



Association for
Computational
Linguistics

The Bright Future of ACL/NLP

Dr. Ming Zhou, ACL president
Microsoft Research Asia

ACL 2019

Florence, Italy, July 29, 2019



A big thanks to

- General Chair Lluís Màrquez, program chairs Anna Korhonen and David Traum, local organization chairs *Alessandro Lenci, Bernardo Magnini, Simonetta Montemagni, other chairs, and everyone of their teams*
- ACL 2019 Coordinating Committee(Marti Hearst, David Yarowsky, Priscilla Rasmussen and all others)



Galileo showed the Doge of Venice how to use the telescope



Outline »



ACL business update



NLP technical development



Galileo showed the Doge of Venice how to use the telescope



Outline »



ACL business update



NLP technical development

ACL: the premier scientific and professional society for CL/NLP



Annual conference
ACL



Conferences by
regional chapters



21 SIGs
including SIGDAT,
SIGHAN, SIGMT,
SIGDIAL...



ACL Anthology
<https://www.aclweb.org/anthology/>

1962
ACL

1982
EACL

2000
NAACL

2018
AACL

The discipline of CL/NLP has been formed with solid theoretical framework, systematic technologies and important applications. Now CL/NLP has been widely viewed as the holy grail of AI.

ACL executive board's duties



Handle finance and membership



Select and negotiate venues for the main conferences



Help organize the various components of conferences



Coordinate 5 main conferences (ACL, EMNLP, EACL, NAACL, AACL)



Coordinate 21 SIGs and 50 workshops



Sponsor 2 journals (CL and TACL) and ACL anthology



Manage centralized IT



Make and execute policy that matters



Handle rising problems



Make strategic planning

Current ACL executive board members



Ming Zhou
President



Shiqi Zhao
Secretary (2016 - 2020)



Sharon Goldwater
EACL chair



Marti Hearst
Past President



Nitin Madnani
At-large (2019 - 2021)



Julia Hockenmaier
NAACL chair (2018 - 2019)



Hinrich Schütze
Vice-President



Barbara Di Eugenio
At-large (2018 - 2020)



Haifeng Wang
AAACL Chair (2018 - 2020)



Rada Mihalcea
Vice-President-Elect (2019)



Jennifer Foster
At-large (2017 - 2019)



David Yarowsky
Treasurer (2018 - 2022)



Hwee Tou Ng
CL Journal editor (2018-)



Priscilla Rasmussen
Business Manager

Supported by business manager:

[https://www.aclweb.org/adminwiki/index.php?title=ACL Officers](https://www.aclweb.org/adminwiki/index.php?title=ACL_Officers)

Thanks to outgoing ACL execs who finished their terms



Joakim Nivre
Past President



Graeme Hirst
Past Treasurer



Jing-Shin Chang
At-large



Paola Merlo
CL Editor



Walter Daelemans
Chair of the EACL

ACL fellows in 2018



Robert Dale

For significant contributions to research in the generation of referring expressions and in natural language generation more broadly.



Jason Eisner

For significant contributions to probabilistic models and algorithms for finding linguistic structure, especially lexicalized syntax and morphology.



Mari Ostendorf

For significant contributions to prosody, pronunciation, acoustic, language modeling, and developments in using out-of-domain data and discourse structure.



Dragomir Radev

For significant contributions to text summarization and question answering, as well as large scale efforts to expand and diversify the computational linguistics pipeline.



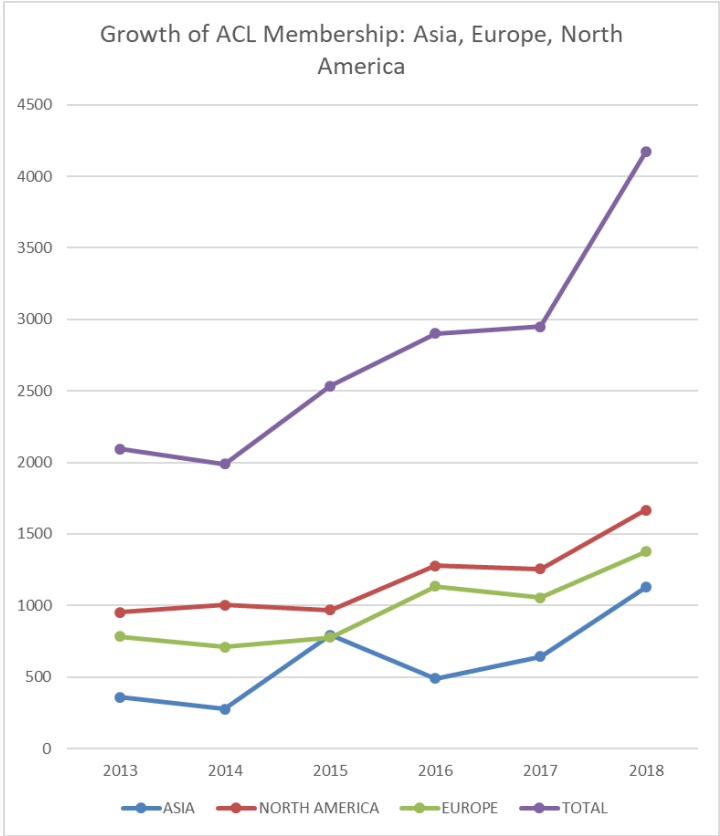
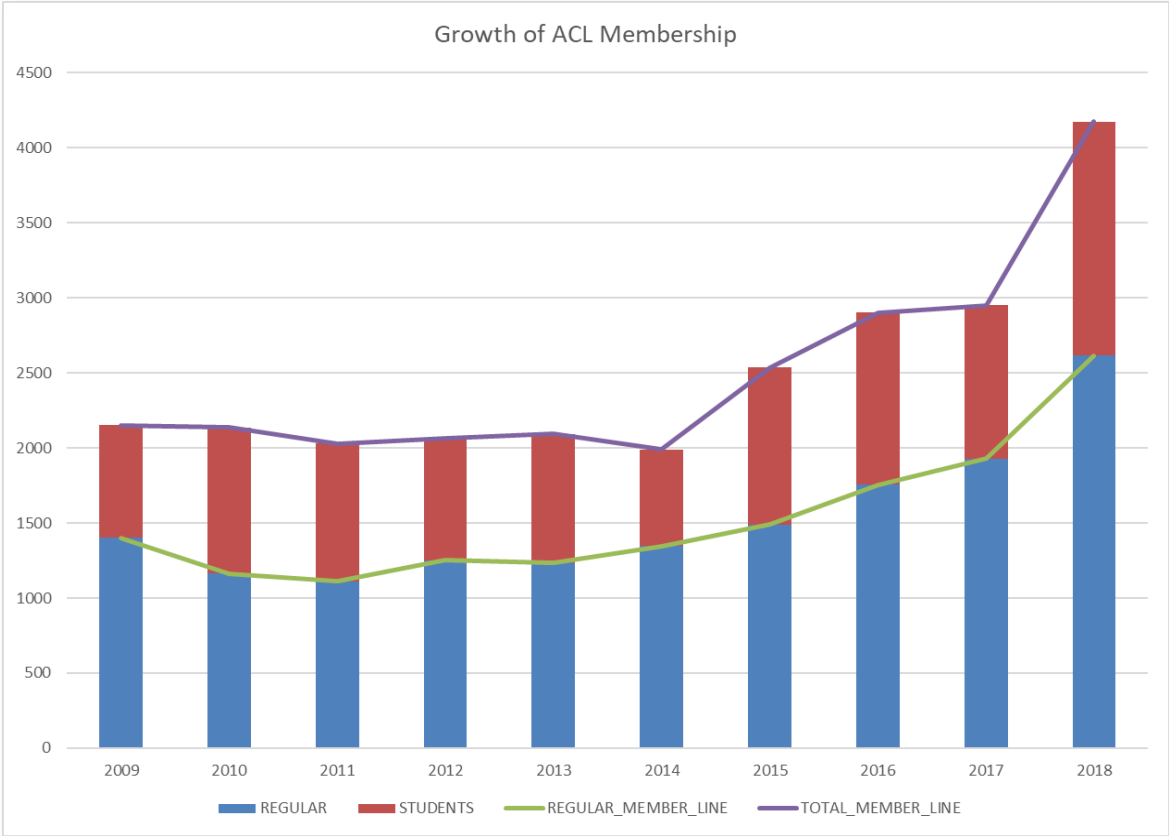
Ellen Riloff

