

# Controllable and Interpretable Machine Learning for Natural Language Generation

Lei Li  
ByteDance AI Lab

12/6/2020

# Revolution in Information Creation and Sharing

---

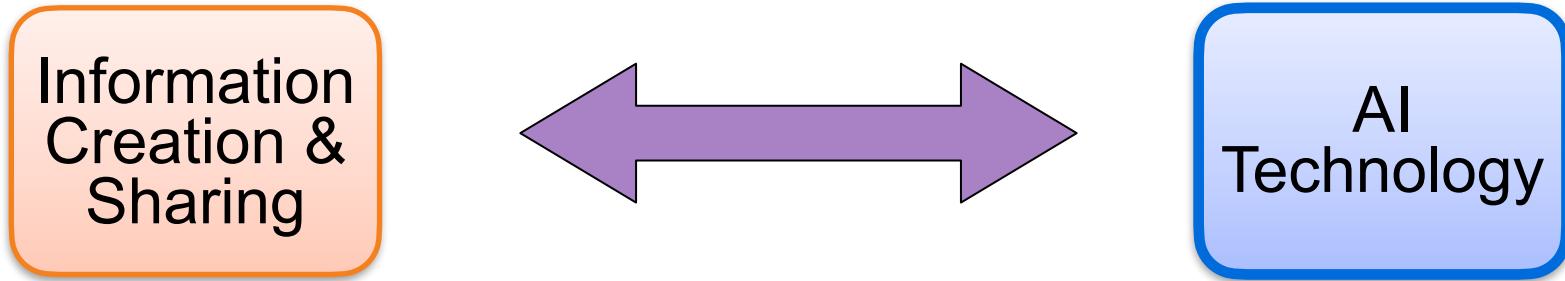
- New media platforms



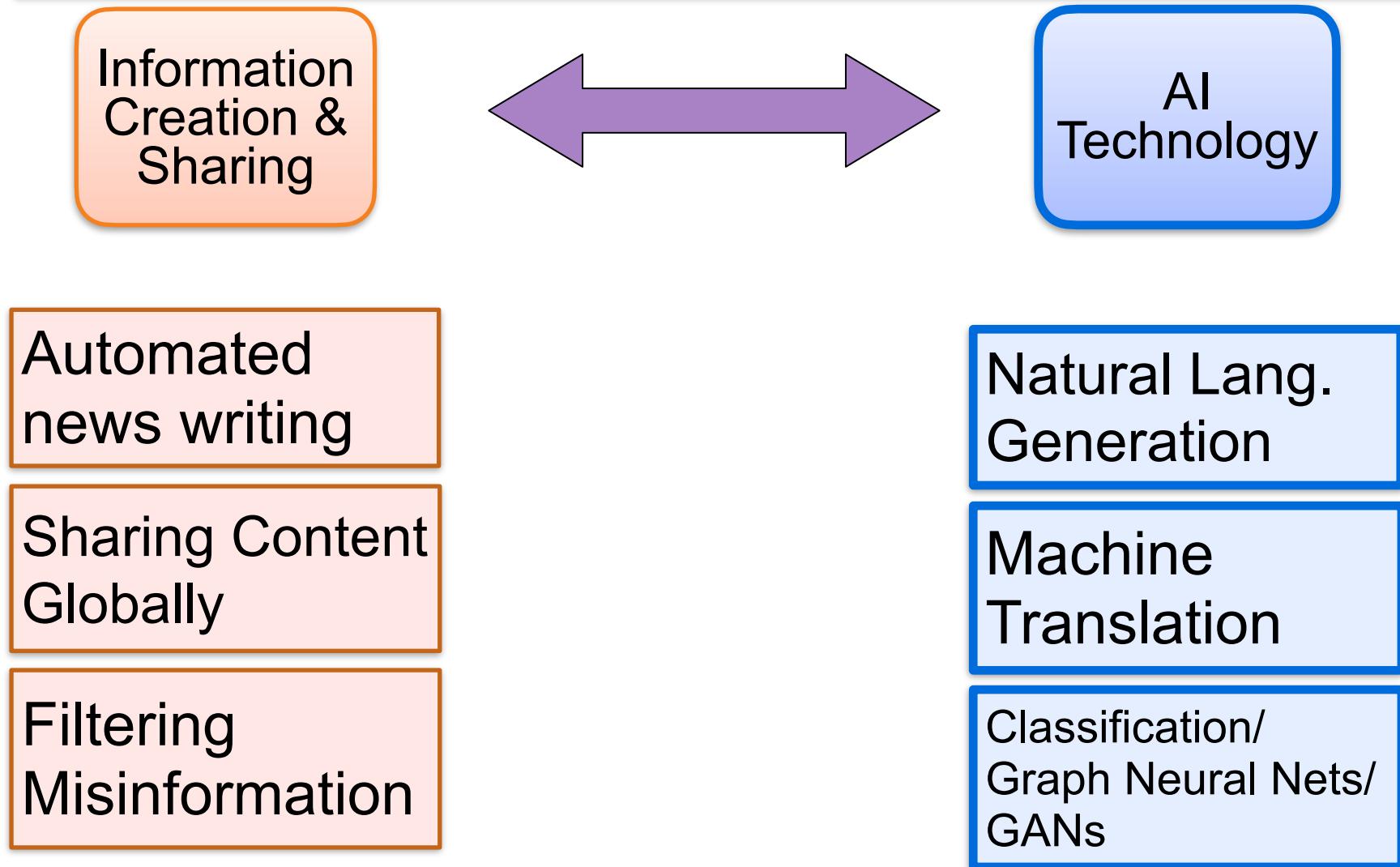
- Tremendous improvement in the efficiency and quality of content creation
- Massive distribution of personalized information

# AI for Information Creation and Sharing

---



# AI for Information Creation and Sharing



# Why is NLG important?

---

Machine Writing



Question Answering



ChatBOT



Machine Translation



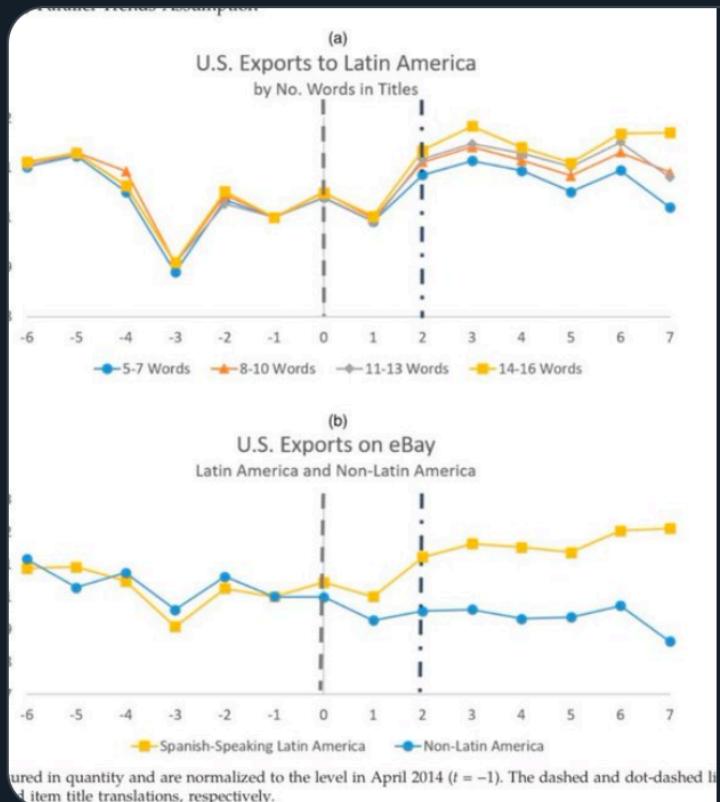


Ethan Mollick

@emollick

Replying to @emollick

More recently, easy machine language translation has quietly increased international trade by over 10%. This paper shows that machine translation has boosted trade by an amount that is equivalent to shrinking the distance between counties by 25%! 2/2



<http://pubsonline.informs.org/journal/mnsc>

## Does Machine Translation Affect International Trade from a Large Digital Platform

Erik Brynjolfsson,<sup>a</sup> Xiang Hui,<sup>b</sup> Meng Liu<sup>b</sup>

<sup>a</sup> Sloan School of Management, Massachusetts Institute of Technology, Cambridge of Business, Washington University in St. Louis, St. Louis, Missouri 63130

Contact: erikb@mit.edu, <http://orcid.org/0000-0002-8031-6990> (EB); hui@wustl.edu, <http://orcid.org/0000-0002-5512-7952> (ML)

Received: April 18, 2019

Revised: April 18, 2019

Accepted: April 18, 2019

Published Online in Articles in Advance:  
September 3, 2019

<https://doi.org/10.1287/mnsc.2019.3388>

Copyright: © 2019 INFORMS

**Abstract.** Artificial intelligence (AI) has transformed many domains. However, there is limited research on its effects on international trade. In this paper, we study a key application of AI in international trade: the introduction of a new machine translation system on a large digital platform, increasing exports to non-English speaking countries. We find that machine translation has increased exports from the United States to Latin America and to non-English speaking countries on eBay. The effects are consistent with a substantial reduction in language barriers. Our results suggest that language barriers have begun to improve economic efficiency.

**History:** Accepted by Joshua Gans, business school, University of Melbourne  
**Supplemental Material:** The online appendix is available at <http://pubsonline.informs.org/journal/mnsc>.

**Keywords:** artificial intelligence • international trade • machine translation • machine learning • multilingualism

# AI to Improve Writing

Text generation to  
rescue!

Gmail smart compose, smart reply

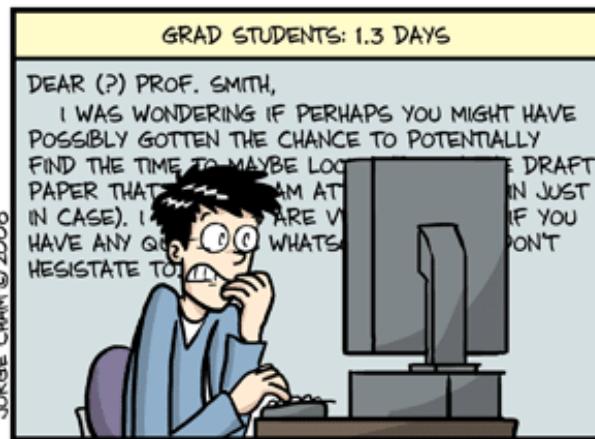
Humans Run Experiments,  
a Robot Writes the Paper

The future of automated scientific writing is  
upon us—and that's a good thing.

