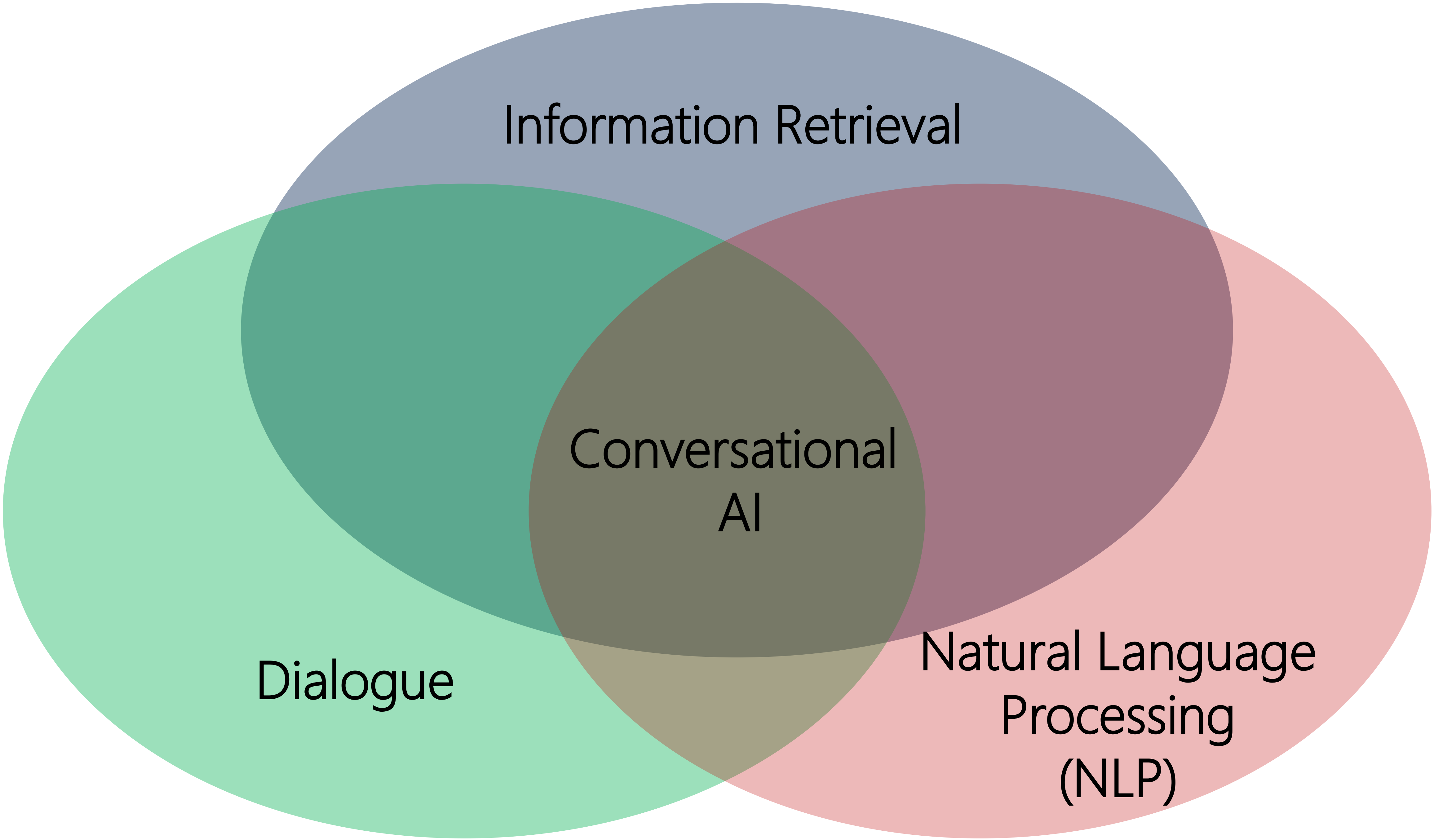
Grounding Neural Conversation Models into the Real World

Michel Galley

SCAI
October 1st, 2017

Microsoft Research AI


Information Retrieval
Conversational
AI
Dialogue
Natural Language
Processing
(NLP)