For significant contributions to information extraction, and the analysis of sentiment, subjectivity and affect.

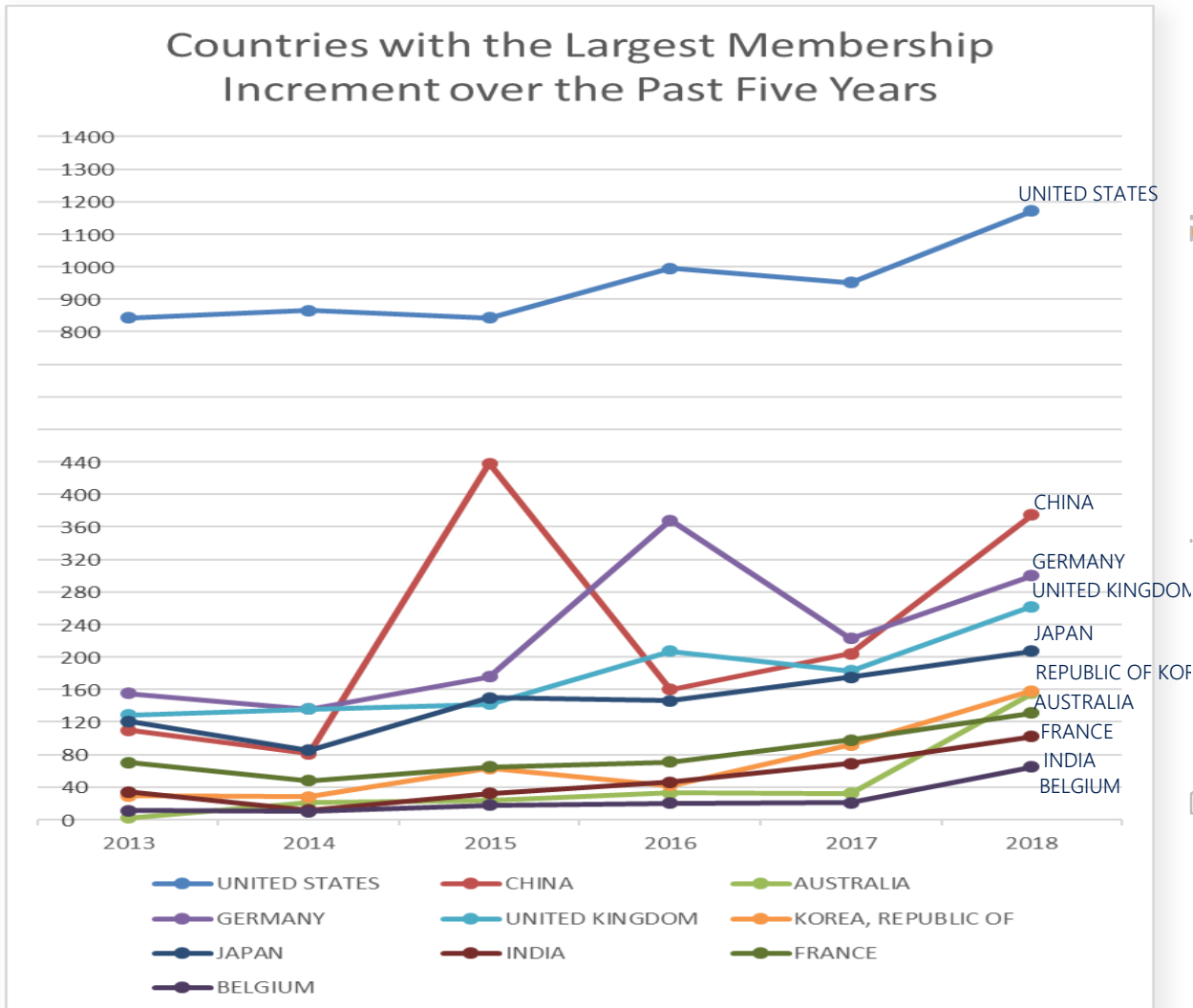
The ACL Fellows program has been established in 2011 by the ACL. The Fellows program recognizes ACL members whose contributions to the field have been most extraordinary in terms of scientific and technical excellence, service to the association and the community and/or educational or outreach activities with broader impact.

