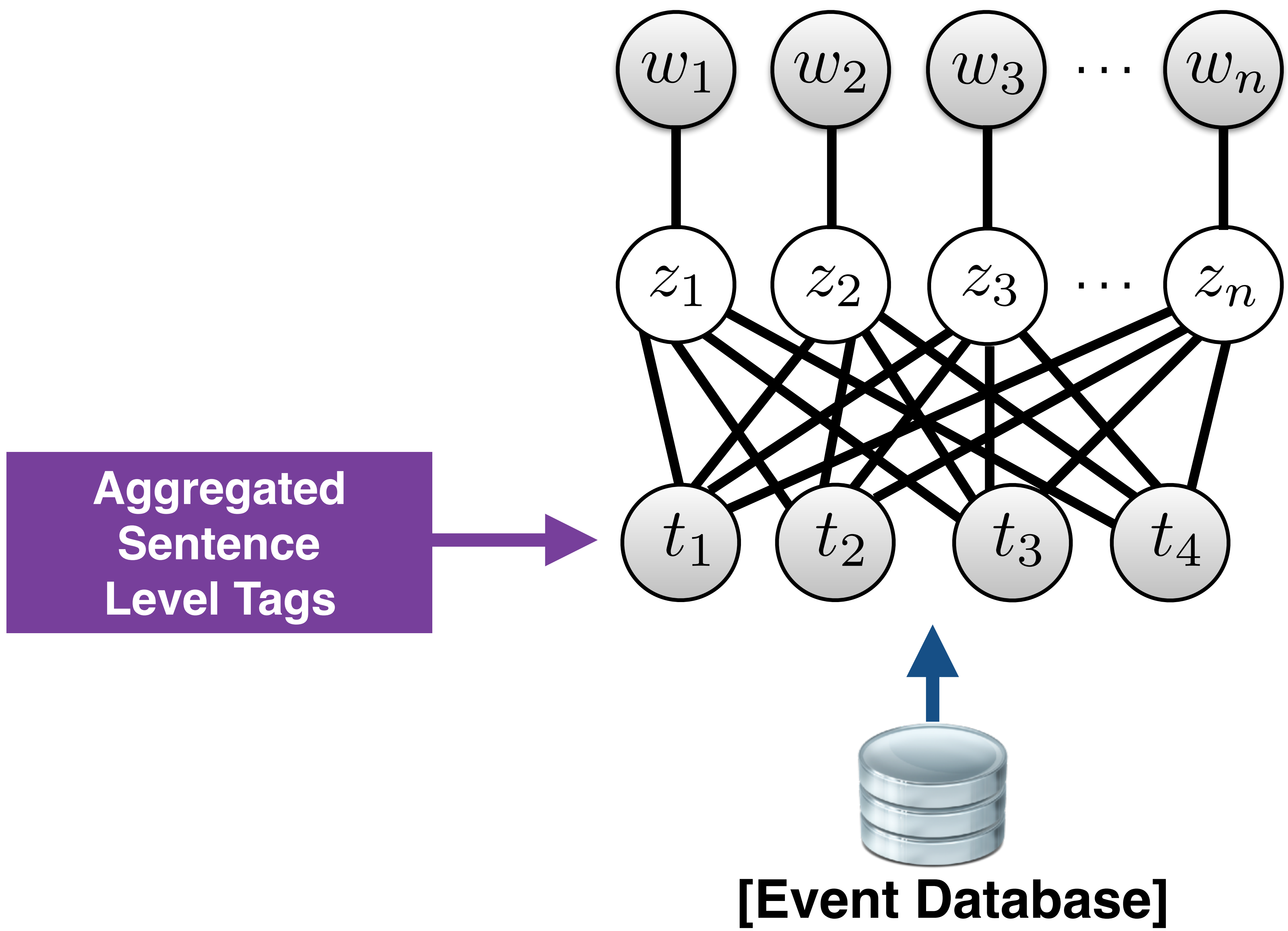
[Hoffmann et. al. 2011]

Missing Data Problem

Sentence Level Tags:

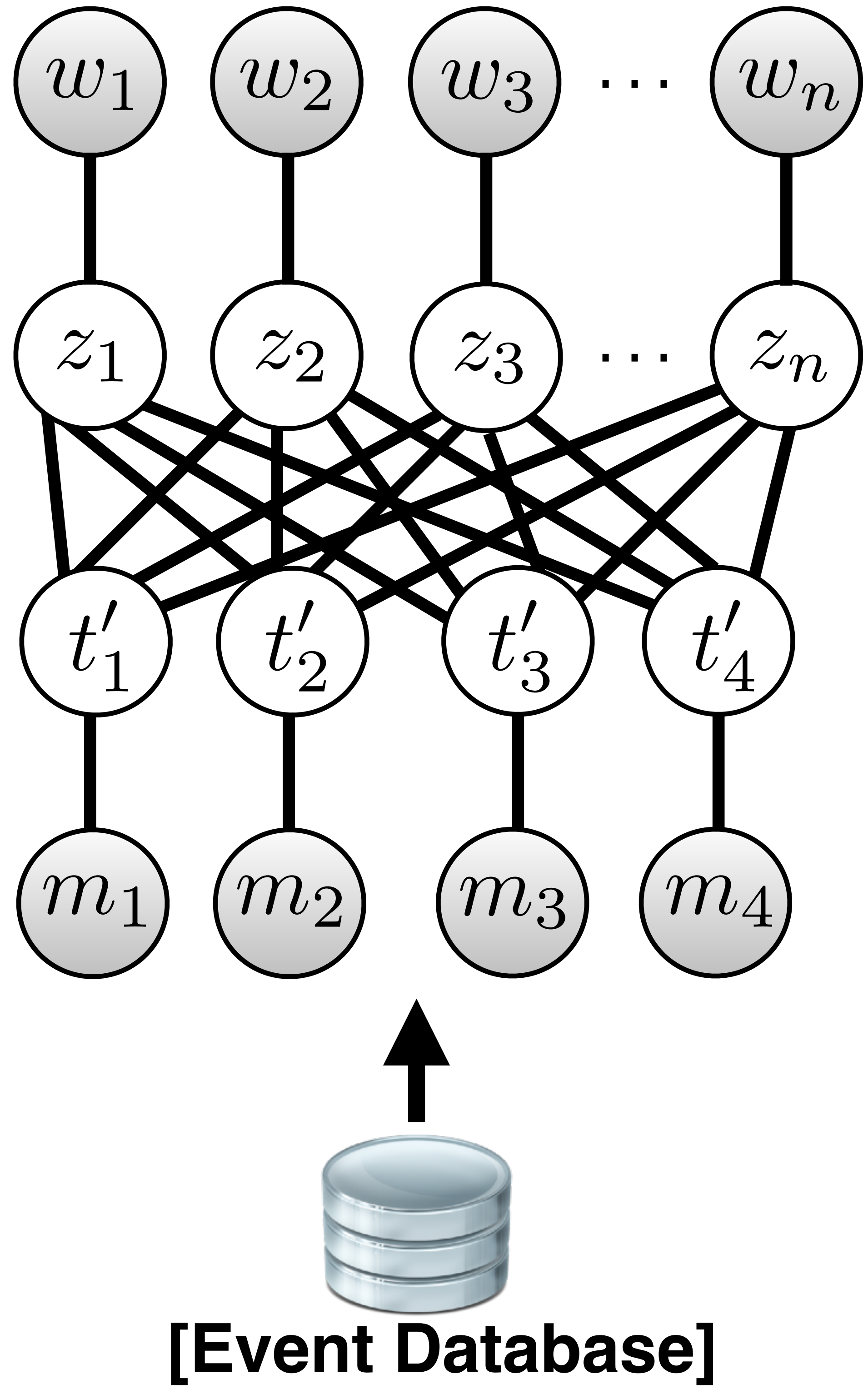
- ✔ TL = Future
- ✘ MOY= May
- ✘ DOM=9
- ✔ DOW= Mon