Natural Language Processing: language in, language out

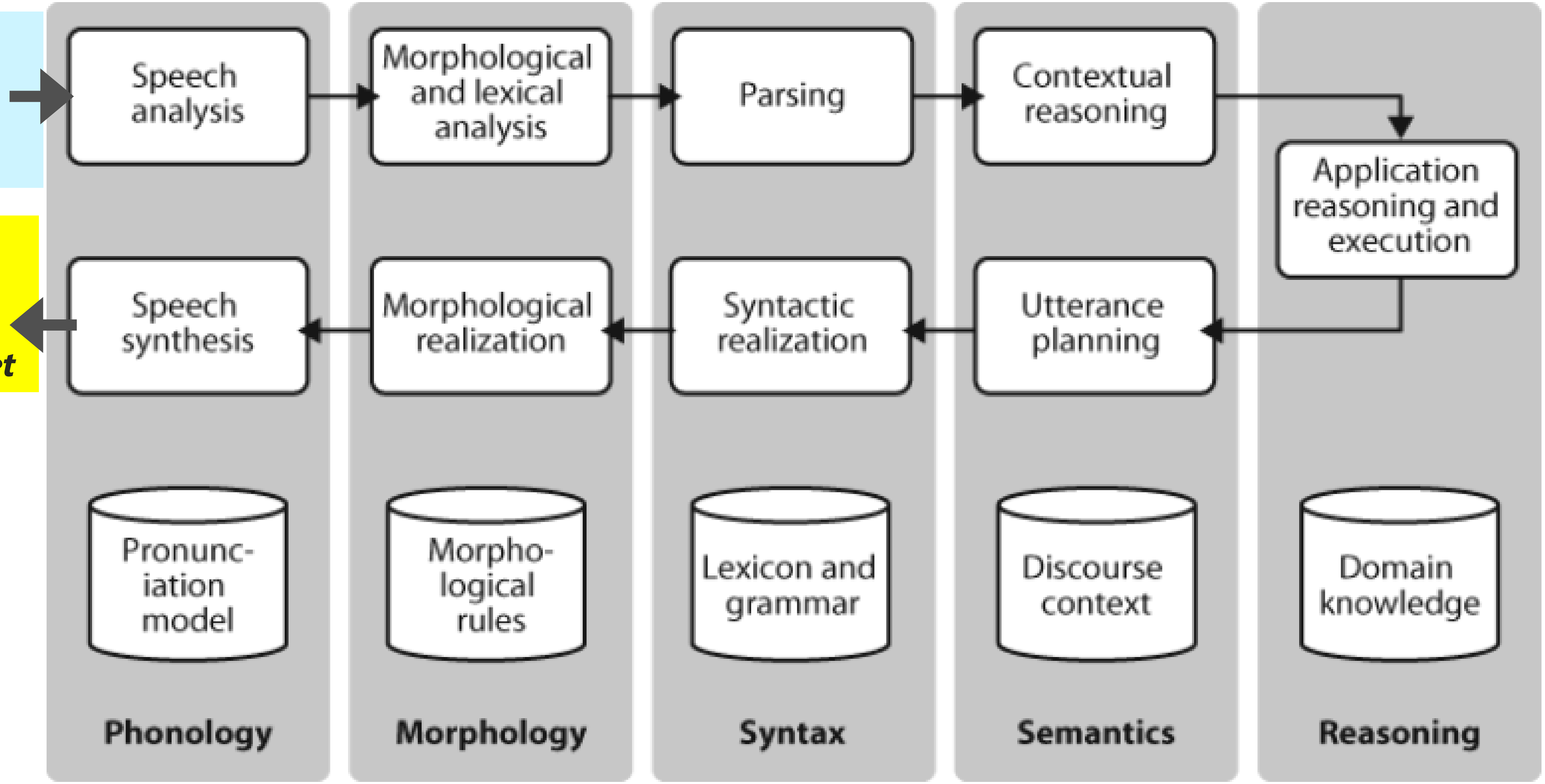
Twitter doubled its character limit

Twitter verdubbelde zijn karakterlimiet

Traditional NLP pipeline

Twitter doubled its character limit

Twitter verdubbelde zijn karakterlimiet



But the technical landscape has shifted

Twitter doubled its character limit

Twitter verdubbelde zijn karakterlimiet

End-to-End Modeling:
Language as
emergent behavior

Deep learning: recent state of the art results

Task	Test set	Metric	Best non-neural	Best neural	Source
Machine Translation	EN-DE newstest16	BLEU	31.4	34.8	http://matrix.statmt.org
	DE-EN newstest16	BLEU	35.9	39.9	http://matrix.statmt.org
Sentiment Analysis	Stanford sentiment bank	5-class Accuracy	71.0	80.7	Socher et al 2013
Question Answering	WebQuestions test set	F1	39.9	52.5	Yih et al 2015
Entity Linking	Bing Query Entity Linking set	AUC	72.3	78.2	Gao et al 2015
Image Captioning	COCO 2015 challenge	Turing test pass%	25.5	32.2	Fang et al 2015
Sentence compression	Google 10K dataset	F1	0.75	0.82	Fillipova et al, 2015

Neural systems beat previous state of the art by wide margins across an array of applications

Conversational AI?

Fully data-driven conversational AI

Twitter:

304M monthly active users
500M tweets per day (6M conversations per day)*

Other sources:

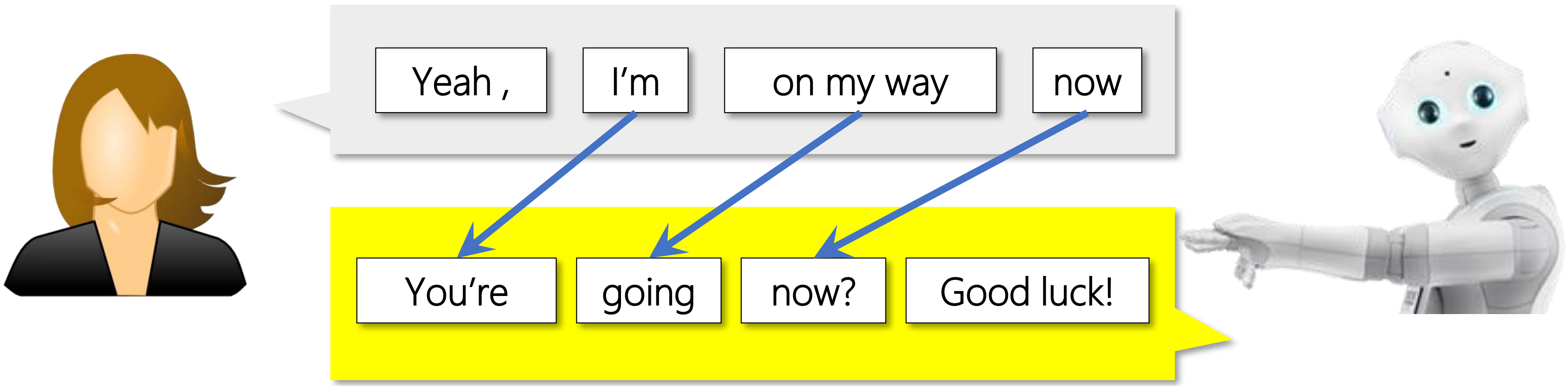
Reddit, movie subtitles,
technical data (Ubuntu), etc.