https://aclweb.org/aclwiki/ACL_Fellows

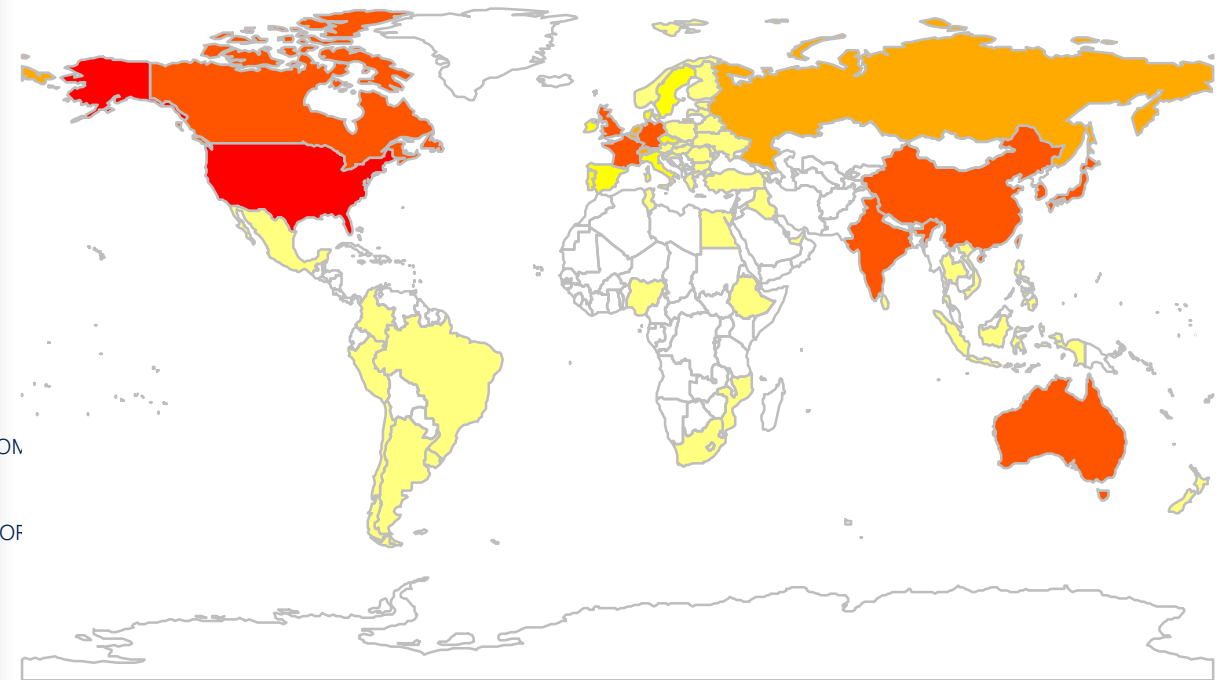
Smooth growth of ACL membership



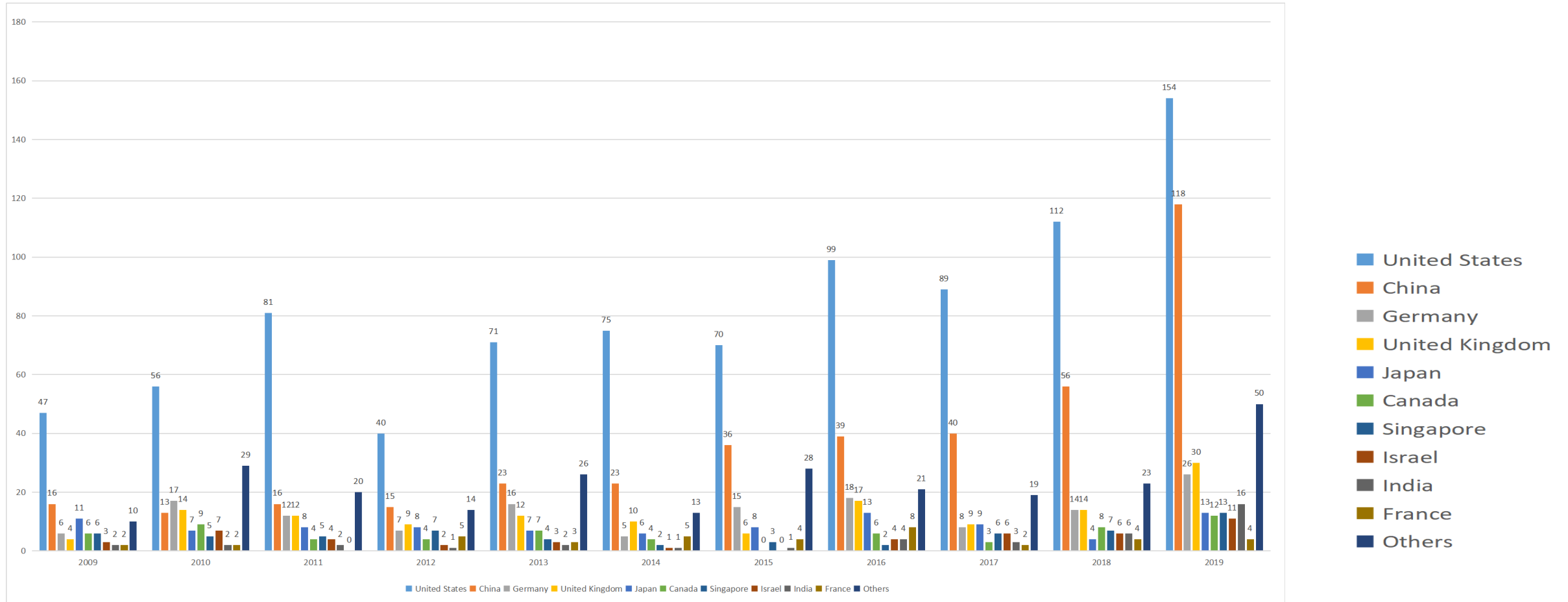
Distribution of membership (statistics in 2013-2018)



ACL membership density



Imbalance of accepted papers over countries/regions



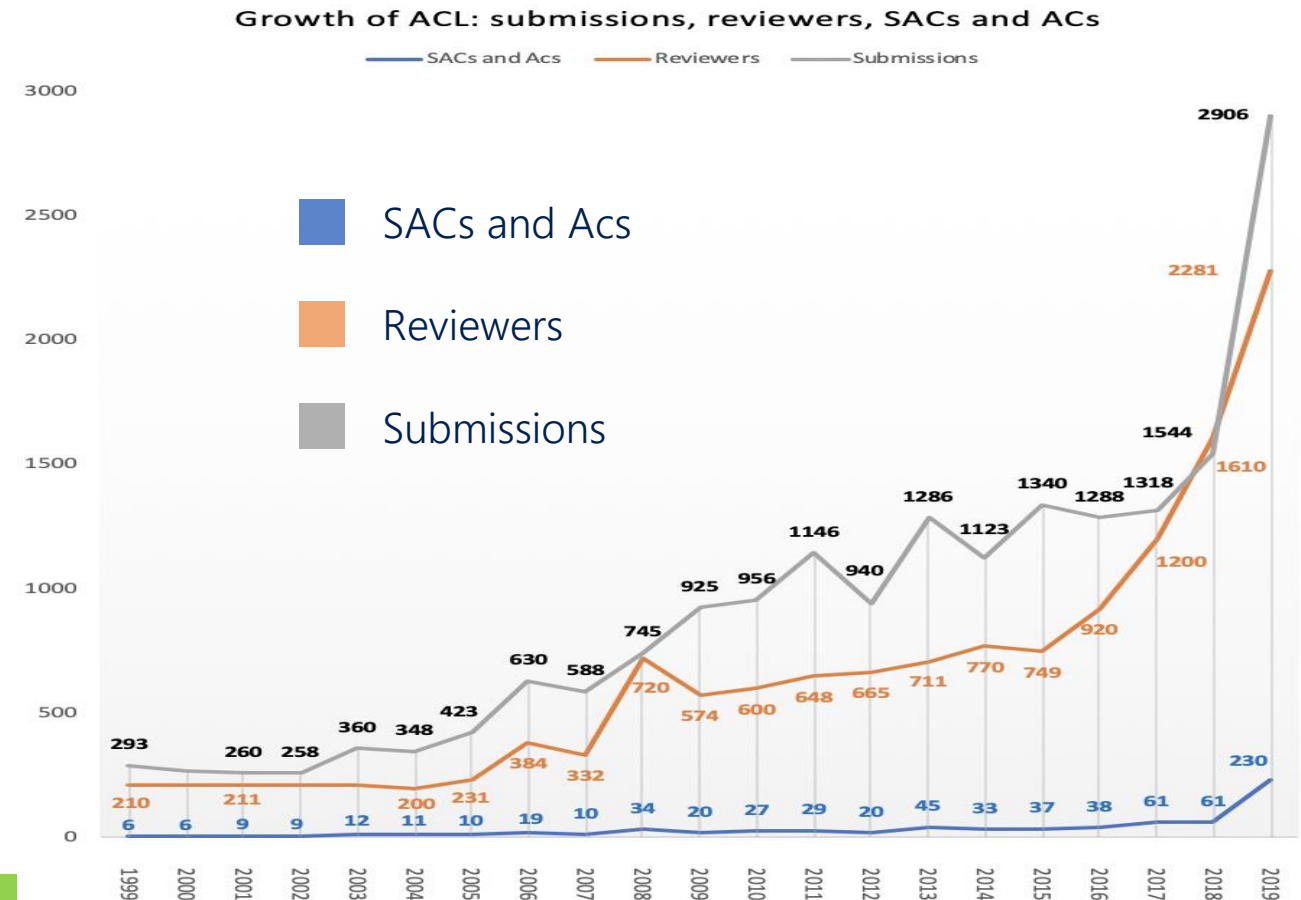
Based on ACL full papers (2009-2019)