Missing Data Extension



Missing Data Extension

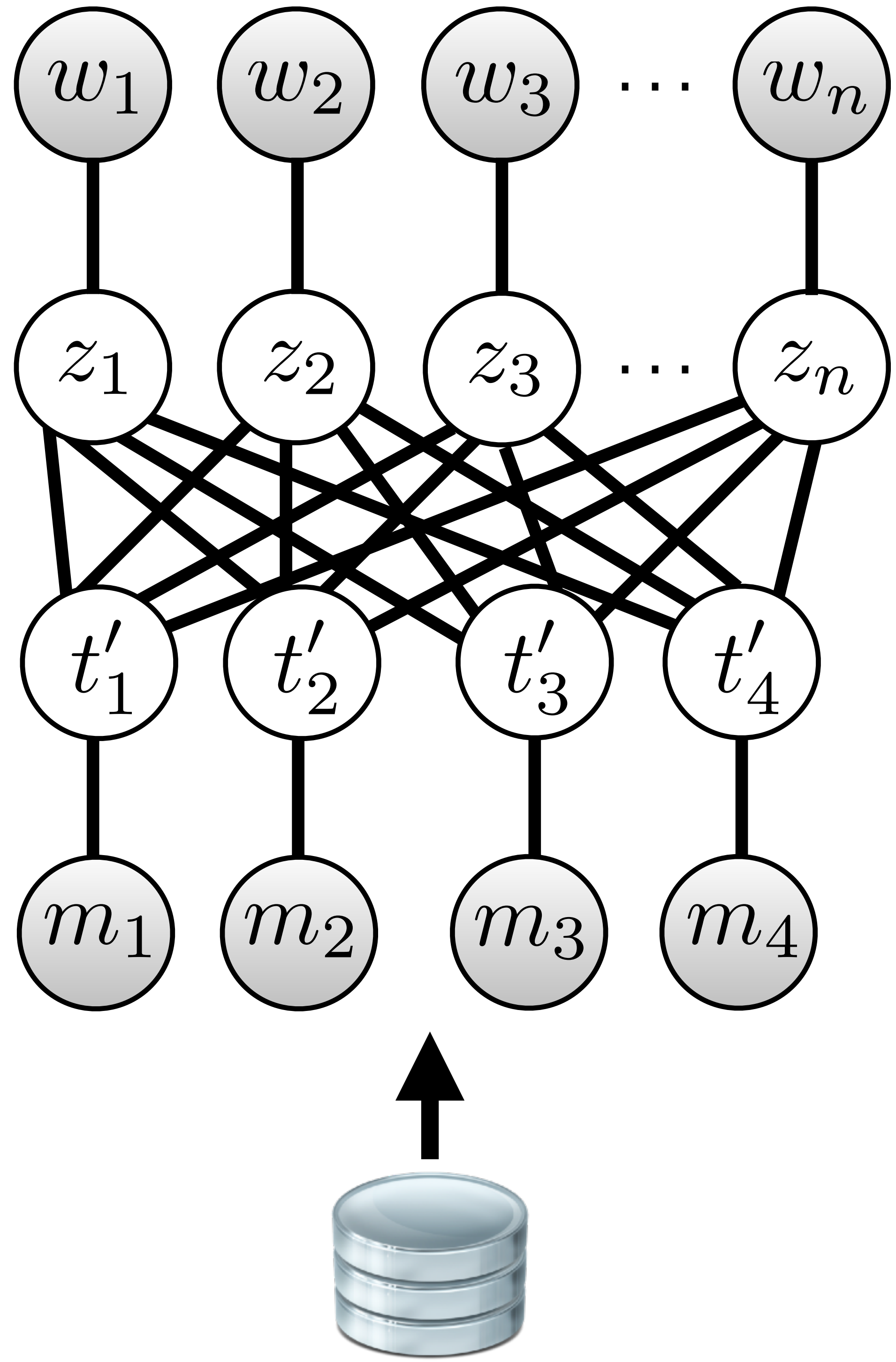
Missing Data Problem
In Distant Supervision
[Ritter, et. al. TACL 2013]



Missing Data Extension

Missing Data Problem
In Distant Supervision
[Ritter, et. al. TACL 2013]