reddit

ubuntu

*: statistics as of 2015

Response Generation as Statistical Machine Translation



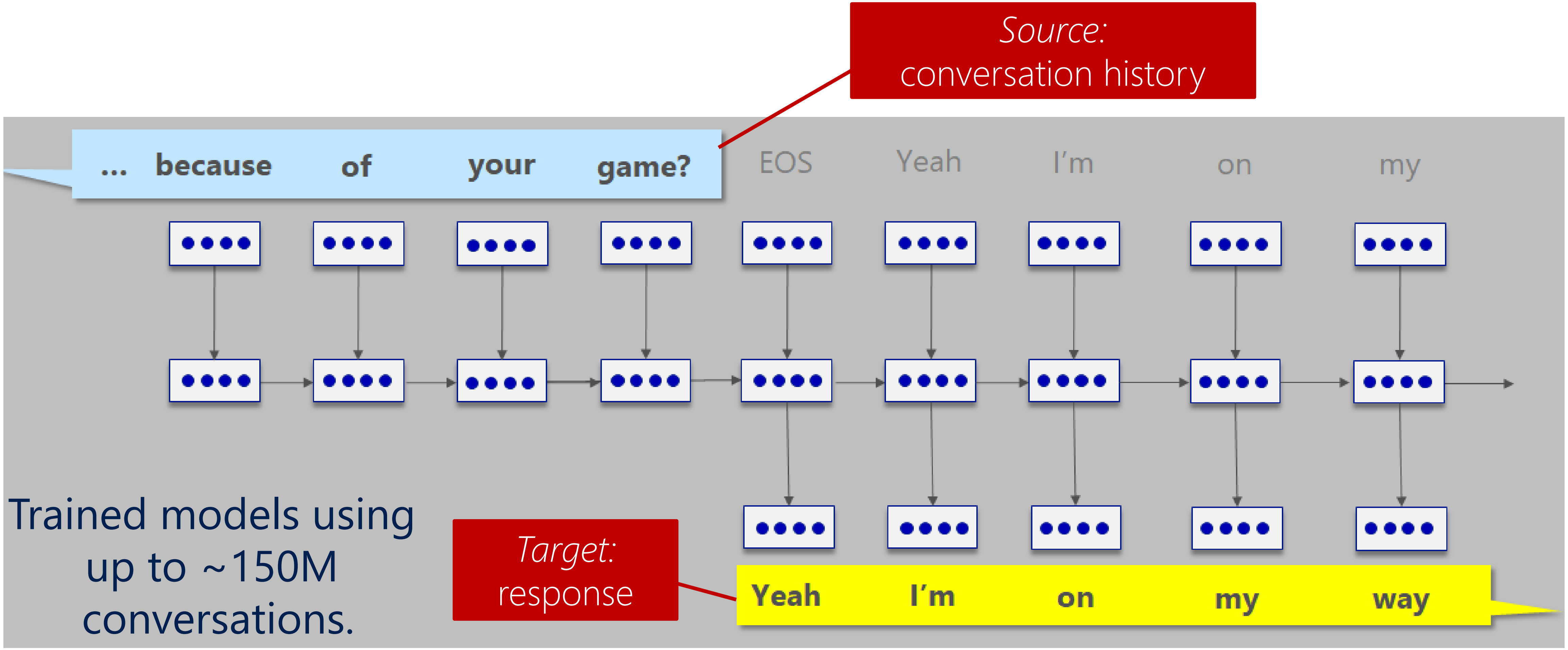
Exploit high-frequency word- and phrase-based mappings

"I'm" → "You're"   "sick" → "get better"   "lovely!" → "thanks!"

"Going to the airport" → "Have a safe flight!"

[Ritter et al., EMNLP 2011]

Neural Conversation Models



Trained models using up to ~150M conversations.

[Sordoni et al., 2015; Vinyals and Le, 2015; Shang et al., 2015; Serban et al., 2016; etc.]

Language as emergent behavior: examples

Conversation

```
Me:    Hey, want to go out for dinner later
this evening?
Sarah: sure, what time?
```

Current N-Best

```
sure, what time?
yeah sure, what time?
yeah sure. what time?
sure. what time?
sure, what time and where?
sure, where do you wanna go
sure, what time is good for you
yes sure, what time?
sure, what time
sure, what time do you wanna go
sure, what time is good for you?
sure, where do you wanna go?
sure, what time do you want to go
sure, what time is best for you
sure, what time suits you?
sure, what time you thinking
sure, what time is best for you?
sure, where do you want to go
sure, what time you thinking?
```

Language as emergent behavior: examples

Conversation

```
------------RESET-------------

Me:    Why would you purposely drop your
phone in the toilet?!
Sarah: because i don't have it in my pocket
```

Current N-Best

```
because i don't have it in my pocket
because it's the only way to get rid of it
because it's the only way to get rid of the
phone
because it's the only way to do it
because that's what i do
because i don't have a phone
because it's not my fault
because that's what i'm doing
why would i do that
because i don't like it
because i don't use it
because it's the only way
because it's the only way i can
cause i don't like it
why would i do that?
it's not my fault i don't have a phone
i don't know why i did it
i'm not that mean
because i don't have it on my phone
```

Pronominal gender, number, case

Conversation

```
------------RESET-------------

Me:    my father is coming to visit
Sarah: tell him i said hi
```

Current N-Best

```
tell him i said hi
tell him i said hello
tell him i said happy birthday
tell him i say hi
tell him i said happy birthday from me
really? i'm sure he will be happy to see you
tell him i said happy birthday to him
really? i'm sure he will be pleased to see
you
tell him i said hey
pls tell him i said hi
tell him i said happy birthday from him
please tell him i said hi
really? i'm sure he'll be pleased to see you
really? i'm sure he will be pleased to hear
that
i'm sure he will be glad to see you
really? i'm sure he will be pleased to see
you again
```