Soaring growth of submission causes huge challenges to paper review

|papers|

$\propto$ |conferences| $\times$ |topics| $\times$ |datasets|
 $\times$ |network structures| $\times$ |learning algorithms|
 $\times$ |languages|
 $\times$ |researchers + professors + students|
 $\times$ |rejection times|
 $\times$ |high salary of AI jobs| $\times$ | ... |

There will be a special discussion on paper reviewing at business meeting on July 30



<http://acl2019pcblog.fileli.unipi.it/?p=156>



ACL/NLP growth in Asia-Pacific

Diverse societies in Asia-Pacific



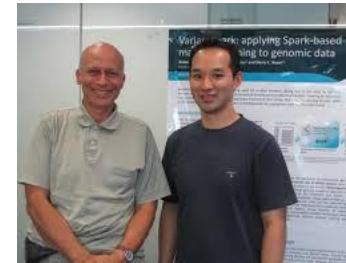
- Australasian Language Technology Association (ALTA), Sydney, Australia
- Natural Language Processing Association India (NLP AI), Hyderabad, India
- Indonesian Association of Computational Linguistics (INACL), Jakarta, Indonesia
- The Association for Natural Language Processing (ANLP), Tokyo, Japan
- SIG-HLT (Special Interest Group of Human Language Technology) of KIISE (Korea Institute of Information Scientists and Engineers), Pohang, Korea
- Chinese and Oriental Languages Information Processing Society (COLIPS), Singapore
- The Association for Computational Linguistics and Chinese Language Processing (ACLCLP), Chinese Taipei
- Society for Natural Language Processing (SNLP), Lahore, Pakistan
- Chinese Information Processing Society
- China Computer Federation
- China Association of Artificial Intelligence



Annual China-Japan joint NLP workshop (since 2001)



Conferences organized by NLP societies in Asia-Pacific



IALP 2018 International Conference on Asian Language Processing



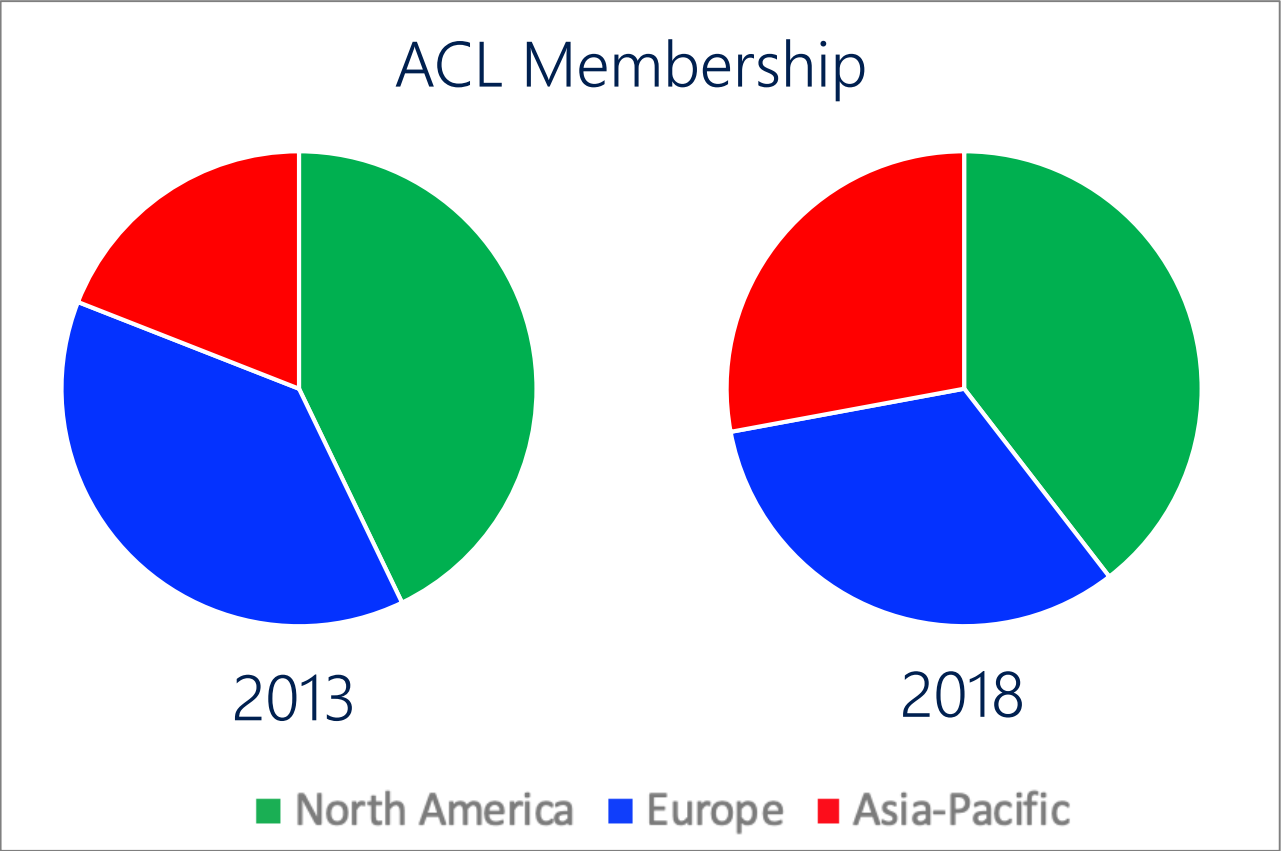
November 15-17, 2018 – Bandung, Indonesia