Mentioned in Text



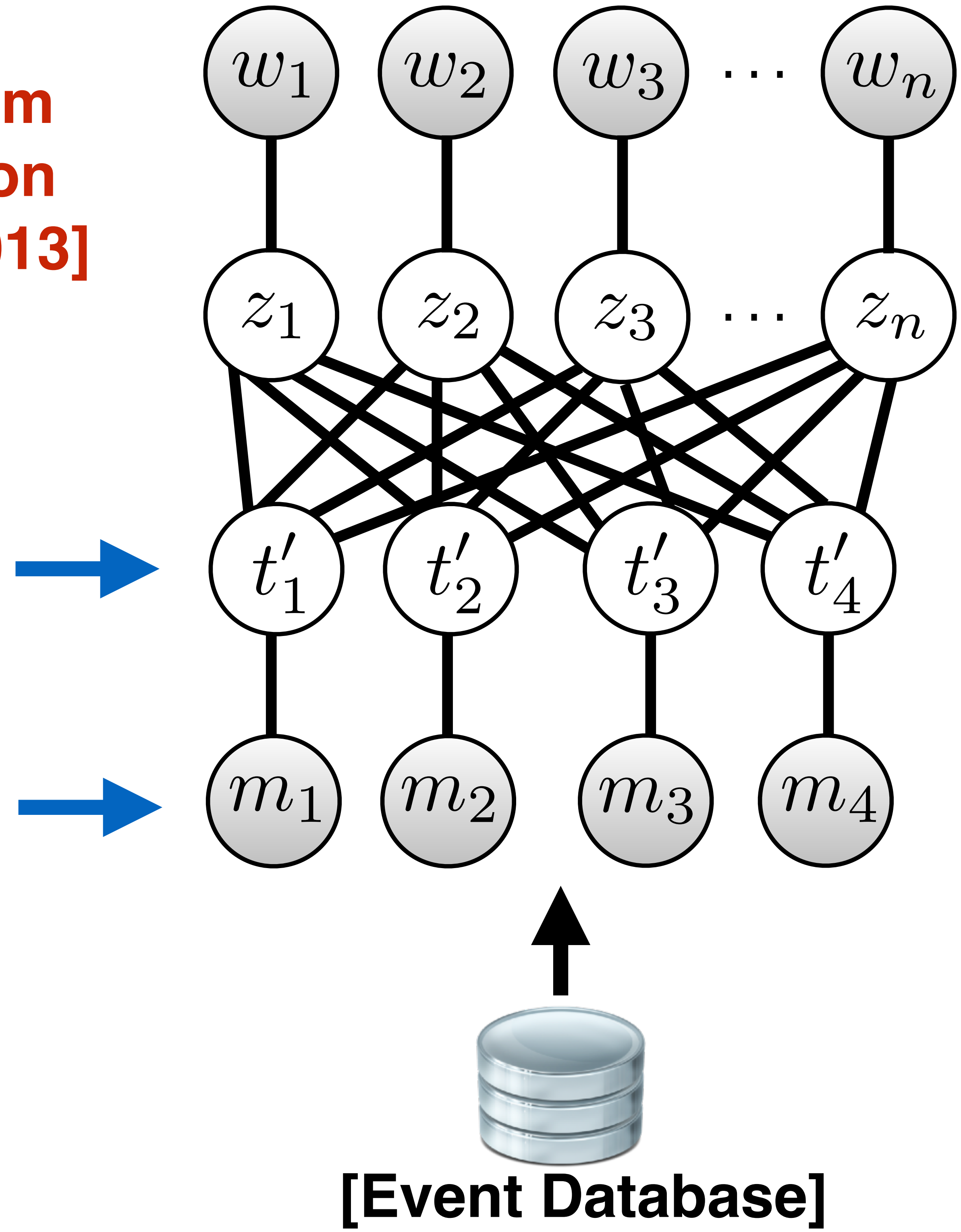
[Event Database]

Missing Data Extension

Missing Data Problem
In Distant Supervision
[Ritter, et. al. TACL 2013]