Pronominal gender, number, case (2)

Conversation

```
-------------RESET--------------

Me:    my mother is coming to visit
Sarah: tell her i said hi
```

Current N-Best

```
tell her i said hi
tell her i say hi
tell her i said hello
really? i'm sure she will be pleased to see
you
tell her i said hey
tell her i love her
really? i'm sure she will be pleased to see
you both
please tell her i said hi
tell her i say hello
tell her i hope she's ok
really? i'm sure she will be pleased to hear
that
really? i'm sure she will love it
really? i'm sure she'll love it
really? i'm sure she will be pleased to hear
from you
im sure she will be glad to see you both
```

It also learns what might be called "inferences"

Conversation

```
------------RESET-------------

Me:    Do you like James Bond movies?
Sarah: personally i don't mind them but i
don't like daniel craig
```

Current N-Best

```
personally i don't mind them but i don't
like daniel craig
personally i don't mind them. i'm not a big
bond fan
personally i don't mind them but i'm not a
big bond fan
personally i don't mind them but i'm not a
bond fan
personally i don't mind them. i'm not a bond
fan
personally i don't mind them. i'm not a huge
bond fan
personally i don't mind them but i don't
like the bond films
personally i don't mind them but i don't
like bond films
personally i don't mind them. i'm a big fan
of bond films
personally i don't think i've seen any of
his films
```

Inferences (2)

Conversation

```
-------------RESET-------------

Me:    What happens if I miss my flight?
Sarah: you won't be able to catch a flight
```

Current N-Best

```
you won't be able to catch a flight
then you won't be able to go to the airport
then you won't be able to go on the plane
you won't miss the flight
then you won't be able to go back
then you won't have to go back
then you won't be able to leave
then you won't have to leave
you won't be able to leave the plane
i don't think you will be able to get a
flight back
you won't be able to go to the airport
then you won't be able to go
you won't be able to get a flight
you won't be able to get to the airport
you won't be able to go 2 the airport
you won't be able to go on the plane
you won't get a flight
then you will have to leave
i don't think you will be able to get a
```

Inferences (3)

Conversation

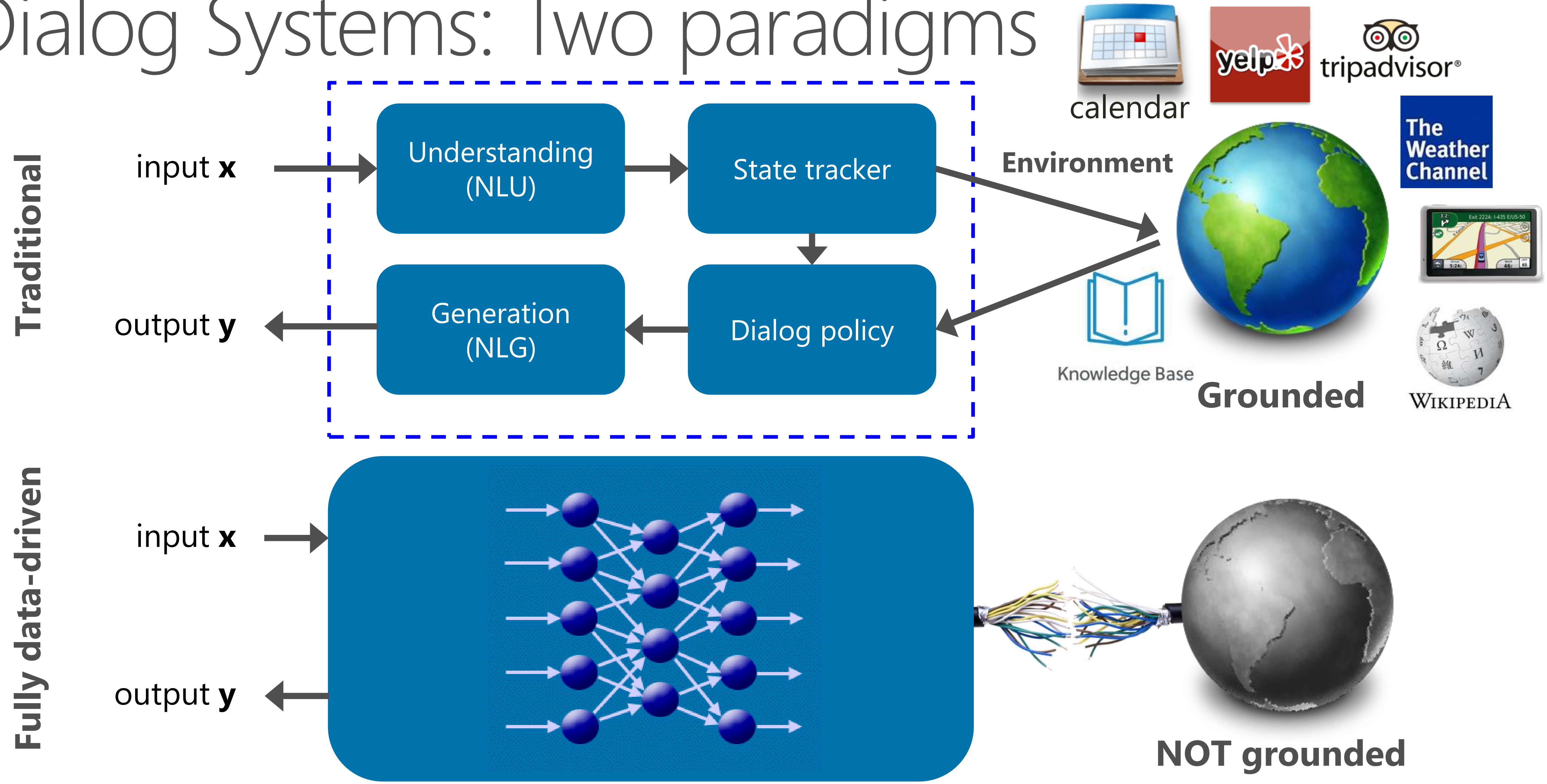
------------RESET-------------

Me: want to to to happy hour later?
Sarah: sure, what time?