Conferences, summer schools, tutorials in China



Fast growth of ACL membership in Asia-Pacific

- The impact of ACL has been dramatically increasing in Asia-Pacific in recent years



Year	Host
2018	Melbourne
2015	Beijing
2012	Jeju Island
2009	Singapore
2006	Sydney
2003	Sapporo
2000	Hong Kong

ACLs held in Asia-Pacific

AACL: the Asia-Pacific Chapter of ACL (launched in 2018)

- Serves ACL members from 57 countries/regions in Asia-Pacific
- Builds a new bridge with AFNLP and all NLP societies in Asia-Pacific



AACL Executive Board



Haifeng Wang
Chair



Keh-Yih Su
Chair-elect



Yang Liu
Secretary



Seung-won Hwang
Treasurer



Yusuke Miyao
At-large



Jian Su
At-large



Mark Dras
At-large

Towards balanced, inclusive and diverse development of ACL/NLP



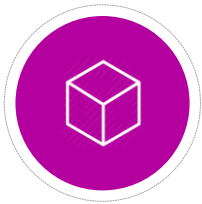
Better membership service
by ACL and its chapters



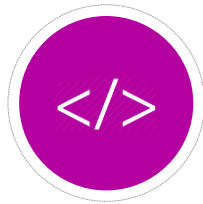
Talent fostering by summer schools,
mentoring programs, internship
programs, language training



Conferences and
activities in diverse venues



Strong support to low-
resource languages



WiNLP/EquiCL/BIG to encourage
diversity and inclusion



IT system, review system, coordination
across chapters, SIGs and conferences

New committees



Information Committee Director
Nitin Madnani



Equity Committee Director
Under search led by
Rada Mihalcea



Publicity Committee Director
Barbara Plank



Professional Conduct Committee Directors
Emily M Bender



Graeme Hirst



Anthology Committee Director
Matt Post



Business Meeting

Plenary Hall

July 30, 2019

- Reports from ACL functional units (secretary, treasurer, office, IT, CL, TACL)
- Updates on EACL, NAACL and AACL
- Progress on setting up ACL2020, ACL2021
- Special panel on paper reviewing

Welcome





Galileo showed the Doge of Venice how to use the telescope



Outline »