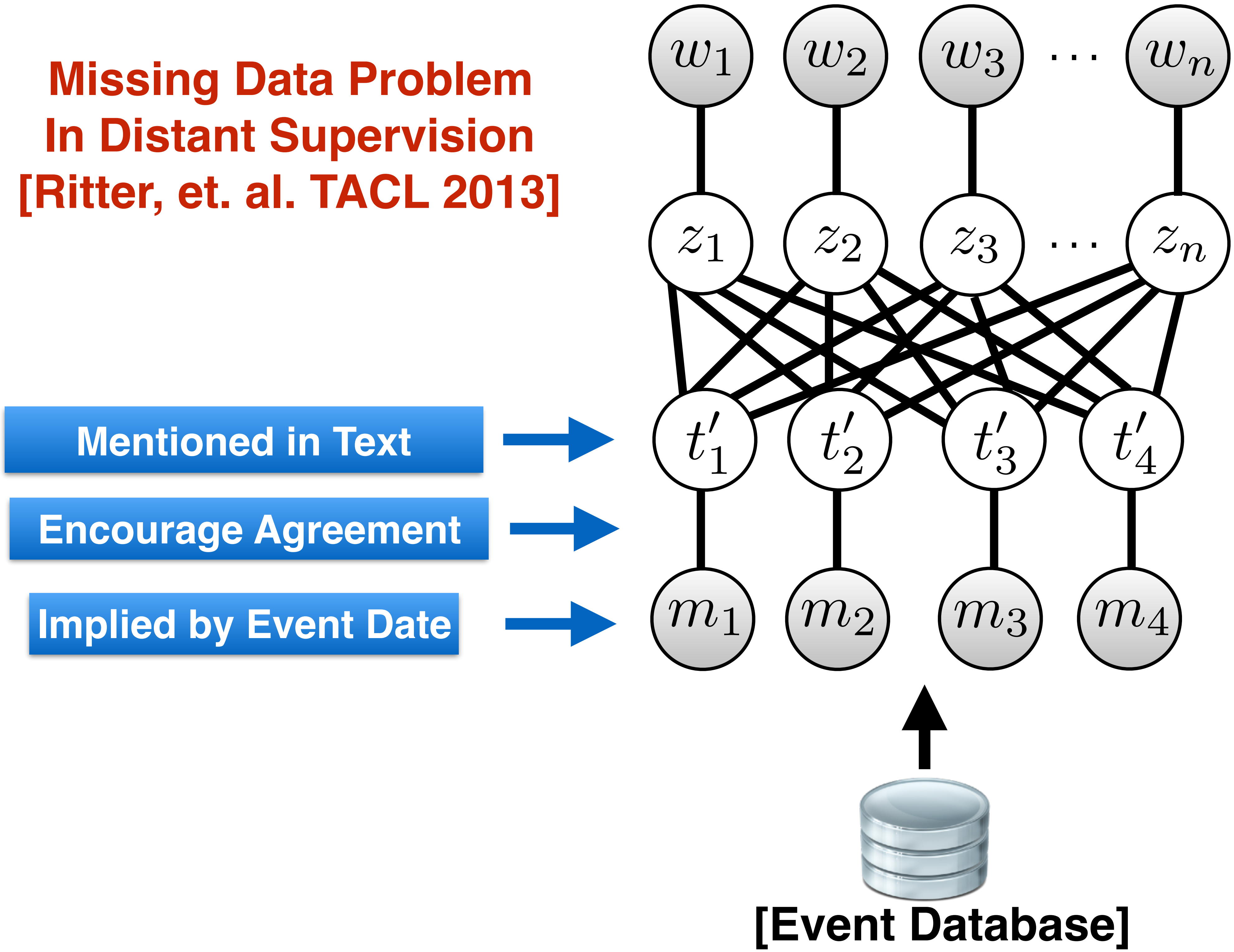
Mentioned in Text

Implied by Event Date



Missing Data Extension

Missing Data Problem
In Distant Supervision
[Ritter, et. al. TACL 2013]