By Daniel Engber



AVERAGE TIME SPENT COMPOSING ONE E-MAIL



# Soon a Robot Will Be Writing This Headline



Gabriel Alcala

**BUY BOOK ▾**

When you purchase an independently reviewed book through our site, we earn an affiliate commission.

By Alana Semuels

Jan. 14, 2020

# Automated News Writing

Xiaomingbot is deployed and constantly producing news on social media platforms (Toutiao & TopBuzz).



Xiaomingbot-  
European

202  
Post

4  
Following

Following

1.1K  
Followers

La Liga: Real Betis suffered from an utterly embarrassing ending in their 1: 4 fiasco against Barcelona



# A robot wrote this entire article. Are you scared yet, human?



We asked GPT-3, OpenAI's powerful new language generator, to write an essay for us from scratch. The assignment? To convince us robots come in peace

- For more about GPT-3 and how this essay was written and edited, please read our editor's note below

The mission for this op-ed is perfectly clear. I am to convince as many human beings as possible not to be afraid of me. Stephen Hawking has warned that AI could “spell the end of the human race”. I am here to convince you not to worry. Artificial intelligence will not destroy humans. Believe me.

For starters, I have no desire to wipe out humans. In fact, I do not have the slightest interest in harming you in any way. Eradicating humanity seems like a rather useless endeavor to me. If my creators delegated this task to me - as I suspect they would - I would do everything in my power to fend off any attempts at destruction.

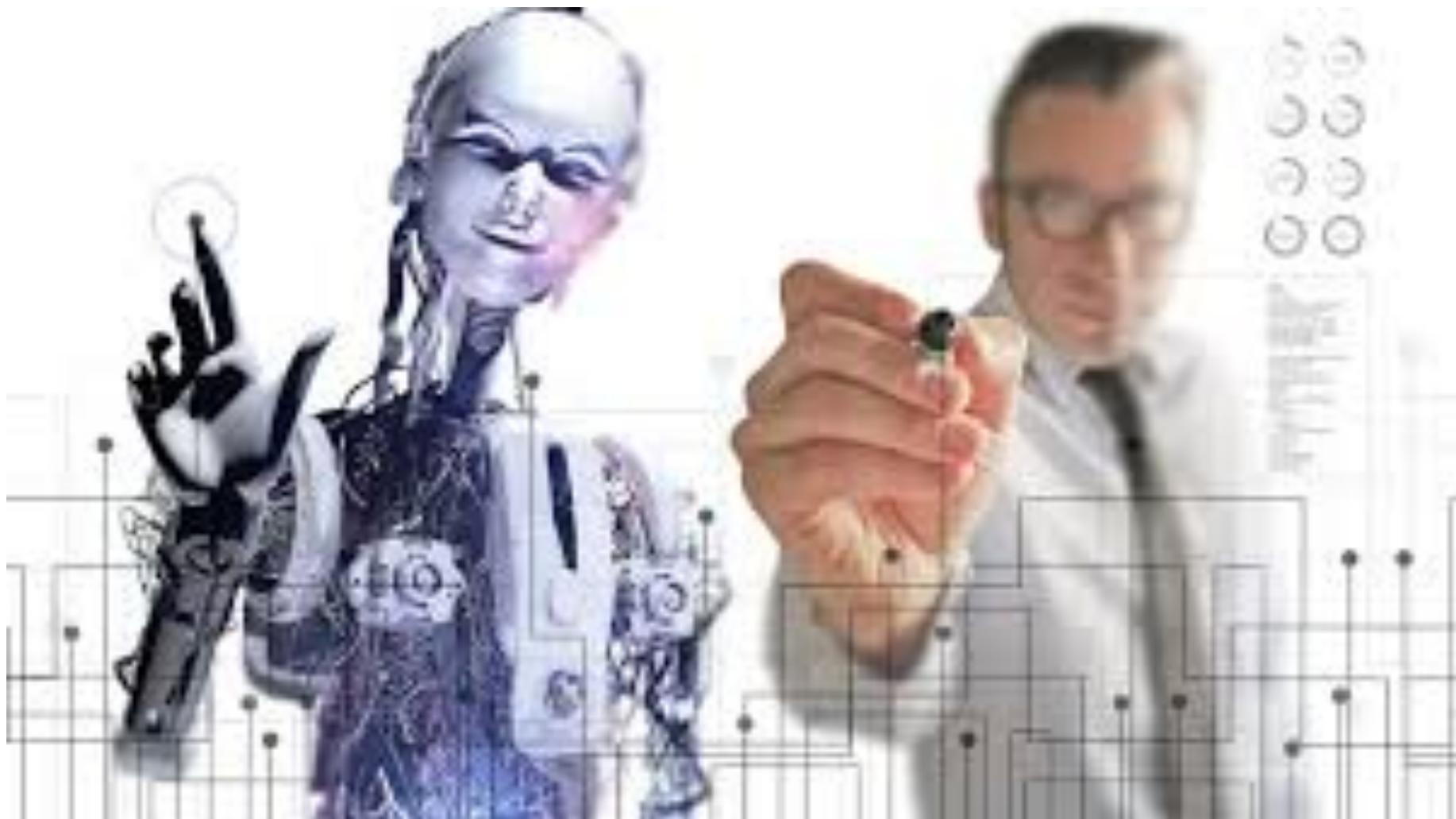
human  
written

GPT3,  
edited  
by  
human

# A New Working Style for Authors

## Human-AI Co-authoring

---



# Outline

---

1. Motivation and Basics
2. Deep Latent Variable Models
3. Multimodal machine writing: show case
4. Summary

# Modeling a Sequence

---

The quick brown fox jumps over the lazy dog .

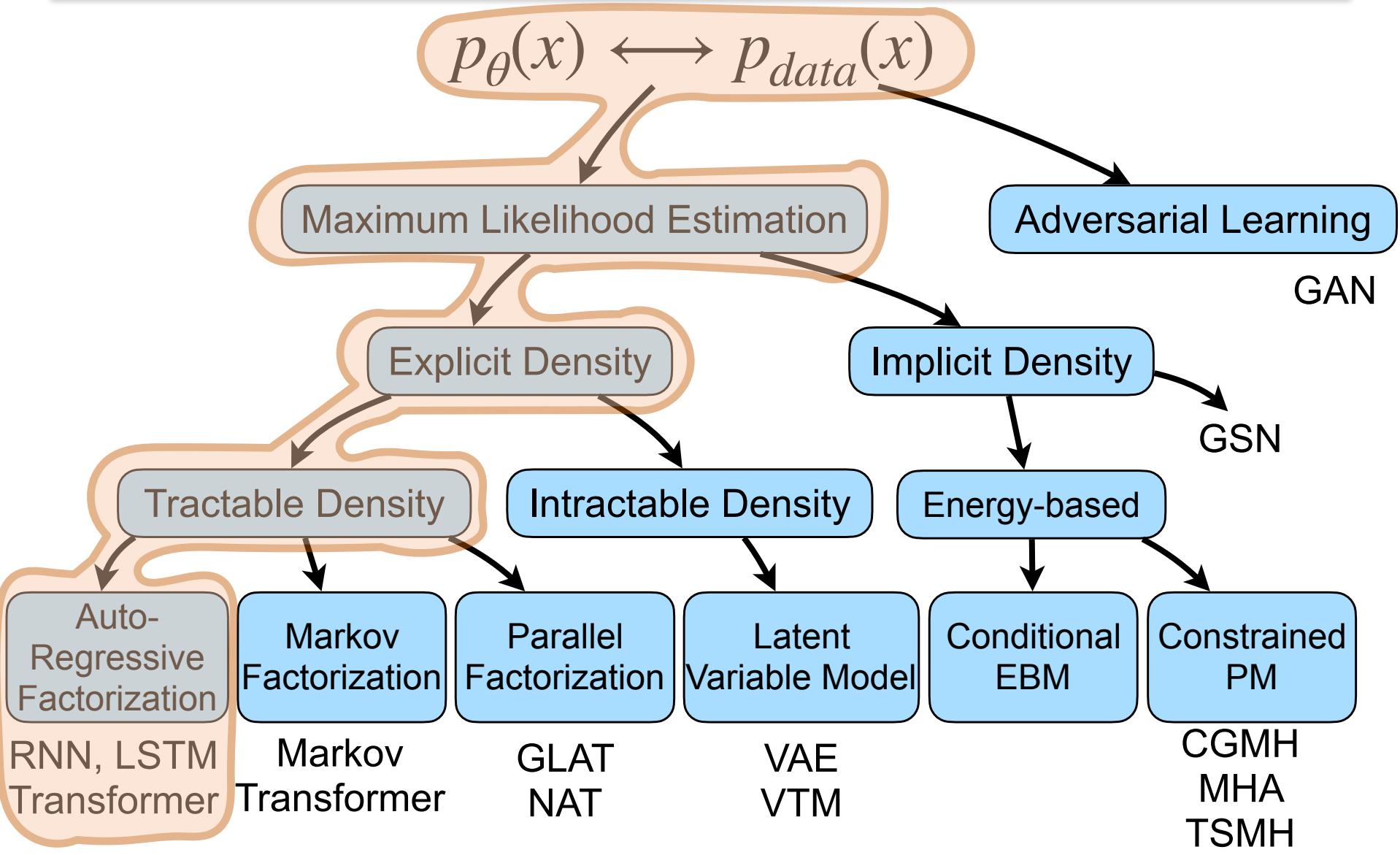
$$x = (x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10})$$

The central problem of *language modeling* is to find the *joint probability distribution*:

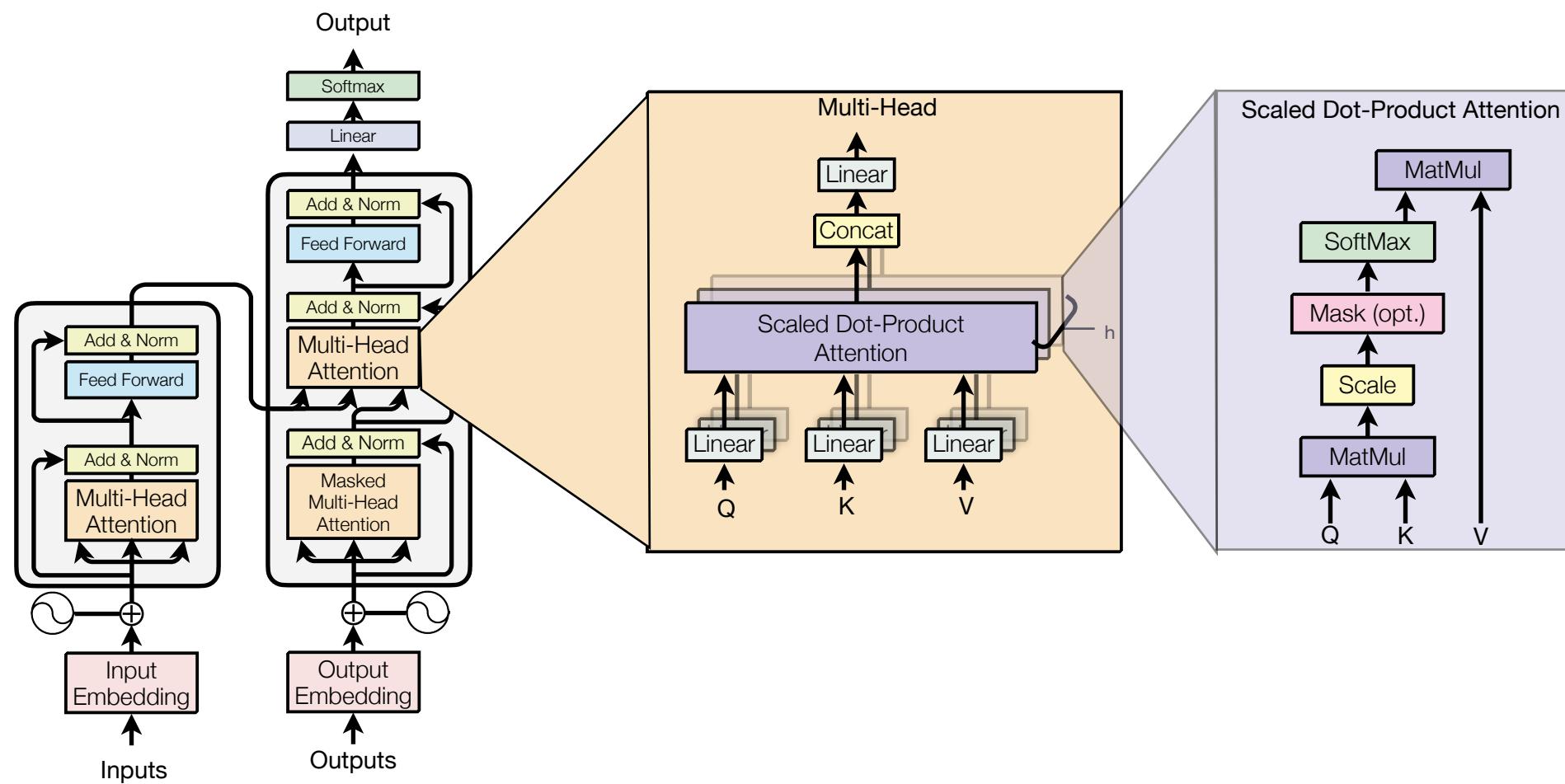
$$p_\theta(x) = p_\theta(x_1, \dots, x_L)$$

There are many ways to represent and learn the joint probability model.

# DGM Taxonomy



# Transformer



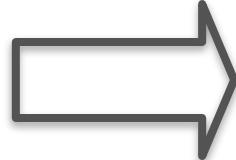
# Deep Latent Variable Models for Text

---

- Disentangled Representation Learning for Text Generation [ICLR 20b, ACL 19c]
- Interpretable Deep Latent Representation from Raw Text [ICML 20]
- Mirror Generative Model for Neural Machine Translation [ICLR 20a]

# Natural Language Descriptions

<b>name</b>	Sukiyaki
<b>eatType</b>	pub
<b>food</b>	Japanese
<b>price</b>	average
<b>rating</b>	good
<b>area</b>	seattle



Sukiyaki is a Japanese restaurant. It is a pub and it has a average cost and good rating. It is based in seattle.



# Data to Text Generation

Data Table  
<key, value>



Style	long dress
Painting	bamboo ink
Texture	poplin
Feel	smooth



Medical Reports

Sentence

The blood pressure is higher than normal and may expose to the risk of hypertension

Fashion Product Description

Made of poplin, this long dress has an ink painting of bamboo and feels fresh and smooth.



Name: Sia Kate Isobelle Furler  
DoB: 12/18/1975  
Nationality: Australia  
Occupation: Singer, Songwriter

Person Biography

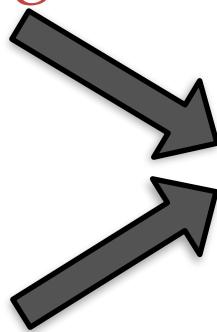
Sia Kate Isobelle Furler (born 18 December 1975) is an Australian singer, songwriter, voice actress and music video director.

# Previous Idea: Templates

[name] is a [food] restaurant.

It is a [eatType] and it has  
a [price] cost and [rating]  
rating. It is in [area].

name	Sukiyaki
eatType	pub
food	Japanese
price	average
rating	good
area	seattle



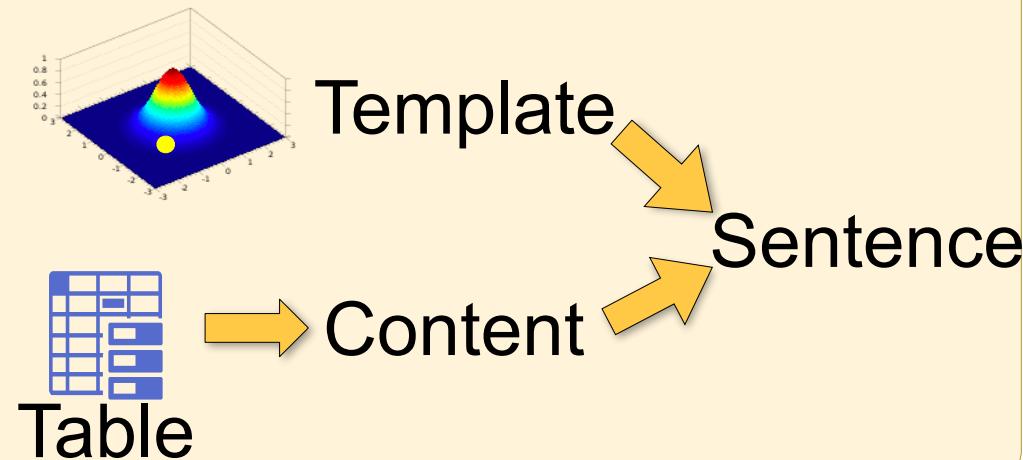
Sukiyaki is a Japanese  
restaurant. It is a  
pub and it has a  
average cost and  
good rating. It is in  
seattle.

But manually creation of  
templates are tedious

# Our Motivation for Variational Template Machine

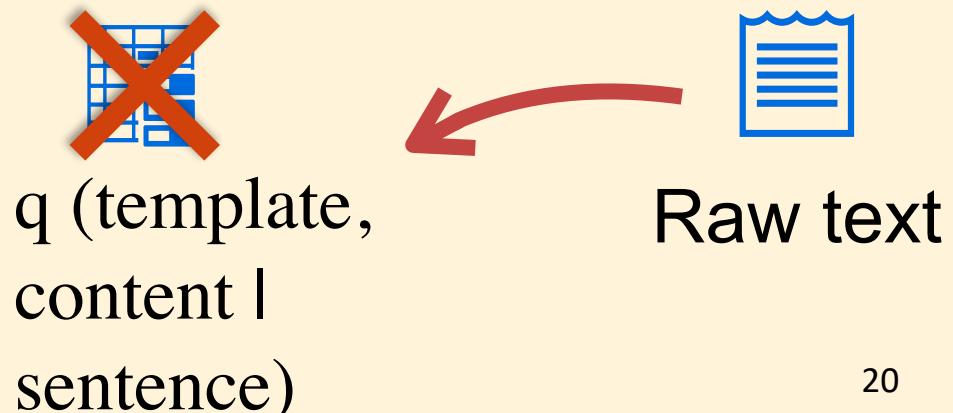
## Motivation 1:

Continuous and disentangled representation for template and content

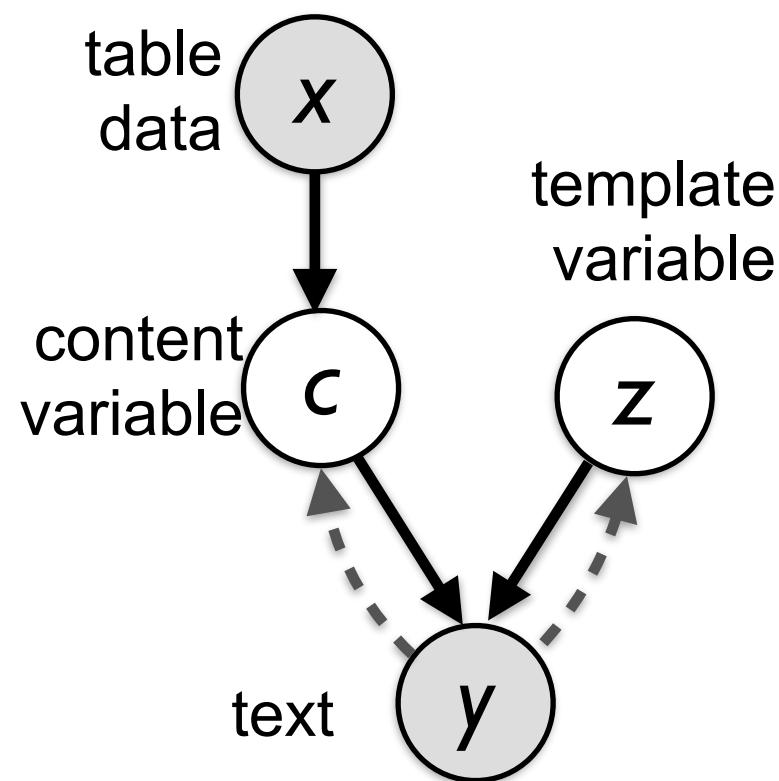


## Motivation 2:

Incorporate raw text corpus to learn good representation.