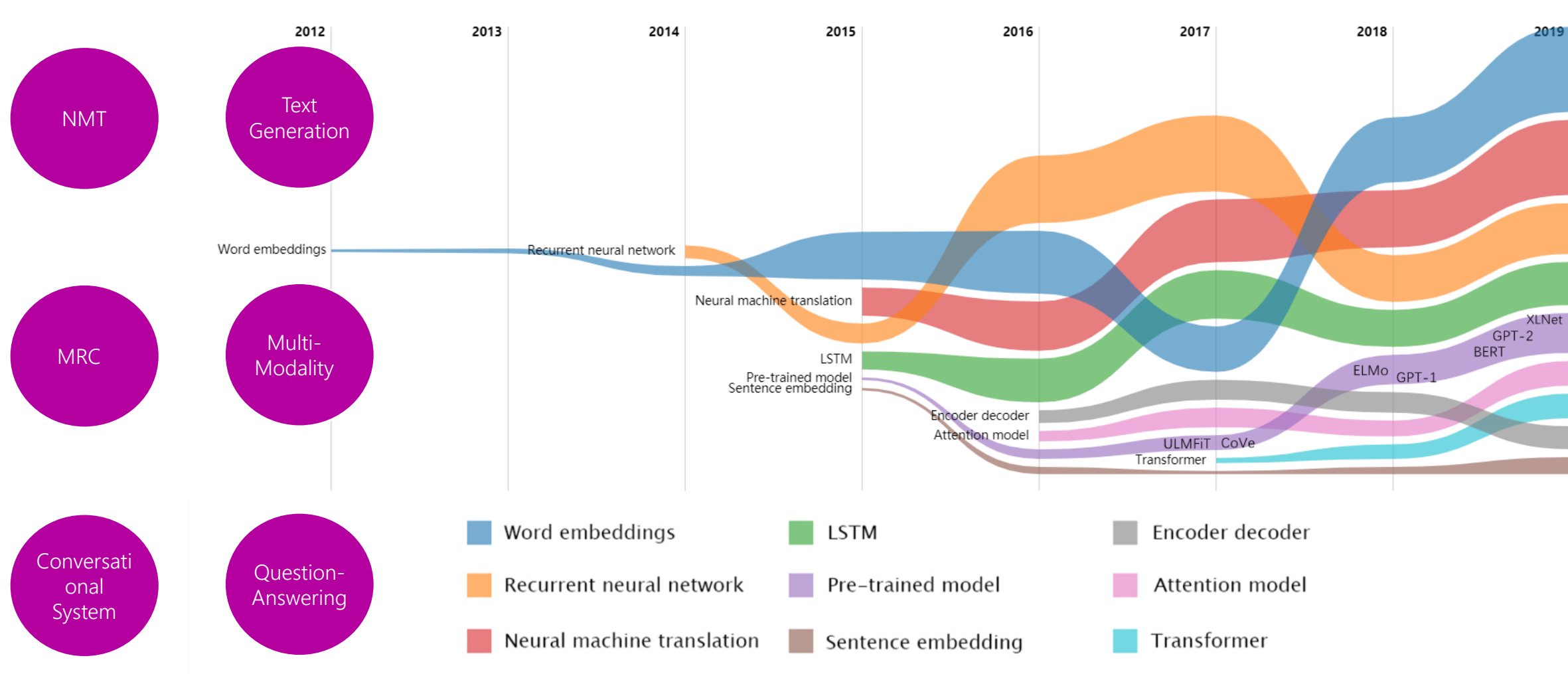


ACL business update



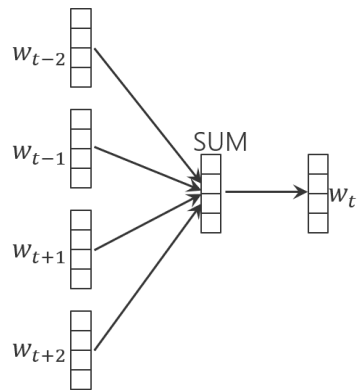
NLP technical development

DNN-NLP progress

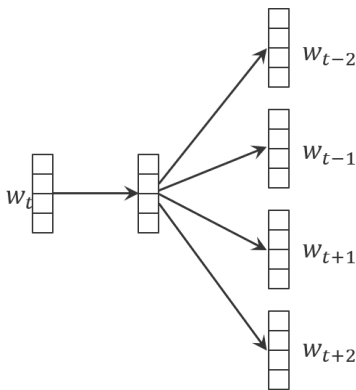


This figure was credited to AMiner, Tsinghua University, 2019

Key techs of DNN-NLP

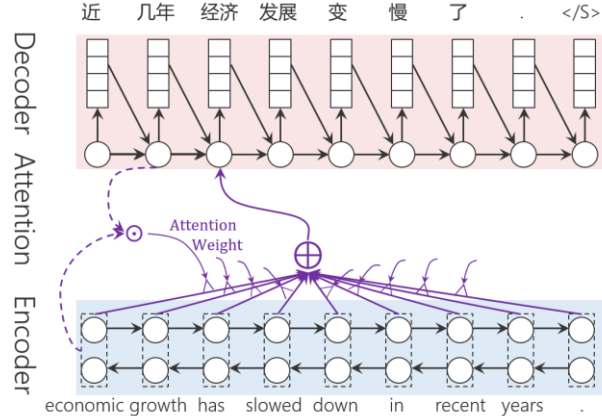
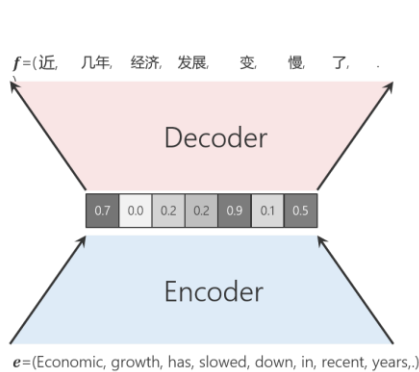


CBOW (Continuous Bag-of-Words): using the context words in a window to predict the central word.

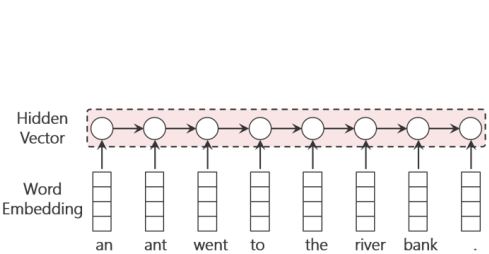


Skip-gram (Continuous Skip-gram): using the central word to predict the context words in a window.

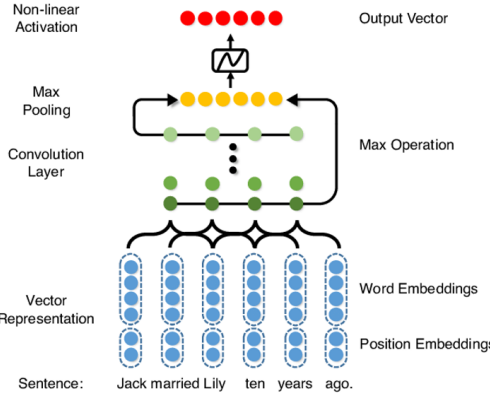
Word embedding (Mikolov et al., 2013)



Encoder-Decoder with attention (Bahdanau et al., 2014)

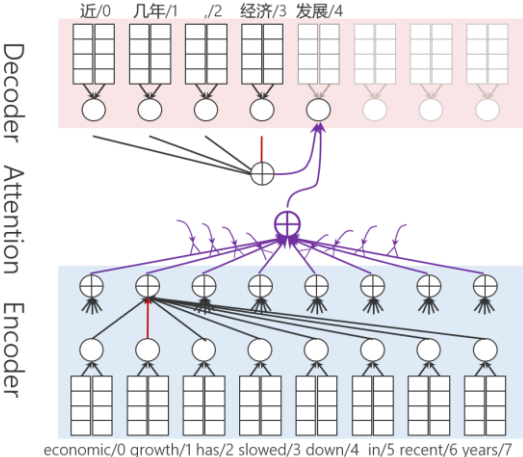


Recurrent Neural Networks (RNN)



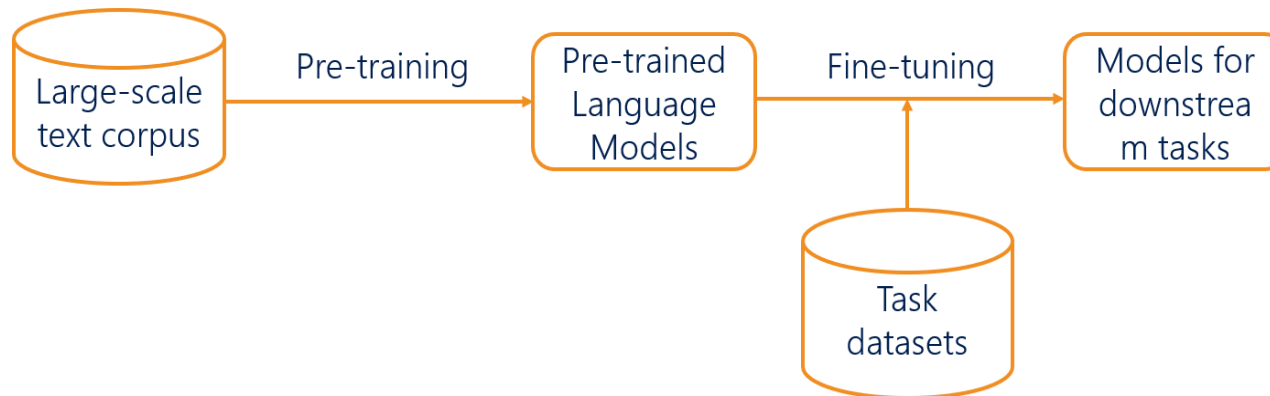
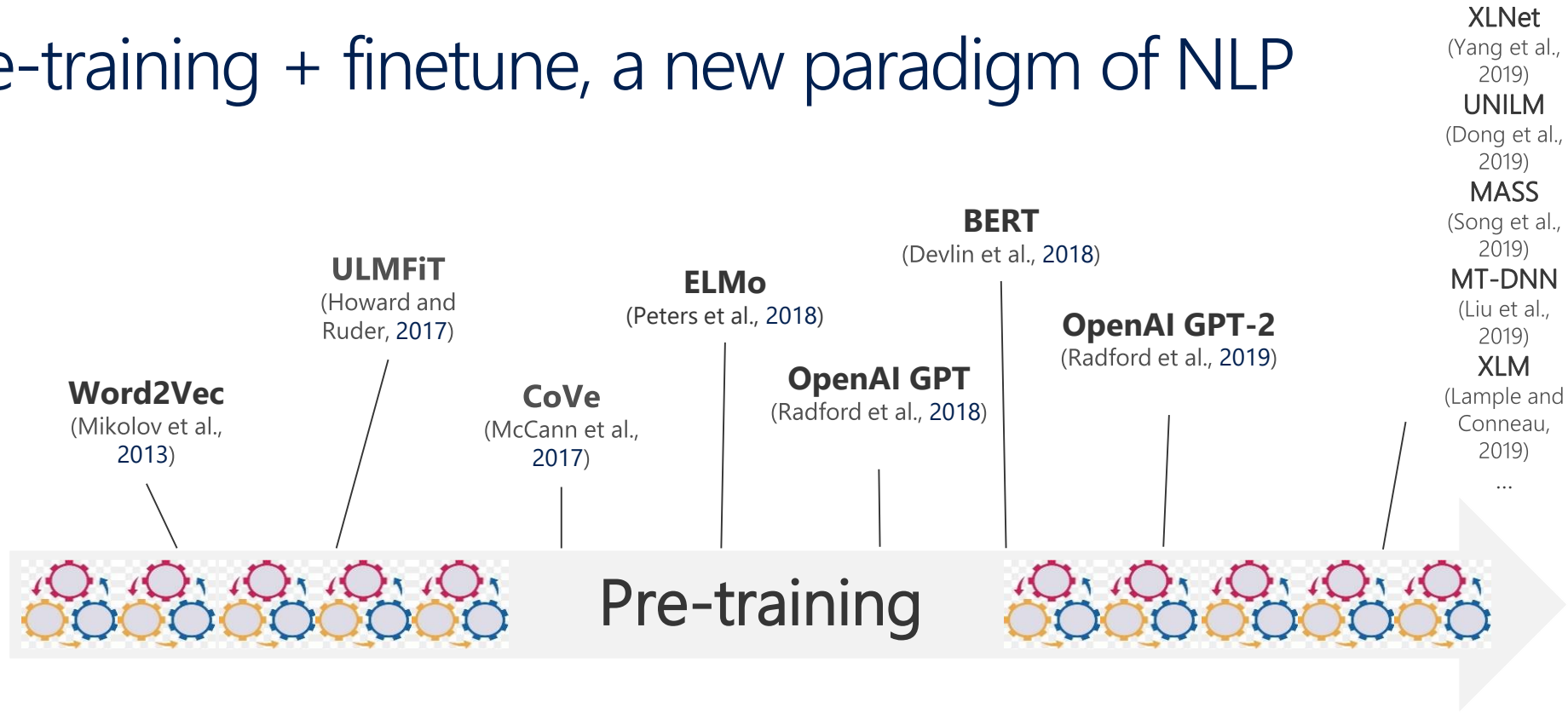
Convolutional Neural Networks (CNN)