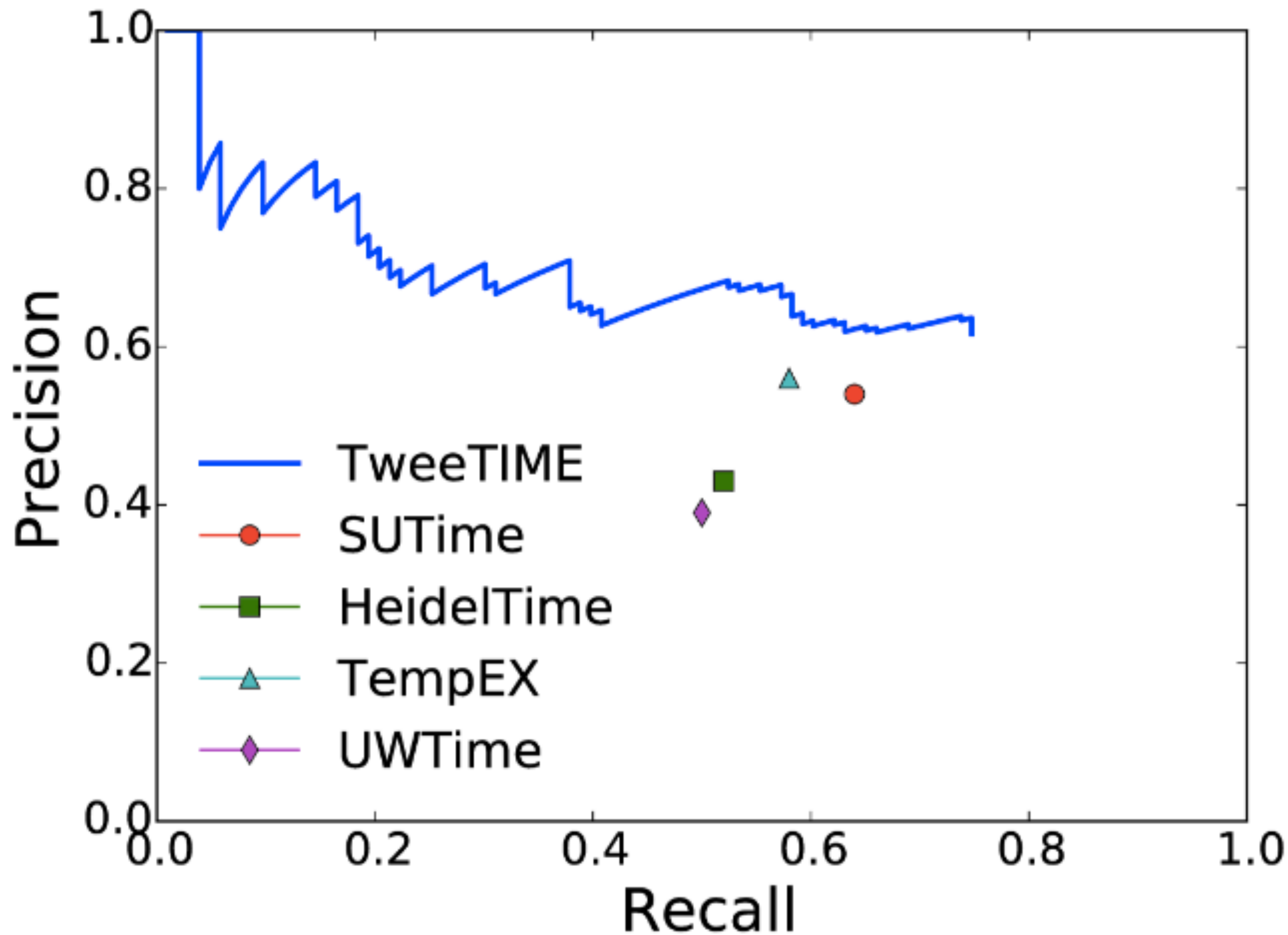


Example Tags

Word	Im	Hella	excited	for	tomorrow
Tag	NA	NA	Future	NA	Future

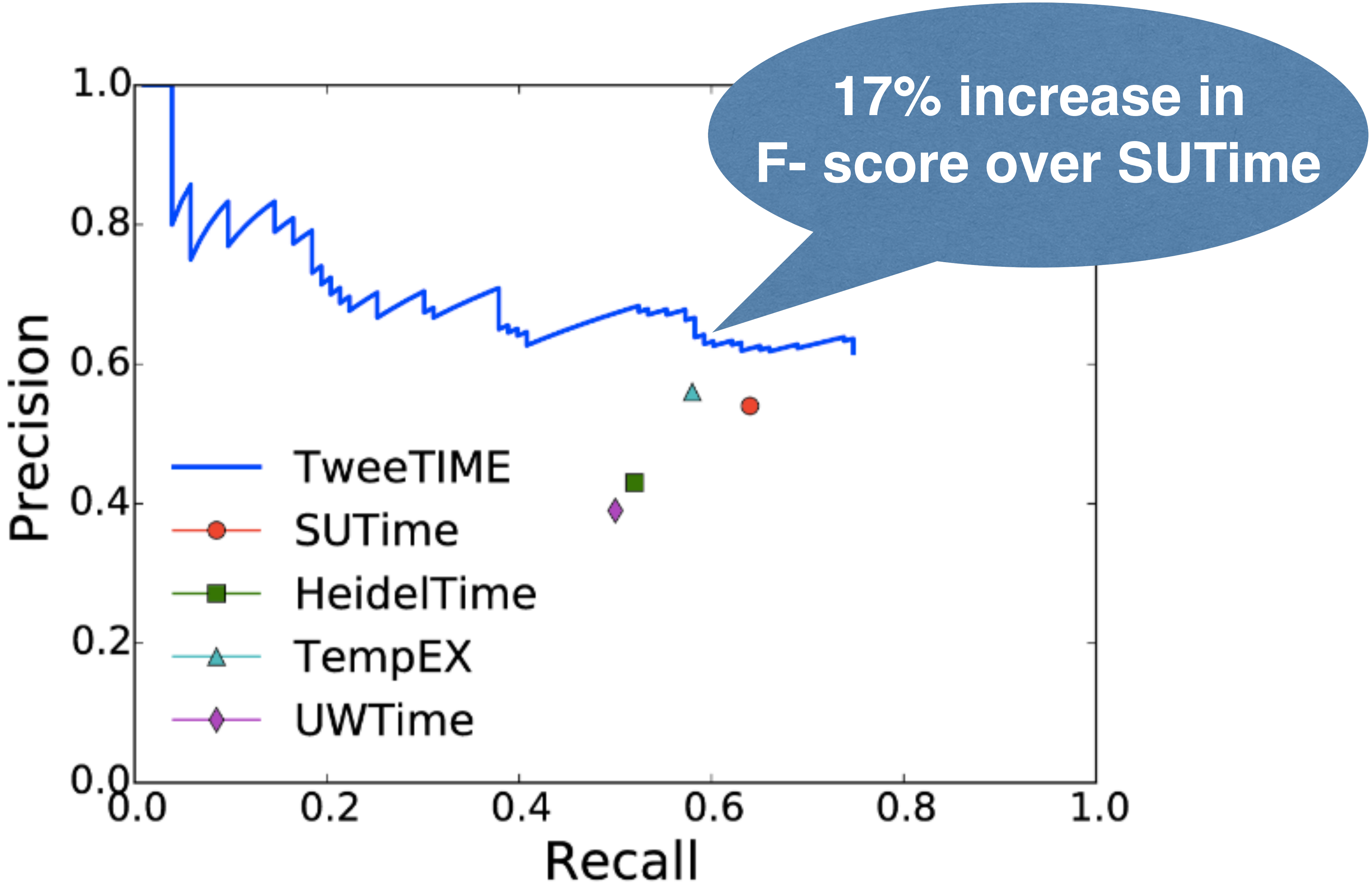
Word	Thnks	for	a	Christmas	party	on	fri
Tag	NA	NA	NA	December	NA	NA	Friday