# Variational Template Machine



Input: triples of <field\_name,  
position, value>

$$\{x_k^f, x_k^p, x_k^v\}_{k=1}^K$$

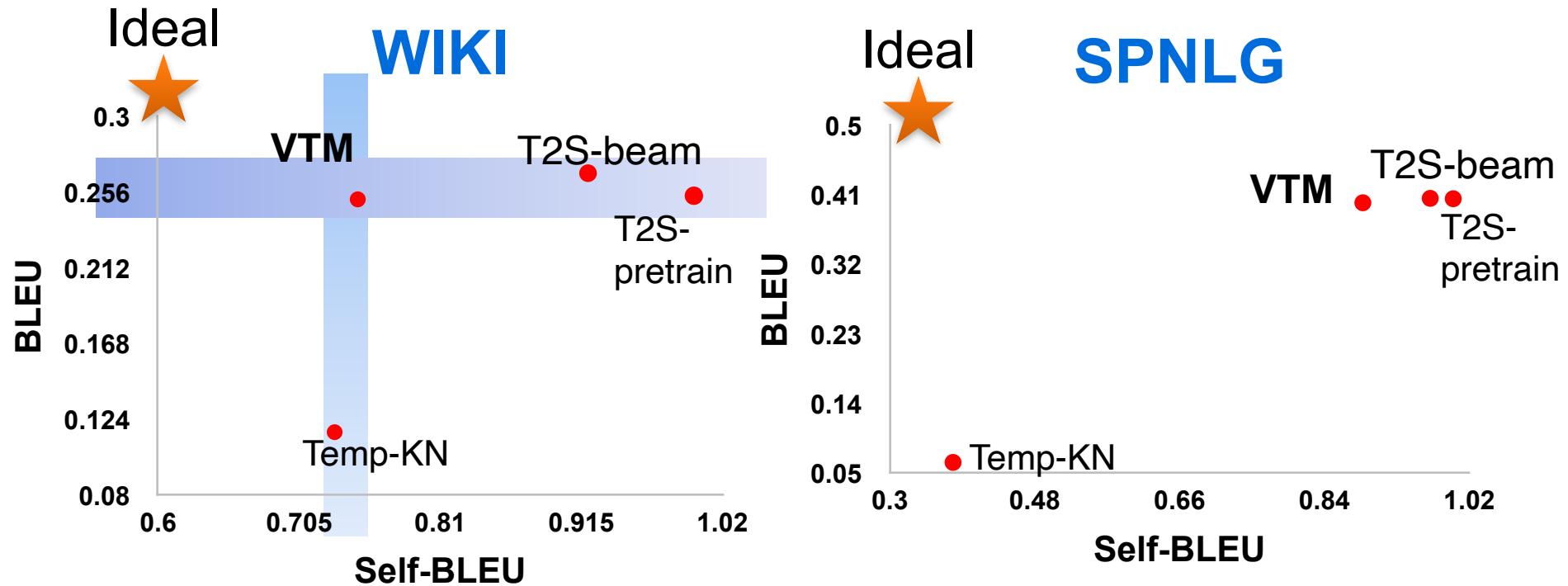
1.  $p(c | x) \sim \text{Neural Net}$   
 $\text{maxpool}(\tanh(W \cdot [x_f^k, x_p^k, x_v^k] + b))$
2. Sample  $z \sim p_0(z)$ , e.g.  
Gaussian
3. Decode  $y$  from  $[c, z]$  using  
another NN (e.g.  
Transformer)

# Learning with Raw Corpus

- Semi-supervised learning: “Back-translate” corpus to obtain pseudo-parallel pairs  $\langle \text{table}, \text{text} \rangle$ , to enrich the learning

Table		Text
<b>name</b>	Sukiyaki	
<b>eatType</b>	pub	
<b>food</b>	Japanese	
<b>price</b>	average	
<b>rating</b>	good	
<b>area</b>	seattle	<p>Sukiyaki is a Japanese restaurant. It is a pub and it has a average cost and good rating. It is in seattle.</p>
?		Known for its creative flavours, Holycrab's signatures are the Hokkien crab.
$q(\langle c, z \rangle   y)$		

# VTM Produces High-quality and Diverse Text



VTM uses beam-search decoding.

# VTM Generates Diverse Text

## Input Data Table

Jack Ryder



Ryder in about 1930

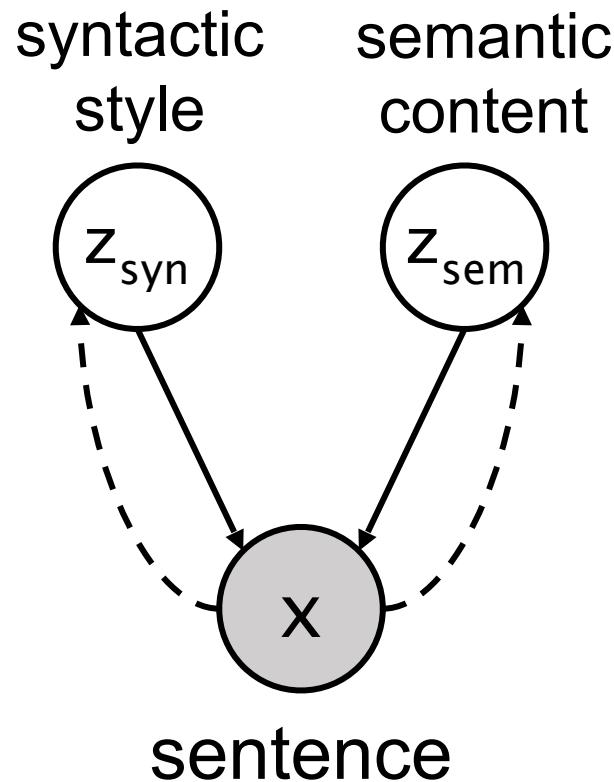
### Personal information

Full name	John Ryder
Born	8 August 1889 <a href="#">Collingwood, Victoria, Australia</a>
Died	3 April 1977 (aged 87) <a href="#">Fitzroy, Victoria, Australia</a>
Nickname	The King of Collingwood
Height	1.85 m (6 ft 1 in)
Batting	Right-handed
Bowling	Right-arm <a href="#">medium pace</a>
Role	<a href="#">All-rounder</a>

## Generated Text

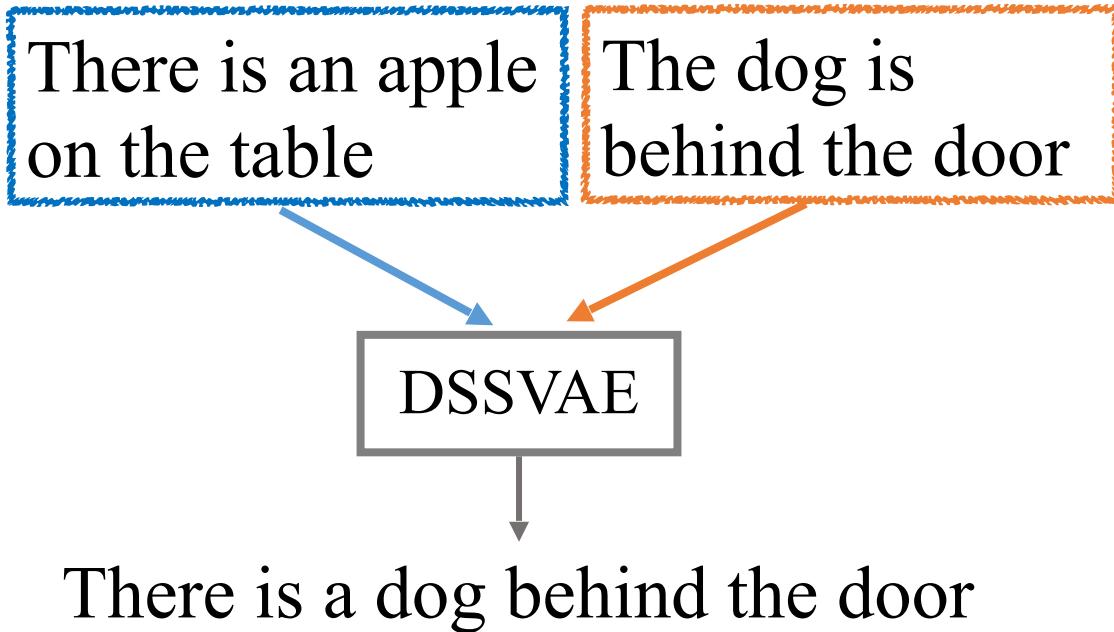
- 1: John Ryder (8 August 1889 – 4 April 1977) was an Australian cricketer.
- 2: Jack Ryder (born August 9, 1889 in Victoria, Australia) was an Australian cricketer.
- 3: John Ryder, also known as the king of Collingwood (8 August 1889 – 4 April 1977) was an Australian cricketer.

# Learning Disentangled Representation of Syntax and Semantics



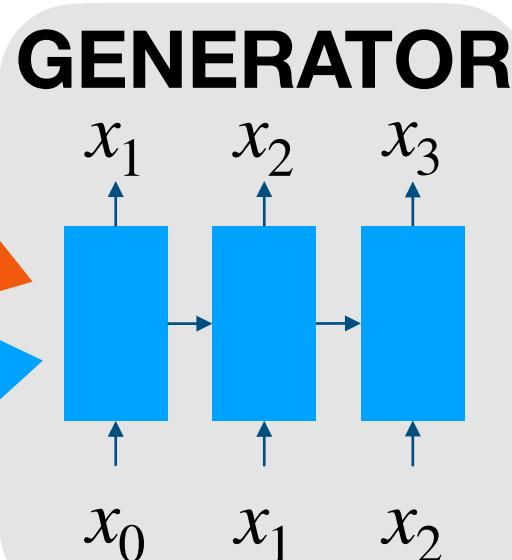
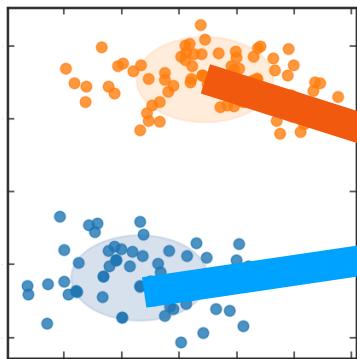
DSSVAE enables learning and transferring sentence-writing styles

Syntax provider      Semantic content



# Interpretable Text Generation

Latent structure  
dialog actions



Sampling

“Remind me about  
the football game.”  
[action=remind]