Current N-Best

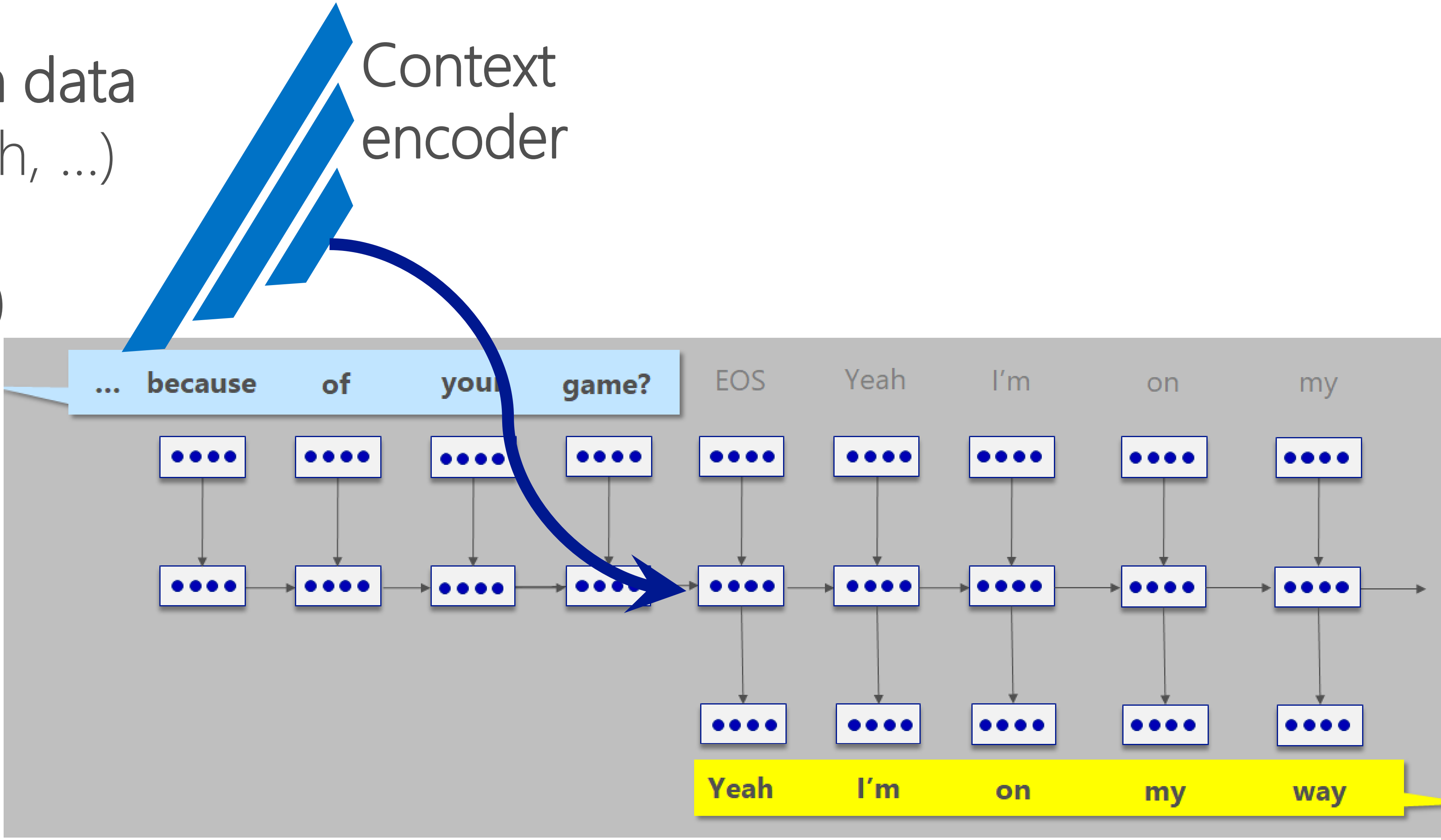
sure, what time?
sure why not
yeah sure, what time?
yeah sure why not
depends what time i finish work
yeah sure, what time
sure, what time
what time?
i'd love 2 but i'm not sure what time i'll be home
depends what time you finish work
i'd love 2 but i'm not sure what time i'll be back
i'd love 2 but i don't think i can afford it
i'd love to but i don't think i'll be able to
i'd love to but i'm on my way to work
i'd love 2 but i don't think i'll be able to
i'd love 2 but i don't have the time
let's do it, what time do you finish work