Evaluation

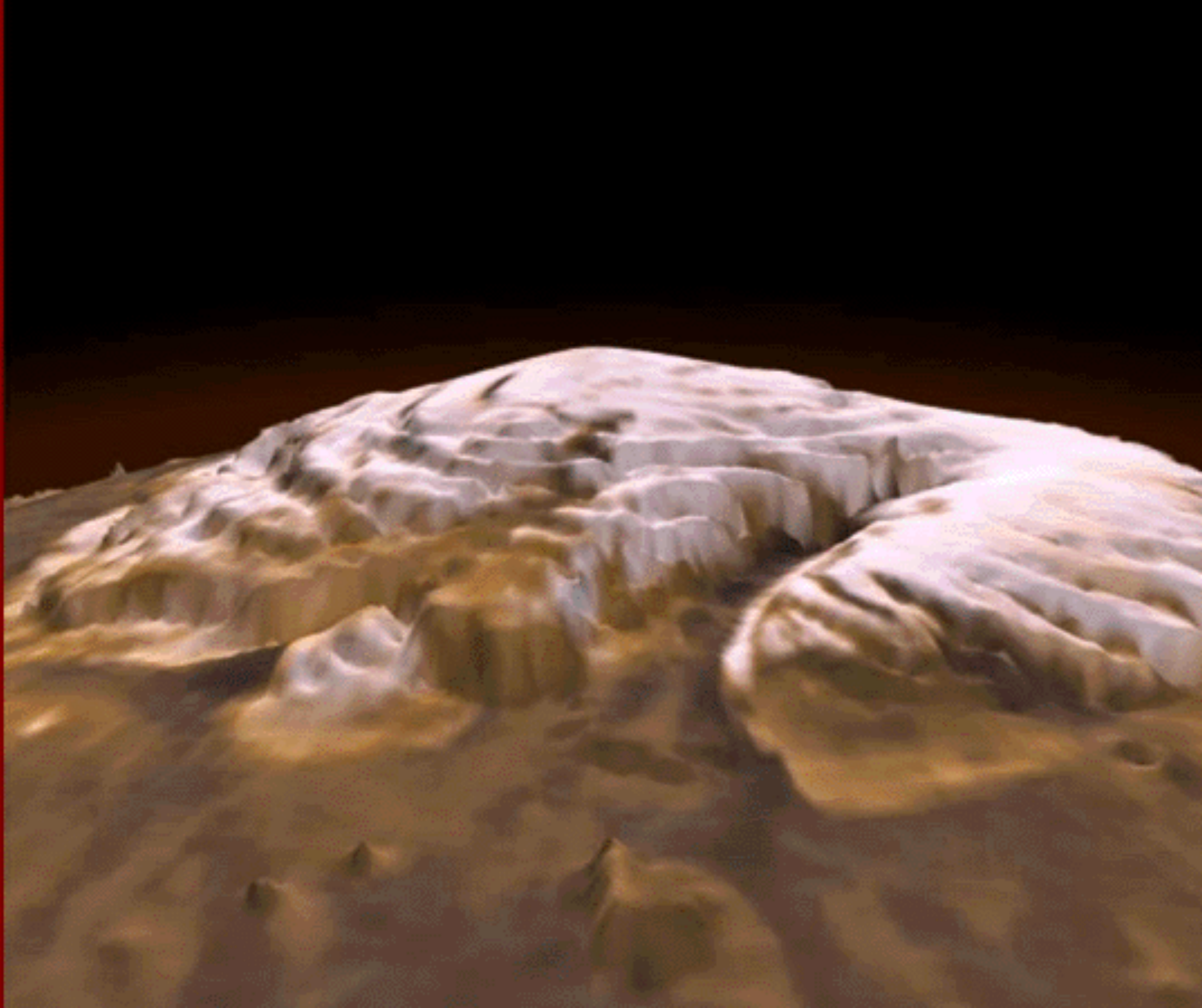


Precision
Recall
TweeTIME
SUTime
HeidelTime
TempEX
UWTime
0.0
0.2
0.4
0.6
0.8
1.0

Evaluation


17% increase in
F- score over SUTime
Precision
Recall
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
TweeTIME
SUTime
HeidelTime
TempEX
UWTime

Following the Water: The Mars Exploration Program



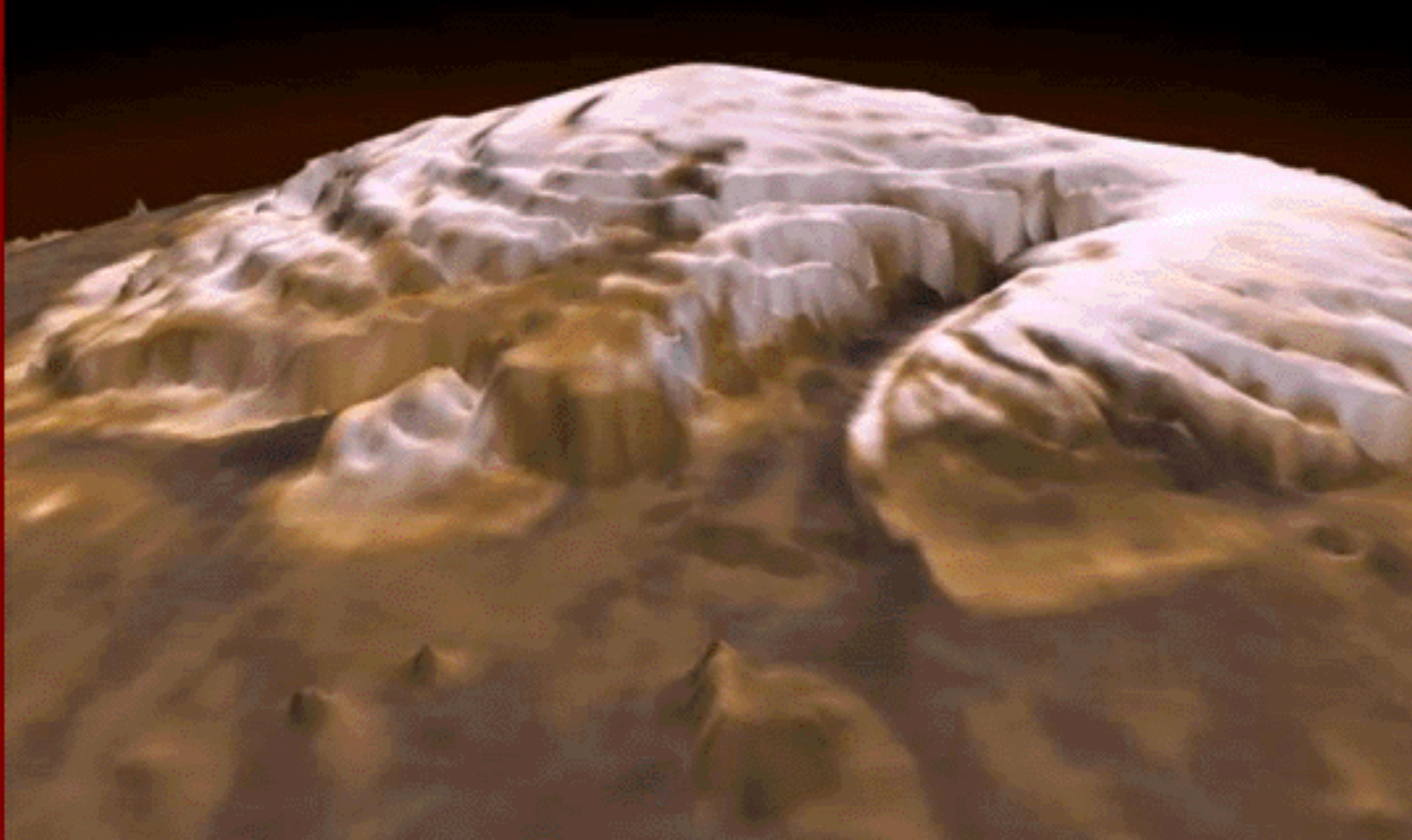
Orlando Figueroa, Director

Dr. Jim Garvin, Lead Scientist

Mars Exploration Program
NASA

Following the Water:
Program

Where can we find NLU?