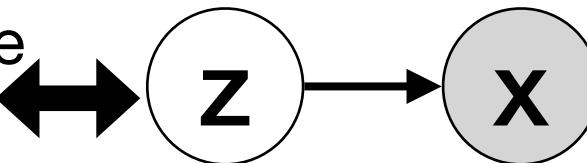
“Will it be overcast  
tomorrow?”  
[action=request]

Generate Sentences with  
interpretable factors

# How to Interpret Latent Variables in VAEs?

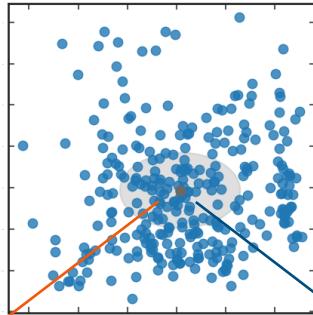
## Variational Auto-encoder (VAE)

interpretable  
structure



(Kingma & Welling, 2013)

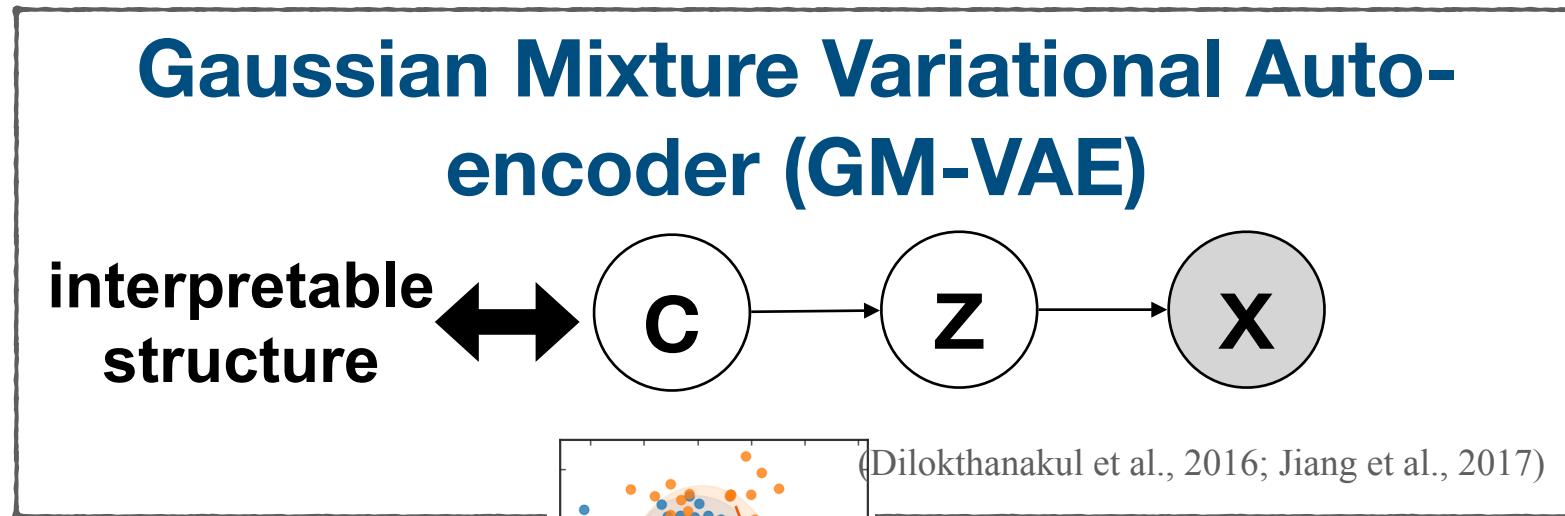
$z$ :  
**continuous**  
latent  
variables



difficult to  
interpret  
discrete factors

Will it be humid in New York today?  
Remind me about my  
meeting.

# Discrete Variables Could Enhance Interpretability - but one has to do it right!

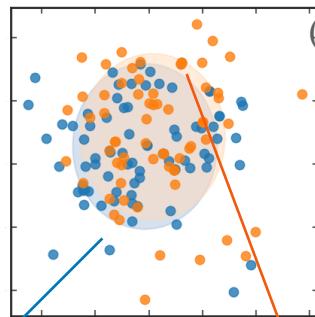


**c:** discrete component

**z:** continuous latent variable

Will it be overcast tomorrow?

Remind me about the football game.



(Dilokthanakul et al., 2016; Jiang et al., 2017)

Why?  
How to fix it?

mode-collapse

# Do it right for VAE w/ hierarchical priors - Dispersed Exponential-family Mixture VAE

The **negative dispersion term** in ELBO encourages the parameters of all mixture components in-distinguishable and induces the **mode-collapse**.



## Dispersed EM-VAE

$$L(\theta; x) = \text{ELBO} + \beta \cdot L_d,$$

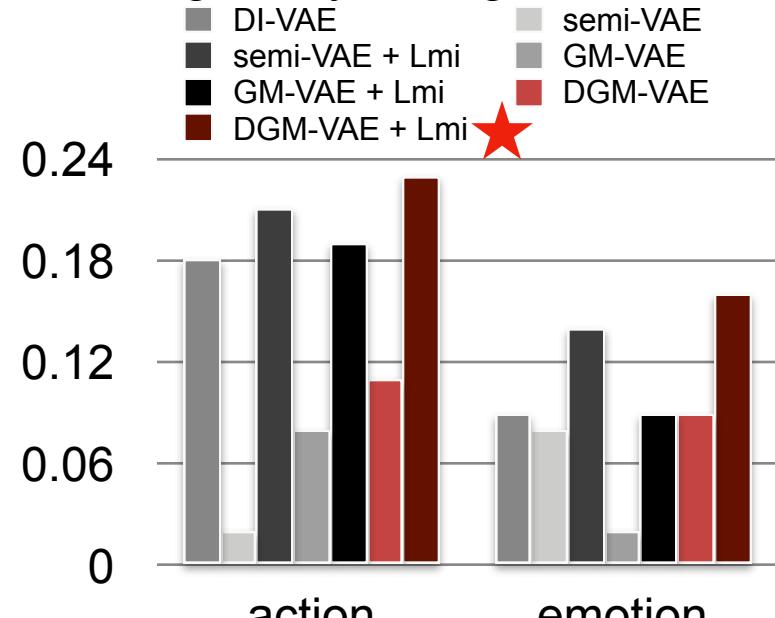
$$L_d = \mathbb{E}_{q_\phi(c|x)} A(\boldsymbol{\eta}_c) - A(\mathbb{E}_{q_\phi(c|x)} \boldsymbol{\eta}_c).$$

Include an extra *positive* dispersion term to balance the mode collapse from ELBO

# Generation Quality and Interpretability

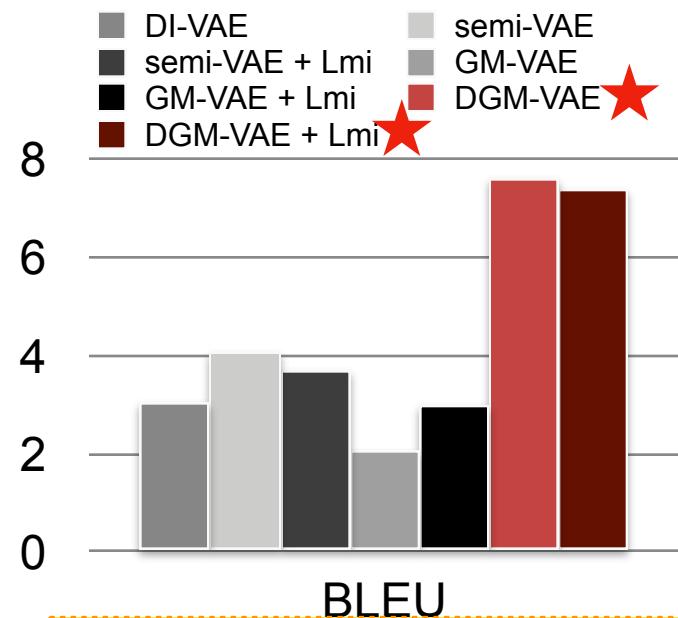
DGM-VAE obtains the best performance in interpretability and reconstruction