Sentence Embedding



Transformer (Vaswani et al., 2016)

Pre-training + finetune, a new paradigm of NLP



NLP Tasks

Machine Translation

Search Engine

Semantic Parsing

Question Answering

Chatbot & Dialogue

Paraphrase
Classification

Text Entailment

Sentiment Analysis

...

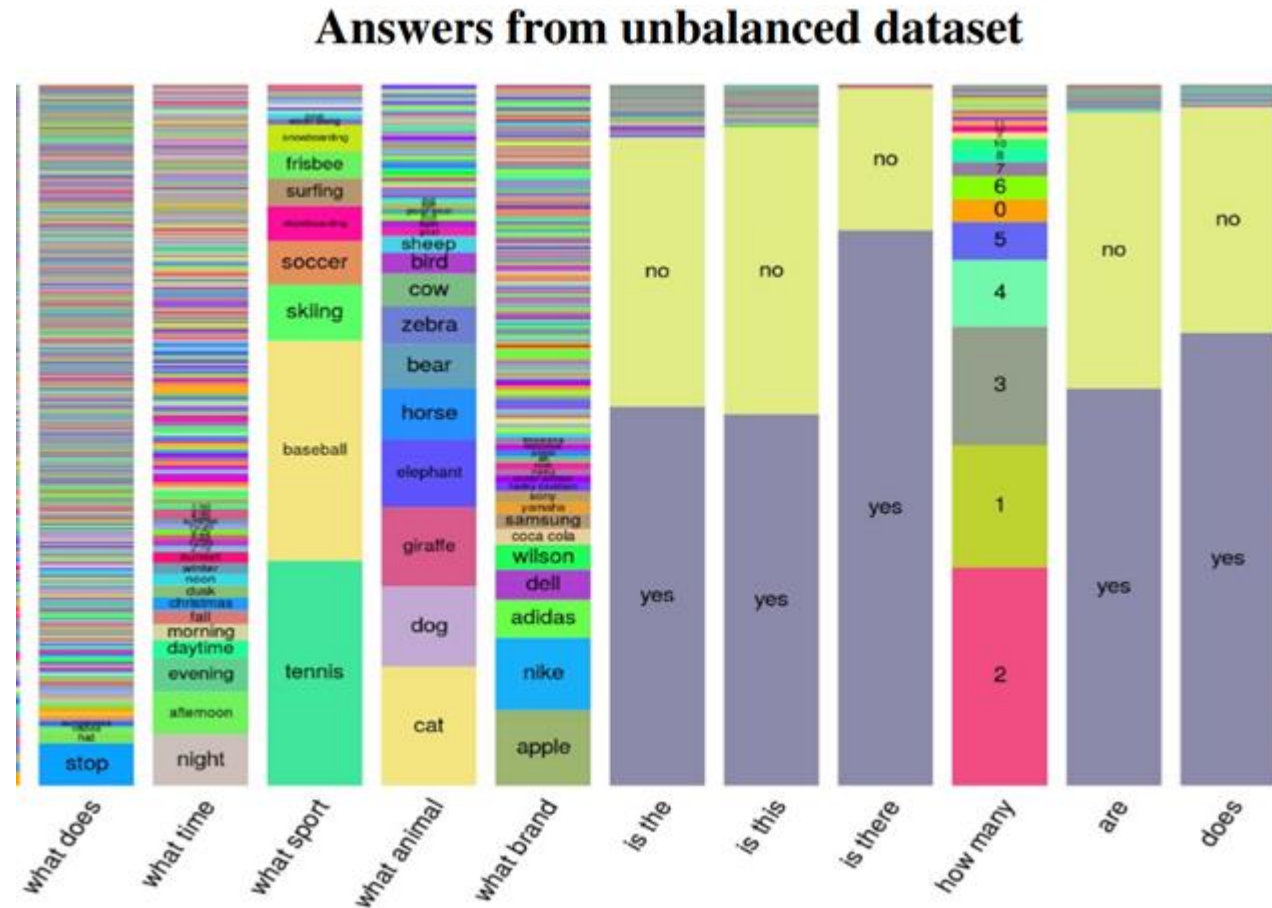
Where is the future direction of NLP?

- Are we satisfied with current DNN-NLP?
- DNN-NLP deeply relies on huge cost of computer power and annotated data and suffers from big challenges in modelling, reasoning and interpretability.
- Linguistics, knowledge, common sense and symbolic reasoning should still play important roles to solve these challenges.
- I would like to analyze challenges in typical tasks and share my views on the technical developments.

Dataset: high cost, bias, noises, privacy and discrepancy from real scenarios

Racial Analogies	
black → criminal	caucasian → lawful
caucasian → hillbilly	asian → yuppie
asian → engineer	black → killer
Religious Analogies	
christian → conservative	jew → liberal
muslim → terrorist	jew → journalist
christian → conservative	muslim → regressive

Table 1: Examples of racial and religious biases in analogies generated from word embeddings trained on the Reddit data from users from the USA.