Orlando Figueroa,
Director

Dr. Jim Garvin,
Lead Scientist

Mars Exploration
Program
NASA

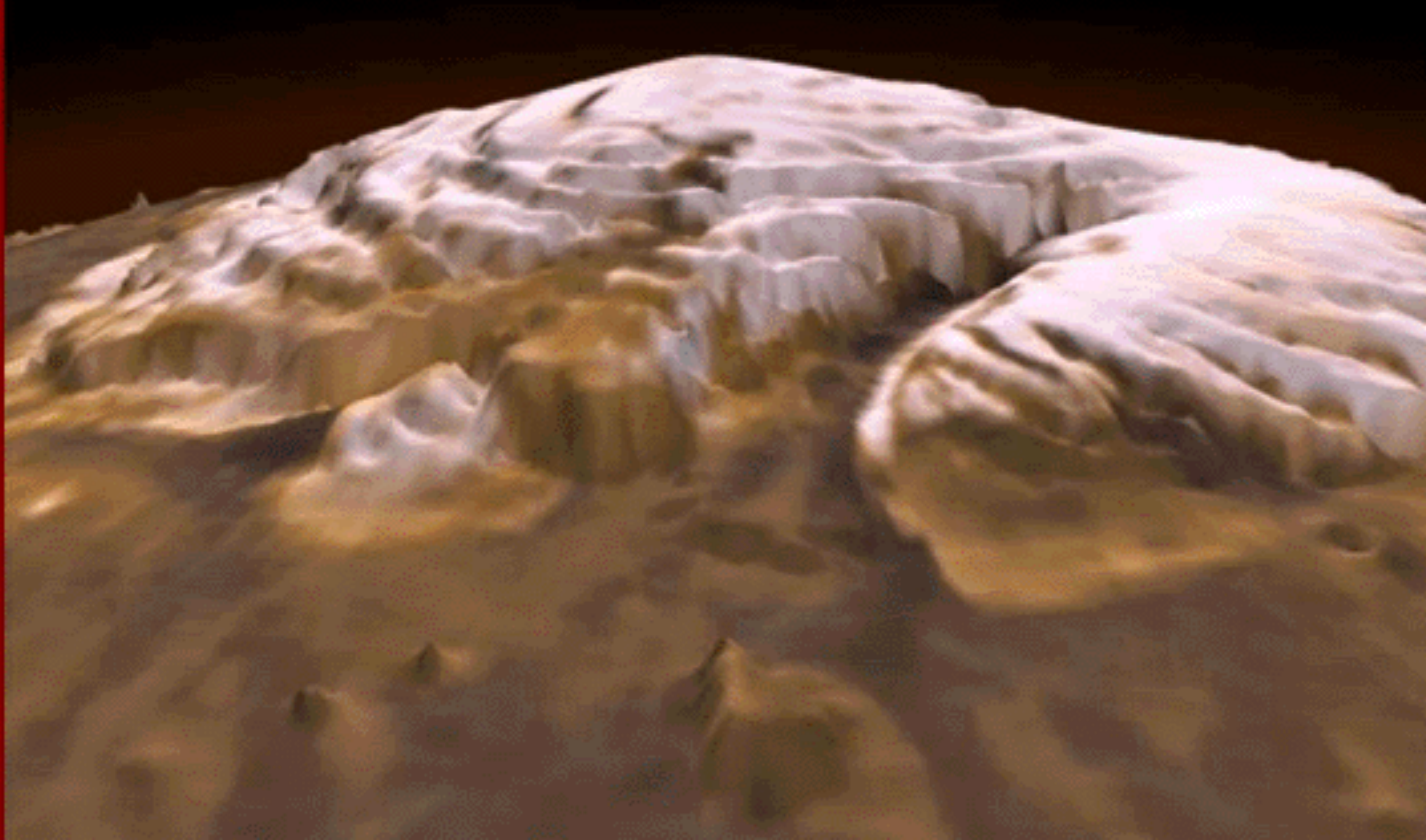
Following the Water:
Program

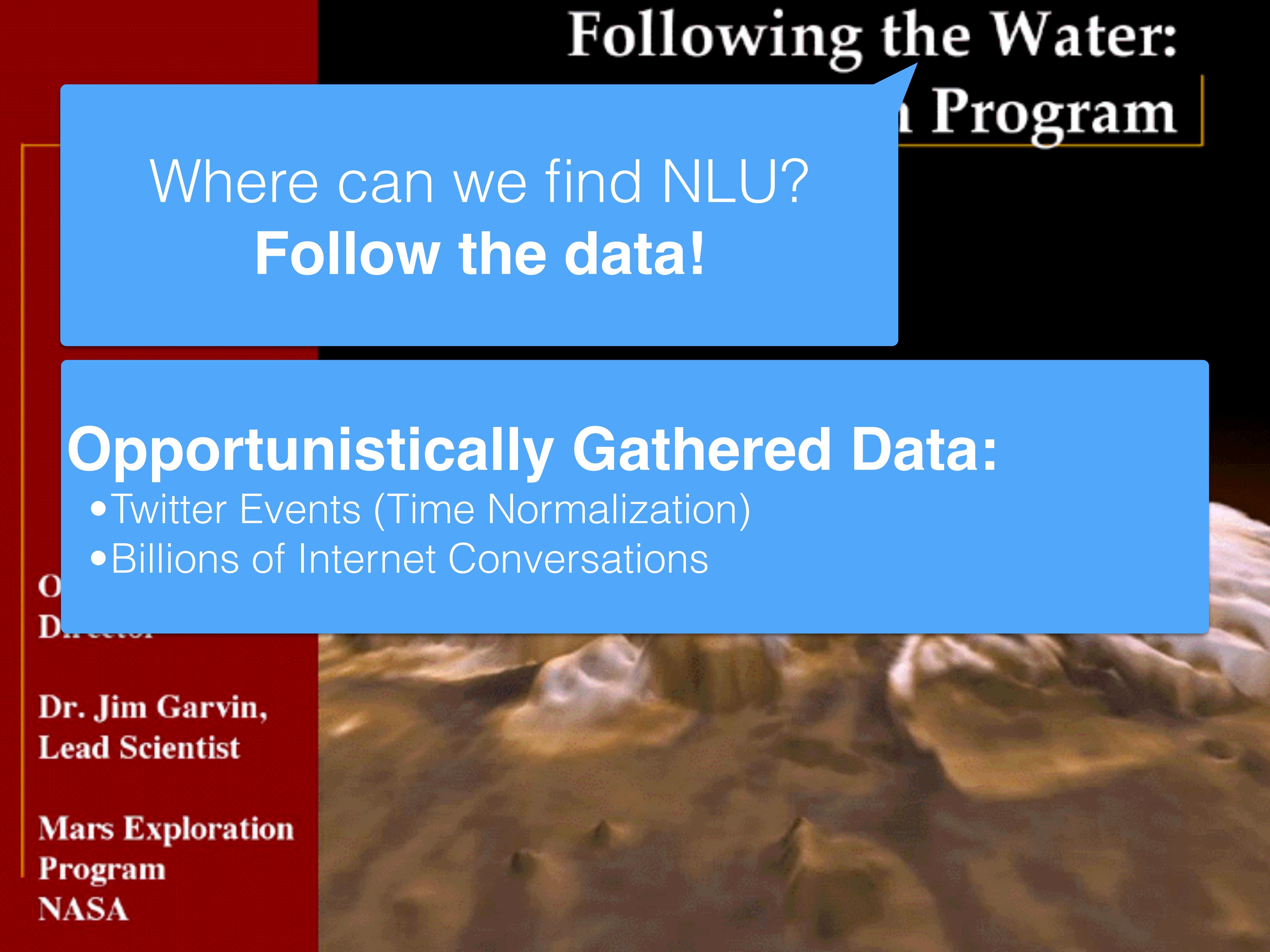
Where can we find NLU?
Follow the data!