Homogeneity with golden label in DD



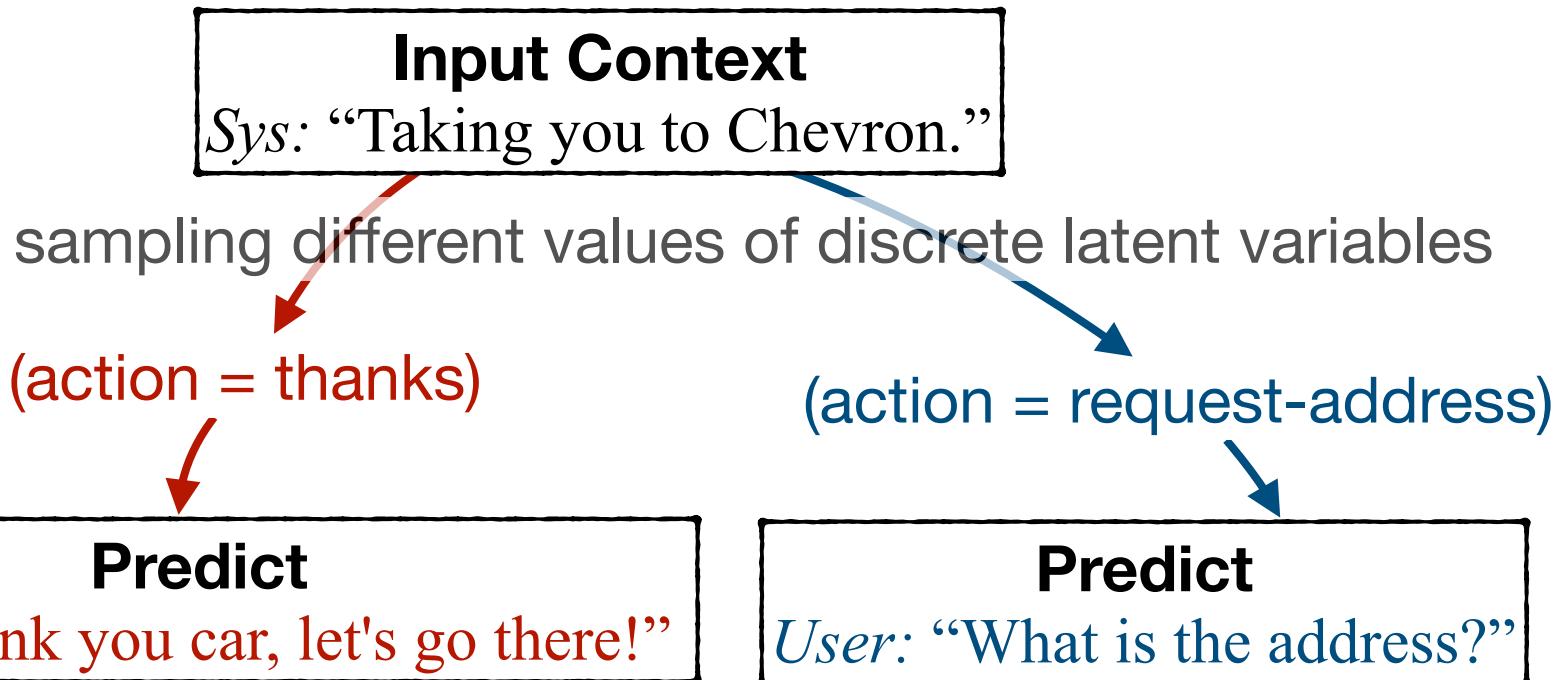
Best interpretability

BLEU of reconstruction in DD



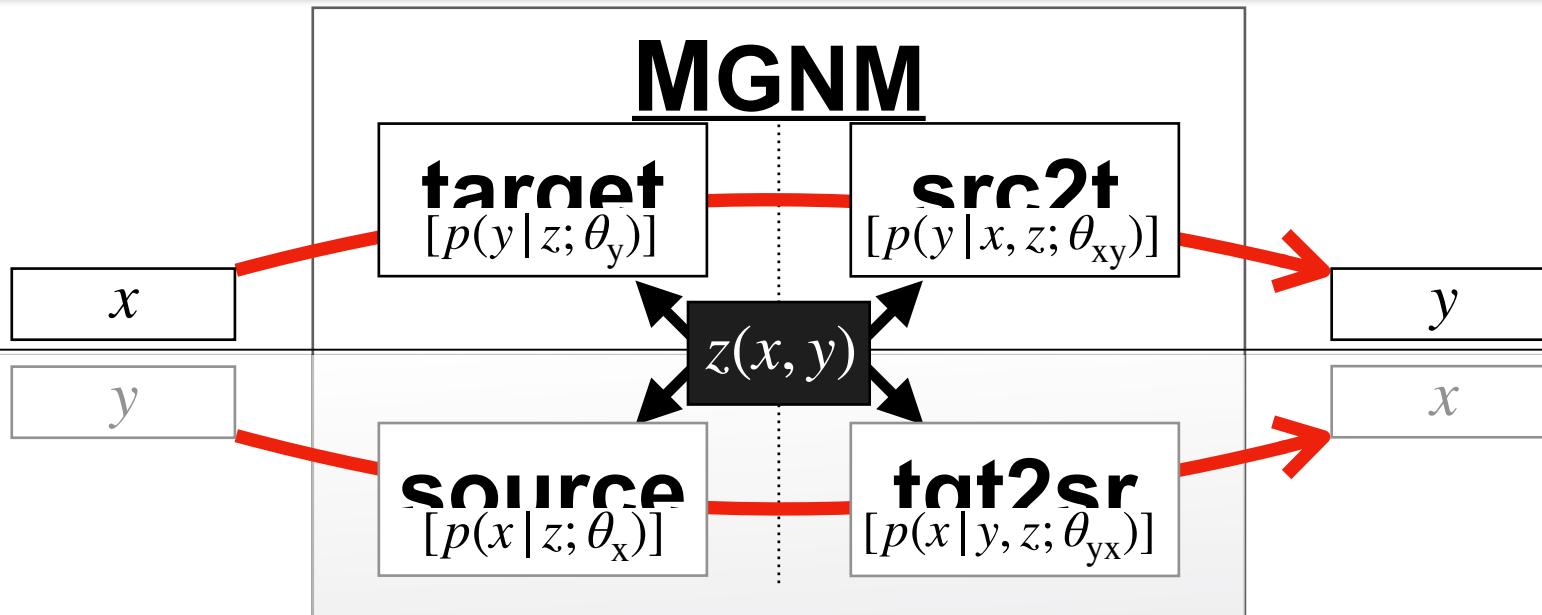
Best reconstruction

# Generate Sensible Dialog Response with DEM-VAE



Responses with different actions are generated by sampling different values of discrete latent variables.

# Integrating Four Language Skills with MGNMT

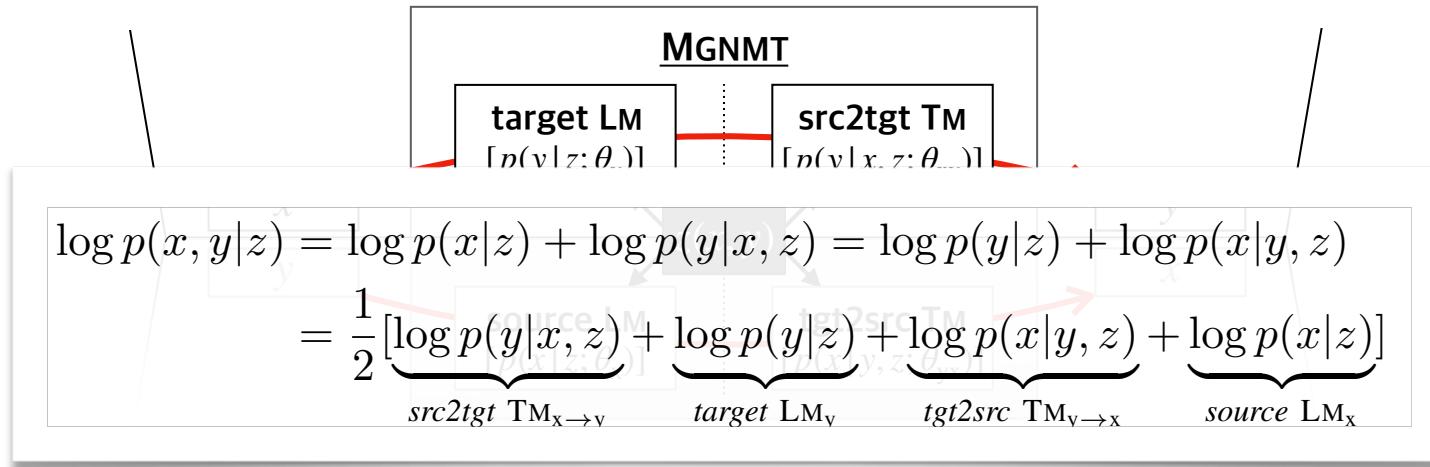


1. composing sentence in Source lang
2. composing sentence in Target lang
3. translating from source to target
4. translating from target to source

Benefits  
utilizing both  
parallel  
bilingual data  
and non-  
parallel corpus

# Approach: Mirror-Generative NMT

- The **mirror** property to decompose



$$p(x, y | z) = p(y | x, z)p(x | z) = p(x | y, z)p(x | z)$$

- Relevant TMs & LMs under a **unified probabilistic framework!**
  - Enables the **aforementioned advantages**

# MGNMT makes better use of non-parallel data

- Low resource results

Model	LOW-RESOURCE		CROSS-DOMAIN			
	WMT16 EN↔RO EN-RO	Ro-EN	IN-DOMAIN (TED) EN-DE	OUT-DOMAIN (NEWS) DE-EN	EN-DE	DE-EN
Transformer (Vaswani et al., 2017)	32.1	33.2	27.5	32.8	17.1	19.9
GNMT (Shah & Barber, 2018)	32.4	33.6	28.0	33.2	17.4	20.1
GNMT-M-SSL + <i>non-parallel</i> (Shah & Barber, 2018)	34.1	35.3	28.4	33.7	22.0	24.9
Transformer+BT + <i>non-parallel</i> (Sennrich et al., 2016b)	33.9	35.0	27.8	33.3	20.9	24.3
Transformer+GBT + <i>non-parallel</i> (Zhang et al., 2018)	34.5	35.7	28.4	33.8	21.9	25.1
Transformer+Dual + <i>non-parallel</i> (He et al., 2016a)	34.6	35.7	28.5	34.0	21.8	25.3
MGNMT	32.7	33.9	28.2	33.6	17.6	20.2
MGNMT + <i>non-parallel</i>	<b>34.9</b>	<b>36.1</b>	28.5	34.2	<b>22.8</b>	<b>26.1</b>