Dialog Systems: Two paradigms



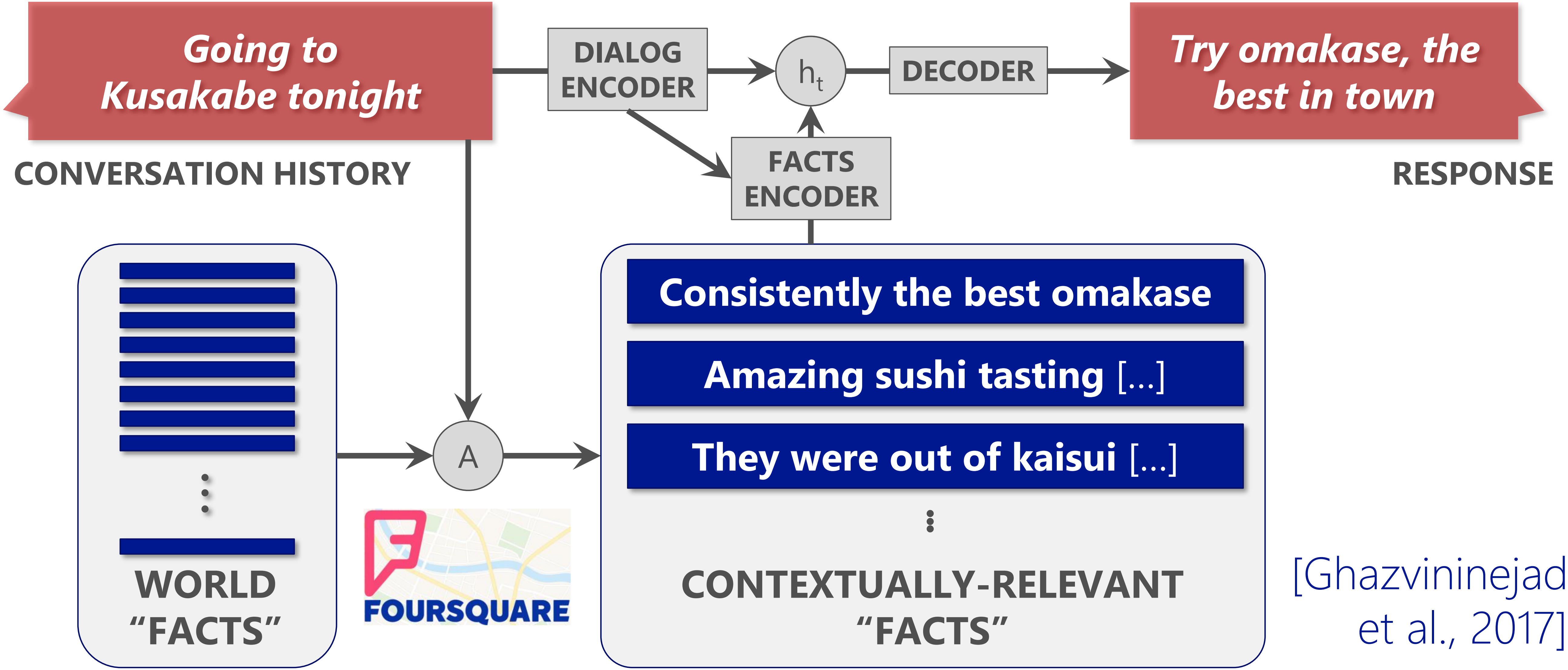
Fully Data-Driven **AND** Grounded Models

Personalization data
(ID, social graph, ...)
Device sensors
(GPS, vision, ...)



[Li et al., 2016]
[Ghazvininejad et al., 2017]
[Luan et al., 2017]
[Mostafazadeh et al., 2017]

A Knowledge-Grounded Neural Conversation Model



[Ghazvininejad et al., 2017]

"Infusing" non-conversational knowledge into conversations

You know any good **A** restaurant in **B**?

Try **C**, one of the best **D** in the city.

You know any good Japanese restaurant in Seattle?

Try Kisaku, one of the best sushi restaurants in the city.

Knowledge-grounded generated responses

FOURSQUARE

I would give ALMOST anything for some [*Mexican restaurant*] right now.

Me too. Creamy jalapeno sauce is really good.