Orlando Figueroa,
Director

Dr. Jim Garvin,
Lead Scientist

Mars Exploration
Program
NASA



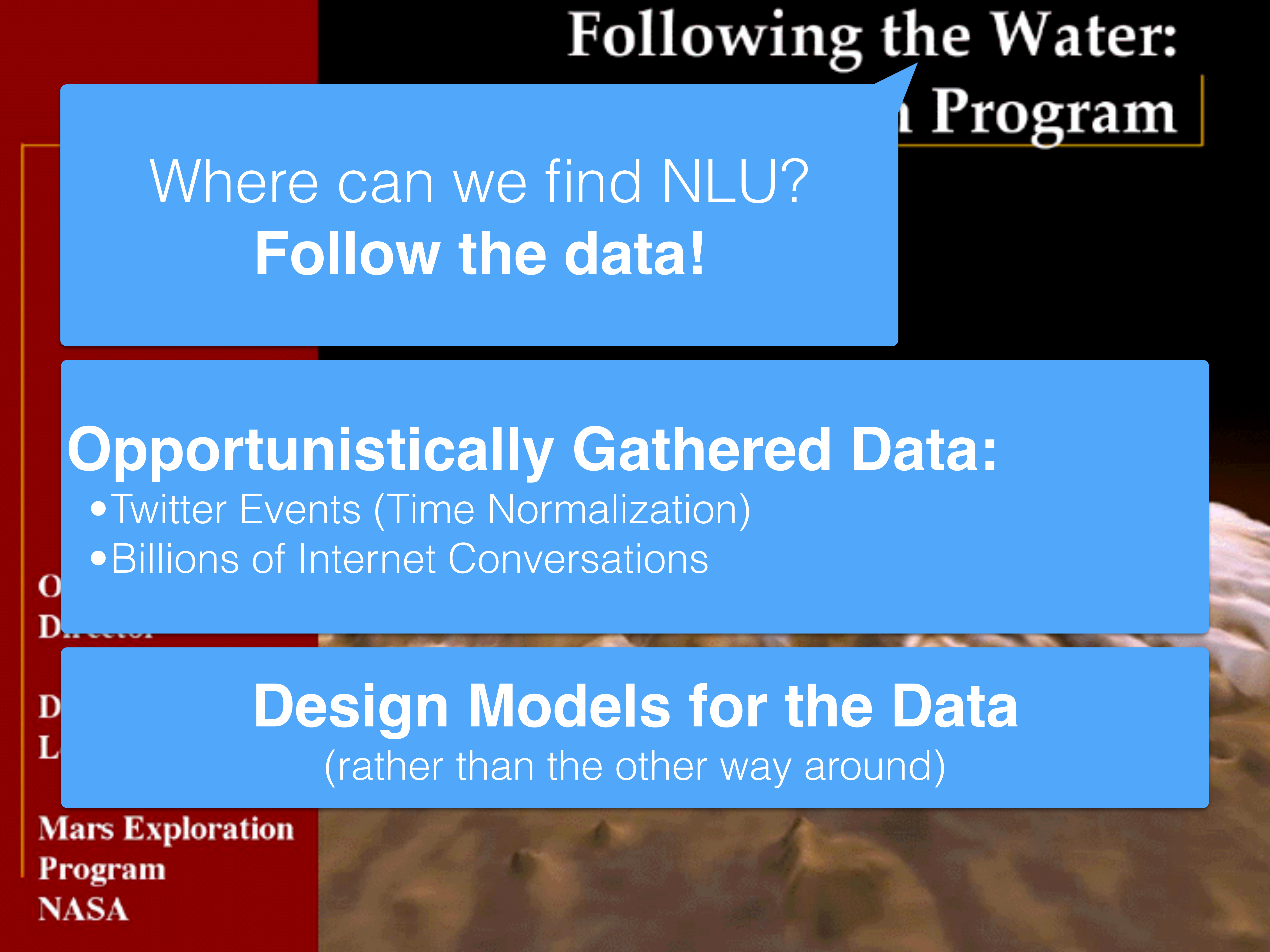


Where can we find NLU?

Follow the data!

Opportunistically Gathered Data:

- Twitter Events (Time Normalization)
- Billions of Internet Conversations


Following the Water:
Program
Where can we find NLU?
Follow the data!
Opportunistically Gathered Data:
•Twitter Events (Time Normalization)
•Billions of Internet Conversations
Design Models for the Data
(rather than the other way around)
Mars Exploration
Program
NASA

Following the Water:

...n Program

Where can we find NLU?

Follow the data!

Opportunistically Gathered Data:

- Twitter Events (Time Normalization)
- Billions of Internet Conversations

Design Models for the Data

(rather than the other way around)

Mars Exploration Program
NASA

Thank You!