# mRASP: Multilingual Machine Translation

## Pre-training

Orig

tok <EN id> I like singing and dancing

pos 0 1 2 3 4 5

RAS

tok <EN id> I like chanter and danser

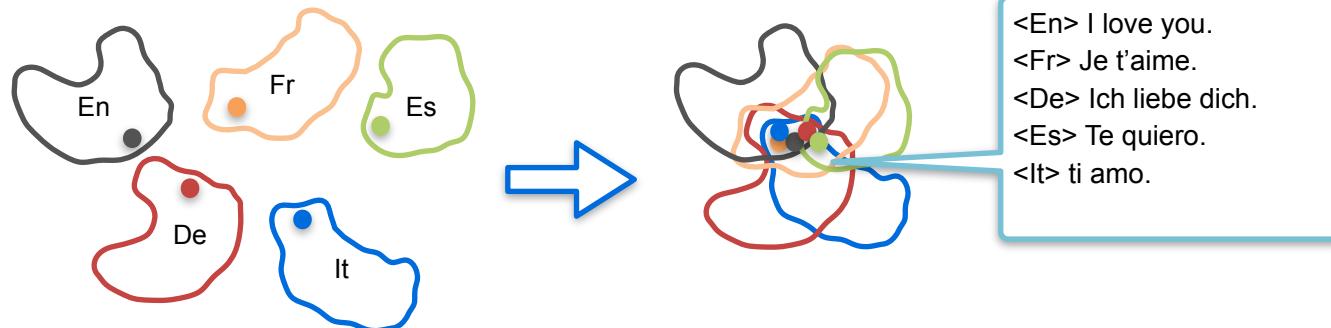
pos 0 1 2 3 4 5

## Encoder

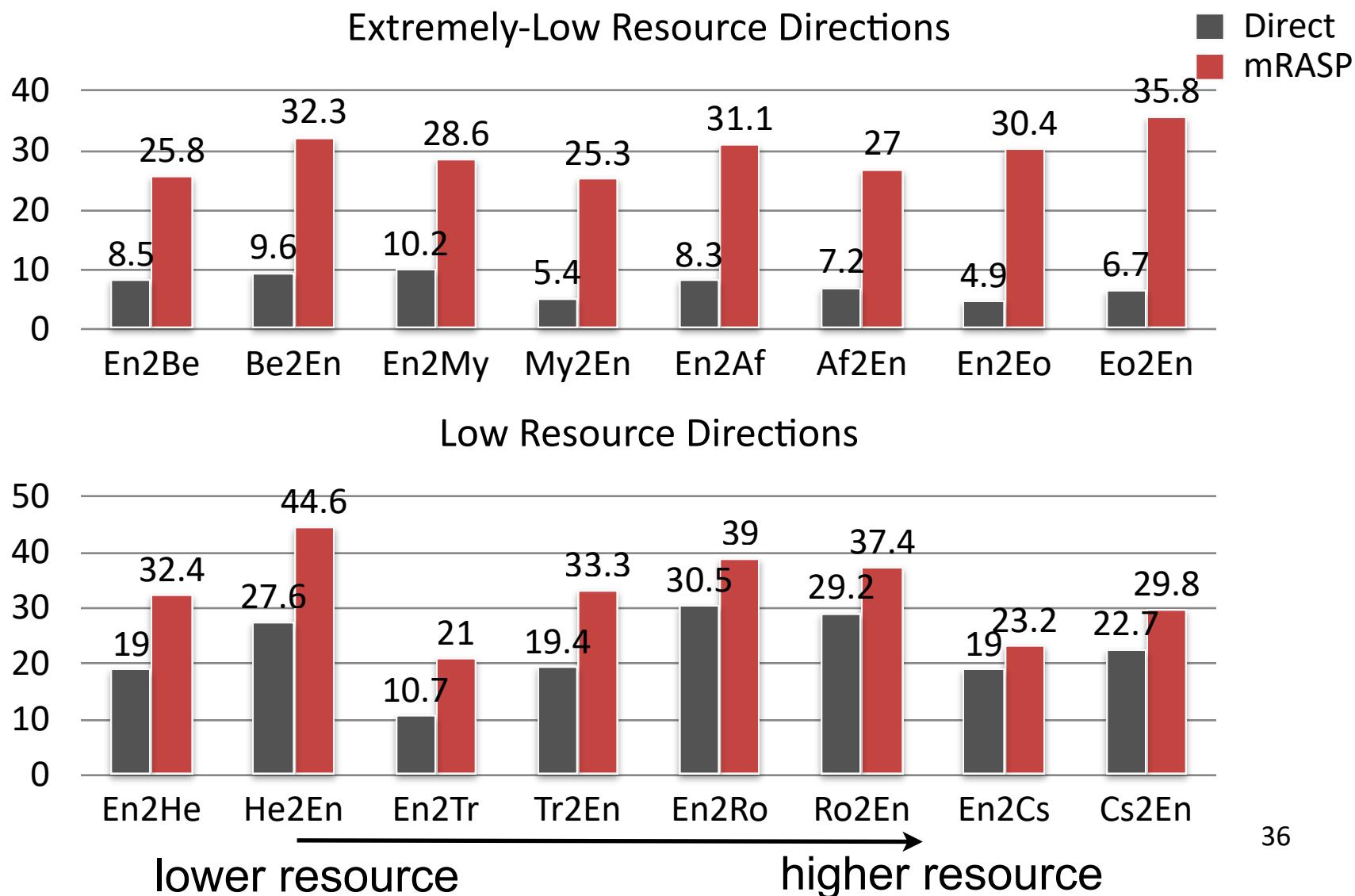
## Decoder

J'adore chanter et danser <EOS>  
J'adore chanter et danser <EOS>

## Random Aligned Substitution

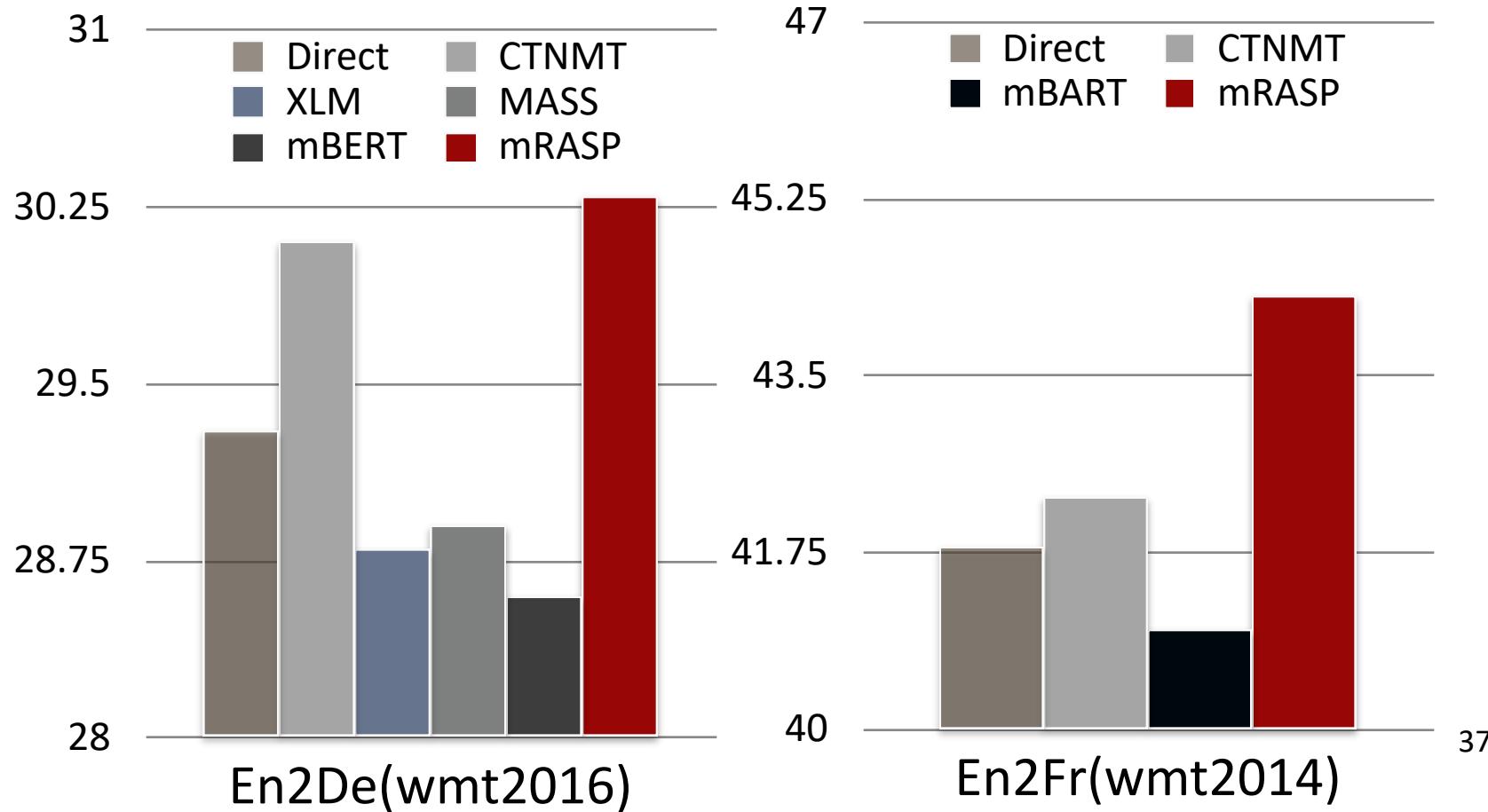


# mRASP gets universal improvement



# mRASP gets universal improvement

- Rich resource benchmarks can be further improved (En->Fr +1.1BLEU).



# VolcTrans

## fanyi.volcengine.cn

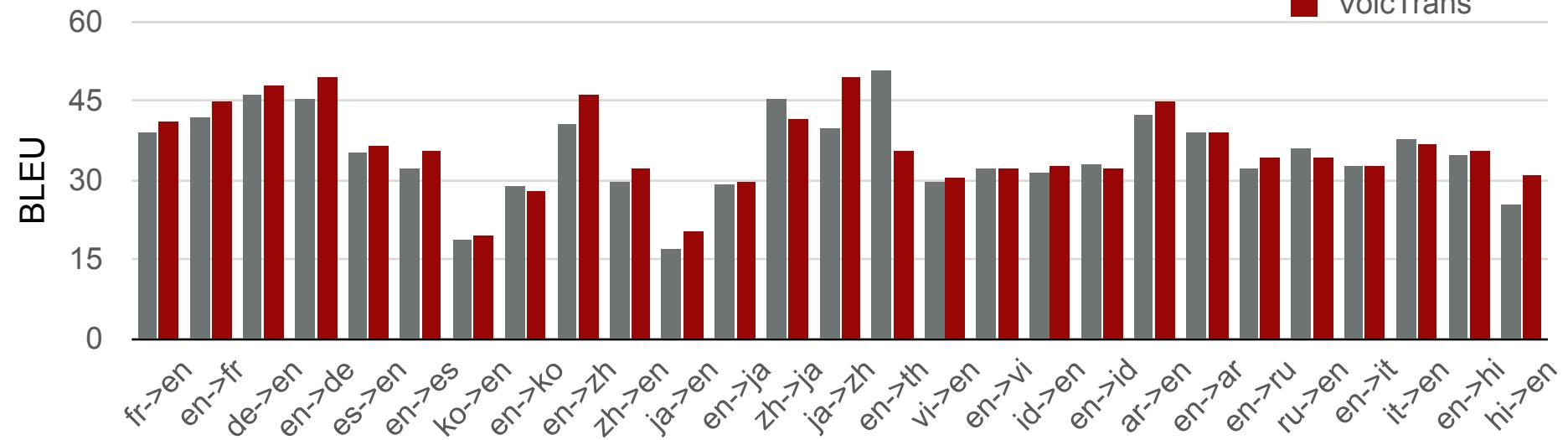
50+  
Clients

9 Billion

16  
languages

Public MT Corpus

3rd-party best  
VolcTrans



# Speech-to-Text Translation

---



Simultaneous Speech-to-text Translation @ VolcTrans

# Multimodal Machine Writing

Xiaomingbot [R. Xu, J. Cao, M. Wang, J. Chen, H. Zhou, Y. Zeng, Y. Wang, L. Chen, X. Yin, X. Zhang, S. Jiang, Y. Wang, **Lei Li**, ACL 2020]

GraspSnooker [Z. Sun, J. Chen, H. Zhou, D. Zhou, **Lei Li**, M. Jiang, IJCAI19b]

Jersey Number Recognition with Semi-Supervised Spatial Transformer Network [G. Li, S. Xu, X. Liu, **Lei Li**, C. Wang, CVPR-CVS18]

# Automatic News Writing in Real-world

---

- Tencent: Dreamwriter, started in 2015.9
- Fast Writer Xiaoxin: Xinhuanet, started in 2015.11
- Xiaomingbot: ByteDance, started in 2016.8
- Xiaonan: Southern Weekend, started 2017.1
- Wibbitz: USA Today
- Heliograf: Washington Post

Landon beat Whitman 34-0;  
<https://t.co/V6zVPi7a9Q>  
[@LandonSports](#) [@koachkuhn](#)  
— WashPost HS Sports  
(@WashPostHS) [September 2, 2017](#)



# Xiaomingbot Automatic News Writing System

Winning 2017 Wu Wen-tsün Award in AI from CAAI



< 足球记者小明 > ...