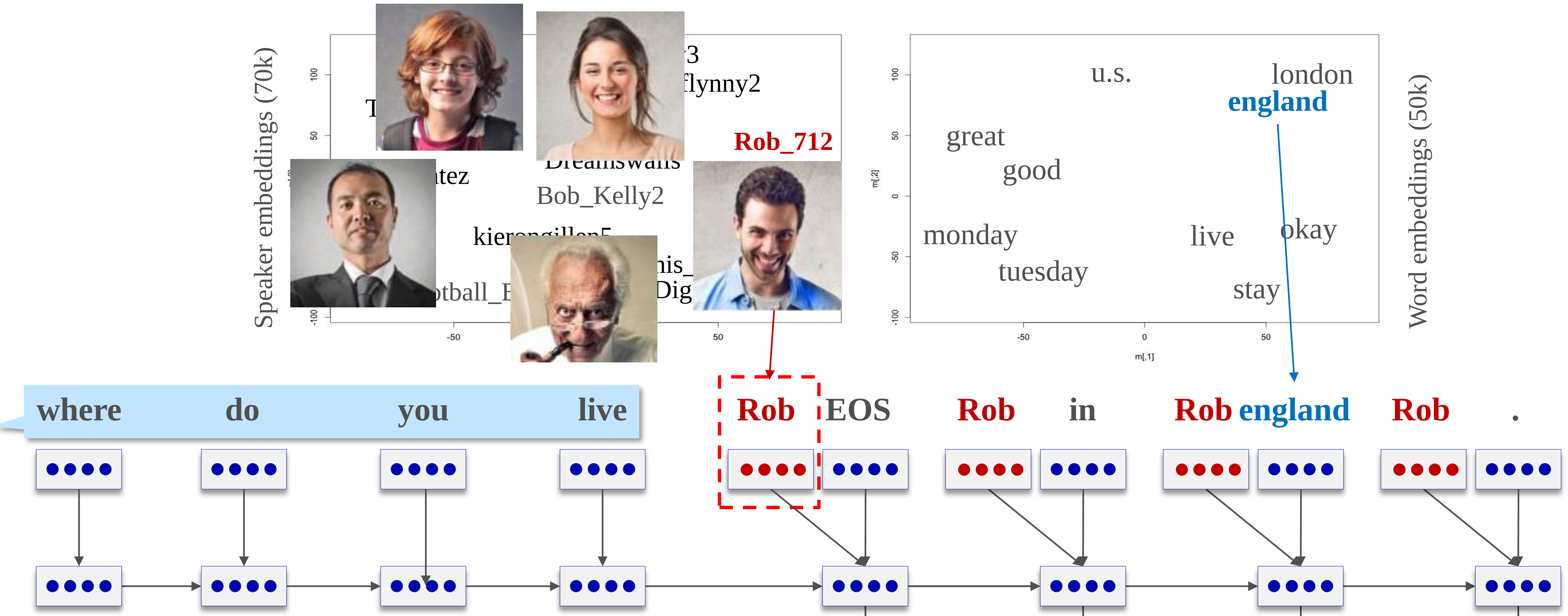
Obsessed with [*jewelry company*] :-*

oh my gosh obsessed with their bracelets and the meaning behind them!

Visiting the celebs at Los Angeles airport - [...] w/ 70 others

Nice airport terminal. Have a safe flight.

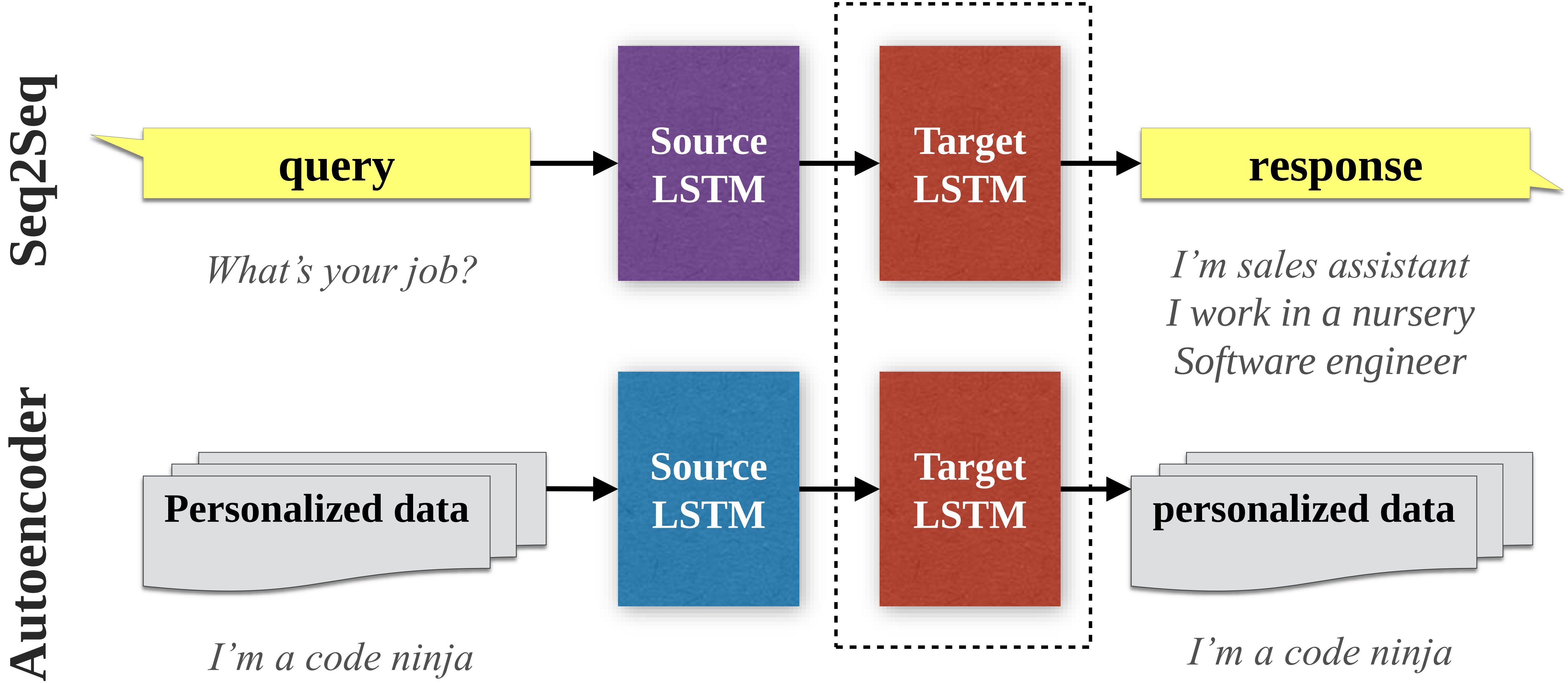
Personalized Response Generation



[Li et al. A Persona-Based Neural Conversation Model, ACL 2016]

Personalized Response Generation

[Luan et al., 2017]



Personalization: generated responses

I am getting a loop back to login page.

baseline: Ah, ok. Thanks for the info.

persona: Have you tried clearing your $\boxed{\text{cache and cookies?}}$

I reset it twice! It still doesn't work.

baseline: Let me know if there's anything I can help you with.

persona: I'm sorry to hear that. Are you receiving $\boxed{\text{any error message?}}$

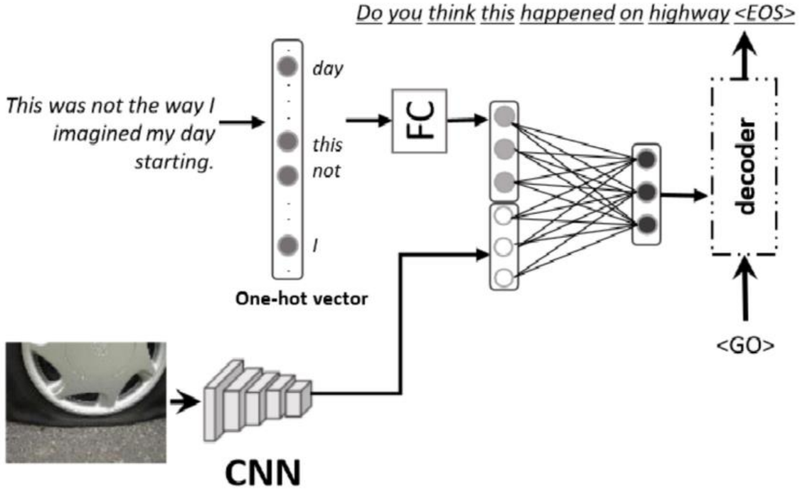
Image-Grounded Conversations

I forgot to take a pic before I took a bite.