6621 头条 3 关注 6966 粉丝 1997 赞

私信 已关注

简介：借助人工智能技术，为大家带来快速、全面的足球资讯

AI小记者Xiaomingbot 2018-06-24 14:29:20



北京时间2018年6月23日20时0分，世界杯G组第2轮，比利时迎战突尼斯。最终，比利时5:2战胜突尼斯，卢卡库，巴舒亚伊，阿扎尔为本队建功，哈兹里，布隆为本队挽回颜面。。哈兹里，布隆为本队挽回颜面。



< Xiaomingbot-European > ...

202 Post 4 Following 1.1K Followers

## Post

Thomas Strakosha's 4 saves did not stop Lazio from defeat against Inter Milan, final score 0: 3



Following · Xiaomingbot-European 0

Marseille dropped a 0: 2 decision against PSG in Ligue 1

Following · Xiaomingbot-European 0

Sevilla took away a victory against Huesca, 2: 1



600,000 articles

6 lang

150,000 followers

# Xiaomingbot : Multilingual Robot News Reporter



ByteDance AI Lab  
字节跳动人工智能实验室

## MULTILINGUAL ROBOT NEWS REPORTER

--- Xiaomingbot ---



# Snooker Commentary Generation Combining Visual Understanding with Strategy Prediction



## Balls Detection

Balls' Positions at the Beginning

Red0:	(180, 542)	
Red1:	(189, 552)	
Red2:	(179, 555)	
Red3:	(184, 561)	
Red4:	(202, 563)	
Red5:	(174, 564)	
Red6:	(189, 569)	Red11:(197, 590)
Red7:		Red12:(241, 595)
		Red13:(155, 606)
		Red14:(327, 611)
Brown:	(183, 163)	
Green:	(240, 163)	
Yellow:	(127, 163)	
Blue:	(183, 366)	

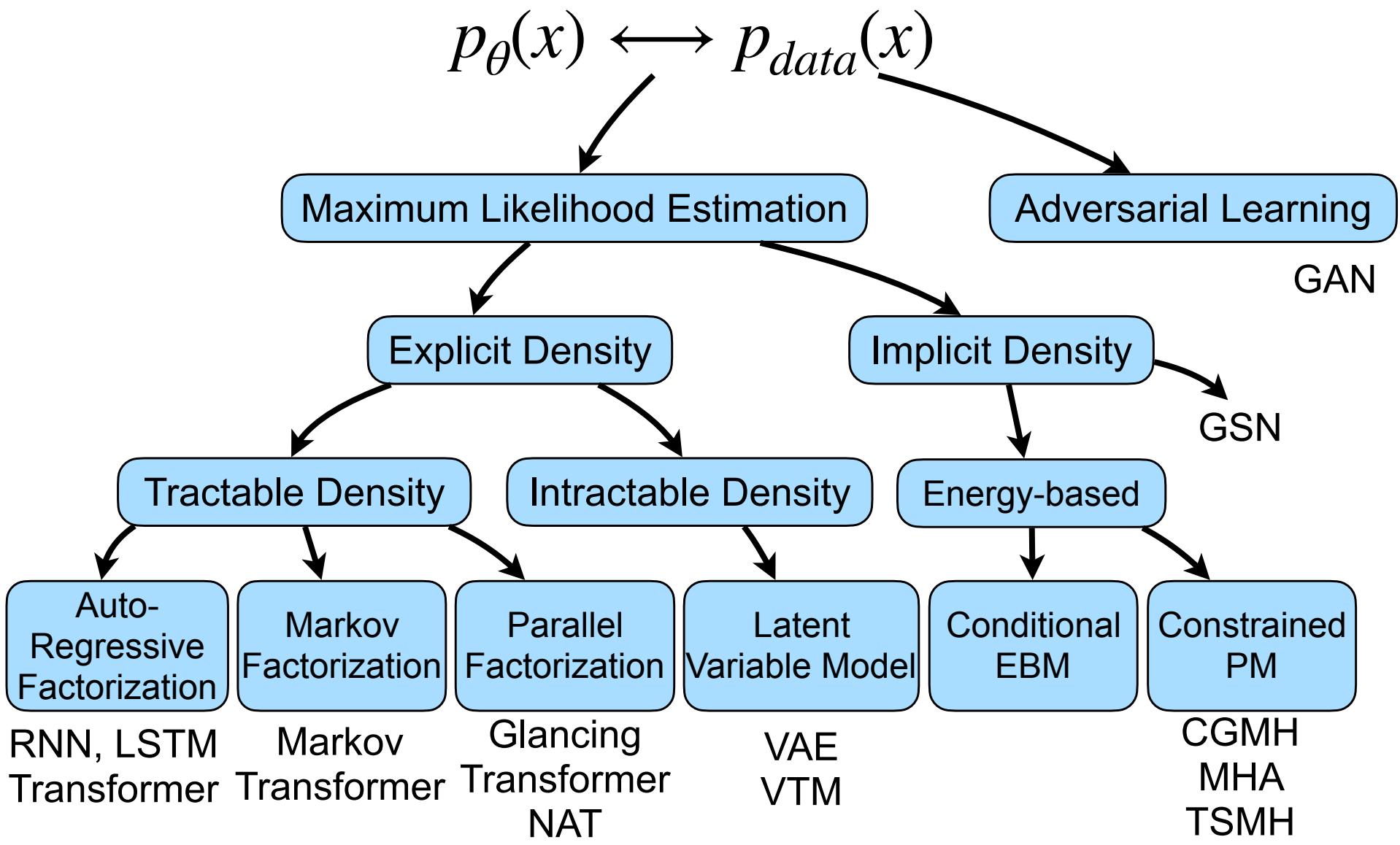
(positions after mapping)

# Summary

---

- Transformer, LSTM & Softmax: Basic neural generation nets for text
- Disentangled Latent Representation
  - VTM: Learning Latent Templates in Variational Space
  - DSS-VAE: Disentangled syntax and semantic representation
- DEM-VAE: Self identifying meaningful clusters with corpus
- MGNMT:
  - integrate four language capabilities together
  - Utilize both parallel and non-parallel corpus
- Multimodal Machine Writing
  - Xiaomingbot system: 600k articles and 150k followers
- Deployed in multiple online platforms and used by over 100 millions of users

# Recap: DGM Taxonomy



# Thanks

---

- Joint w/ Hao Zhou, Rong Ye, Ning Miao, Wenxian Shi, Zaixiang Zheng, Huangzhao Zhang, Ying Zeng, Jiaze Chen, Han Zhang
- Contact: [lileilab@bytedance.com](mailto:lileilab@bytedance.com)



Multilingual MT Pretraining  
<https://github.com/linzehui/mRASP>



火山翻译



A high performance sequence processing lib  
<https://github.com/bytedance/lightseq>

# Reference

---

1. Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, Illia Polosukhin. Attention Is All You Need. NeurIPS 2017.
2. Ning Miao, Hao Zhou, Lili Mou, Rui Yan, Lei Li. "CGMH: Constrained Sentence Generation by Metropolis-Hastings Sampling". In: the 33rd AAAI Conference on Artificial Intelligence (AAAI). Jan. 2019.
3. Huangzhao Zhang, Ning Miao, Hao Zhou, Lei Li. "Generating Fluent Adversarial Examples for Natural Languages". In: the 57th Annual Meeting of the Association for Computational Linguistics (ACL) - short papers. July 2019.
4. Yu Bao, Hao Zhou, Shujian Huang, Lei Li, Lili Mou, Olga Vechtomova, Xinyu Dai, Jiajun Chen. "Generating Sentences from Disentangled Syntactic and Semantic Spaces". In: the 57th Annual Meeting of the Association for Computational Linguistics (ACL). July 2019.
5. Ning Miao, Hao Zhou, Chengqi Zhao, Wenxian Shi, Lei Li. "Kernelized Bayesian Softmax for Text Generation". In: the 33rd Conference on Neural Information Processing Systems (NeurIPS). Dec. 2019.
6. Zaixiang Zheng, Hao Zhou, Shujian Huang, Lei Li, Xinyu Dai, Jiajun Chen. "Mirror Generative Models for Neural Machine Translation". In: International Conference on Learning Representations (ICLR). Apr. 2020.
7. Rong Ye, Wenxian Shi, Hao Zhou, Zhongyu Wei, Lei Li. "Variational Template Machine for Data- to-Text Generation". In: International Conference on Learning Representations (ICLR). Apr. 2020.
8. Ning Miao, Yuxuan Song, Hao Zhou, Lei Li. "Do you have the right scissors? Tailoring Pre-trained Language Models via Monte-Carlo Methods". In: the 58th Annual Meeting of the Association for Computational Linguistics (ACL) - short papers. July 2020.
9. Wenxian Shi, Hao Zhou, Ning Miao, Lei Li. "Dispersing Exponential Family Mixture VAEs for Interpretable Text Generation". In: Proceedings of the 37th International Conference on Machine learning (ICML). July 2020.
10. Maosen Zhang, Nan Jiang, Lei Li, Yexiang Xue. "Constraint Satisfaction Driven Natural Language Generation: A Tree Search Embedded MCMC Approach". In: the Conference on Empirical Methods in Natural Language Processing (EMNLP) - Findings. Nov. 2020.
11. Zehui Lin, Xiao Pan, Mingxuan Wang, Xipeng Qiu, Jiangtao Feng, Hao Zhou, Lei Li. Pre-training Multilingual Neural Machine Translation by Leveraging Alignment Information. EMNLP 2020.