Is that an ice cream?

The weather was amazing at the game.

Who is winning?

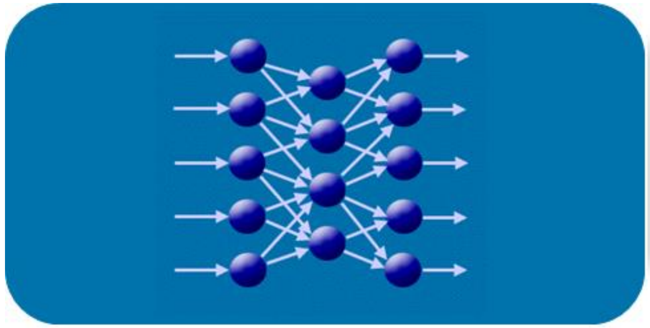


[Image-Grounded Conversations: Multimodal Context for Natural Question and Response Generation]()

N. Mostafazadeh, C. Brockett, B. Dolan, M. Galley, J. Gao, G. Spithourakis, L. Vanderwende, IJCNLP 2017

Data-driven conversation: toward more informational and "useful" dialogs

[Ghazvininejad et al., 2017; etc.]

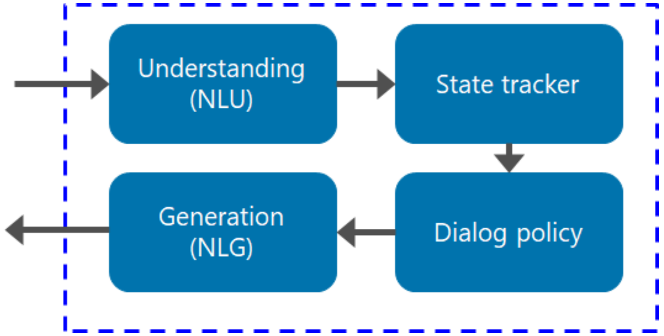


Fully data-driven
(previously ungrounded)

chitchat

GROUNDED!

informational, task-completion dialogue



Traditional dialogue systems
(grounded)

[Ritter et al., 2011, Sordoni et al., 2015; Vinyals and Le, 2015; Shang et al., 2015; Li et al., 2016; ...]

Conclusions

- Language as emergent behavior

  Learn the backbone or shell of open-domain natural conversation (e.g., question → answer, apology → downplay)

  Grammatically **well-formed** and usually socially **well-behaved**

  Capture **commonsense** "inferences": *make mouth water -> delicious*

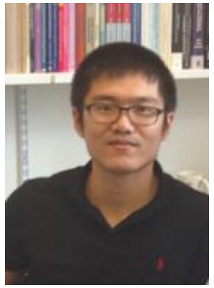
- Grounded conversational AI models

  Exploit external textual knowledge, device sensors, personal information

  Produce more informational and "**useful**" dialogues

Collaborators

Marjan Ghazvininejad
USC/ISI



Jiwei Li
Stanford

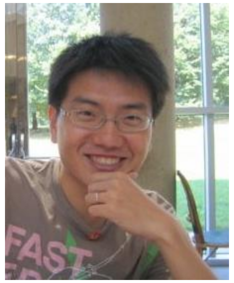
Yi Luan
U. Washington

Nasrin Mostafazadeh
U. Rochester

Alan Ritter
Ohio State U.

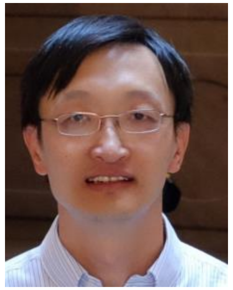
Alessandro Sordoni
Microsoft

Chris Brockett



Ming-Wei Chang

Bill Dolan



Jianfeng Gao

Chris Quirk

Scott Yih

Thank you

- Marjan Ghazvininejad, Chris Brockett, Ming-Wei Chang, Bill Dolan, Jianfeng Gao, Wen-tau Yih, Michel Galley. A Knowledge-Grounded Neural Conversation Model.
- Yi Luan, Chris Brockett, Bill Dolan, Jianfeng Gao and Michel Galley. Multi-Task Learning for Speaker-Role Adaptation in Neural Conversation Models. IJCNLP 2017.
- Jiwei Li, Michel Galley, Chris Brockett, Jianfeng Gao, and Bill Dolan. A Personalized Neural Conversation Model. In preparation for ACL 2016.
- Jiwei Li, Michel Galley, Chris Brockett, Jianfeng Gao, and Bill Dolan, A Diversity-Promoting Objective Function for Neural Conversation Models, NAACL 2016.
- Alessandro Sordoni, Michel Galley, Michael Auli, Chris Brockett, Yangfeng Ji, Meg Mitchell, Jian-Yun Nie, Jianfeng Gao, and Bill Dolan, A Neural Network Approach to Context-Sensitive Generation of Conversational Responses, NAACL 2015.
- Alan Ritter, Colin Cherry, Bill Dolan. Data-Driven Response Generation in Social Media, EMNLP 2011.

mgalley@microsoft.com