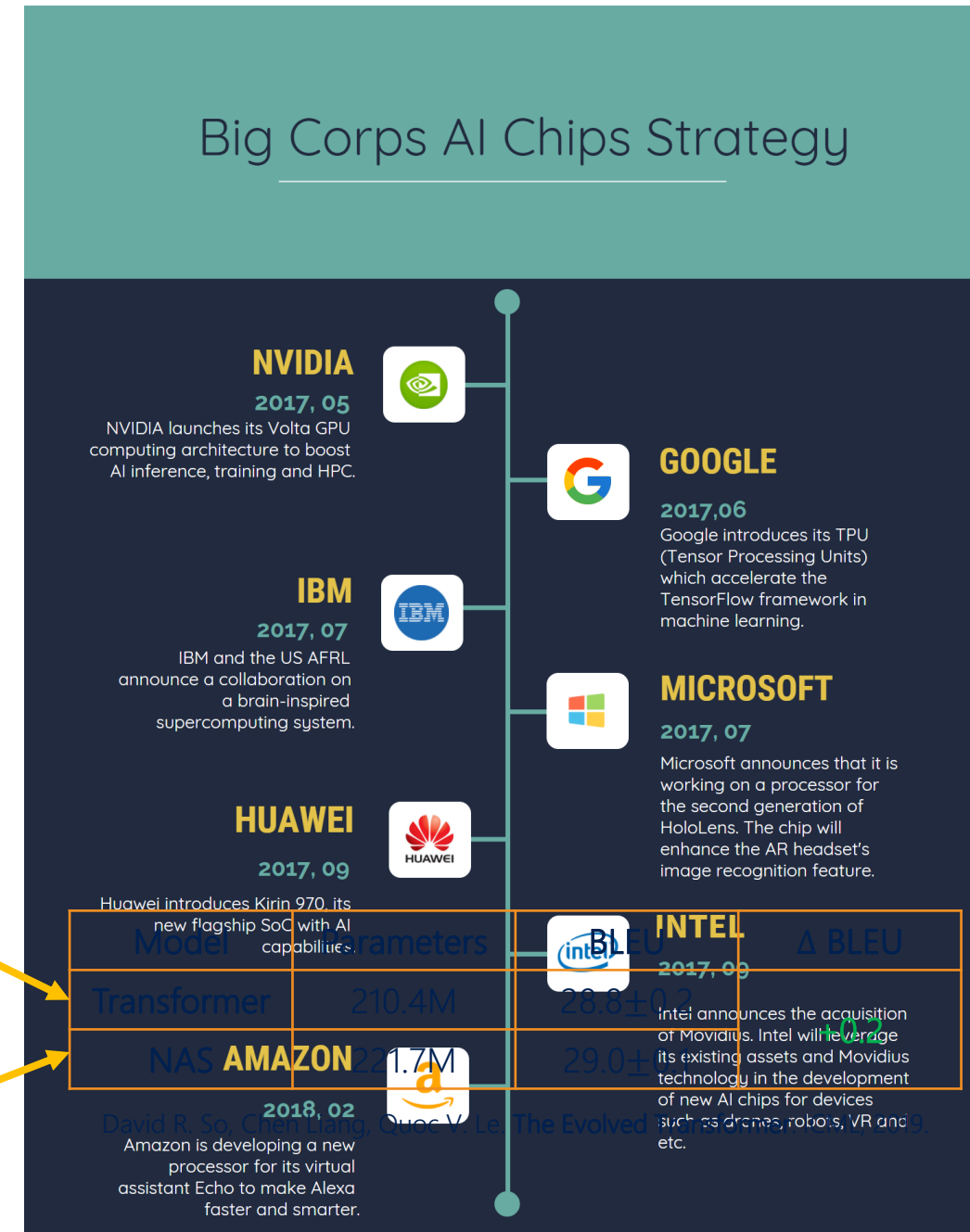


Fierce computing power arm races

- New arms race over AI chips and computing power
- High computation cost for big models
- Low return on investment(sometimes)
- New type of environmental pollution

Model	Hardware	Power (W)	Hours	kWh·PUE	CO ₂ e	Cloud compute cost
Transformer _{base}	P100x8	1415.78	12	27	26	\$41–\$140
Transformer _{big}	P100x8	1515.43	84	201	192	\$289–\$981
ELMo	P100x3	517.66	336	275	262	\$433–\$1472
BERT _{base}	V100x64	12,041.51	79	1507	1438	\$3751–\$12,571
BERT _{base}	TPUv2x16	—	96	—	—	\$2074–\$6912
NAS	P100x8	1515.43	274,120	656,347	626,155	\$942,973–\$3,201,722
NAS	TPUv2x1	—	32,623	—	—	\$44,055–\$146,848
GPT-2	TPUv3x32	—	168	—	—	\$12,902–\$43,008

Emma Strubell, Ananya Ganesh and Andrew McCallum. Energy and Policy Considerations for Deep Learning in NLP. ACL, 2019.



Analysis on typical tasks



Rich-Resource



Low-Resource



Multi-Turn

Analysis on typical tasks



Rich-Resource



Low-Resource



Multi-Turn

Error analysis of NMT results (Ch-En)

Error Category	Fraction [%]
Incorrect Words	7.64
Ungrammatical	6.33
Missing Words	5.46
Named Entity	4.38
Person	1.53
Location	1.53
Organization	0.66
Event	0.22
Other	0.44
Word Order	0.87
Factoid	0.66
Word Repetition	0.22
Collocation	0.22
Unknown Words	0

Error distribution, as fraction of sentences that contain specific error categories.

← The translation of 500 sentence were manually checked and error types were labeled

← This analysis indicate that there is still big room to improve the translation quality

Achieving Human Parity on Automatic Chinese to English News Translation

<https://arxiv.org/abs/1803.05567>

Missing words

英国最高法院随后裁定引渡他到瑞典。但阿桑奇于2012年6月在保释期间进入位于**伦敦的厄瓜多尔驻英国使馆**寻求庇护至今。阿桑奇创建的“维基揭秘”网站因公布大量美国有关阿富汗和伊拉克战争的秘密文件，引起轰动和争议。

The British Supreme Court subsequently decided to extradite him to Sweden. But Assange entered **the British embassy in London** in June 2012 on bail to seek refuge so far. The WikiLeaks website, created by Mr. Assange, has caused a stir and controversy by publishing a large number of U.S. secret documents about the war in Afghanistan and Iraq.



“厄瓜多尔”
(Ecuador) is
missed

Acronym understanding

德国在参与打击极端组织的多国联合行动时，向土耳其空军基地派驻约250名军人。土耳其政府此前指责德国为参与去年7月土耳其未遂政变的人员提供政治避难。作为报复，**土方**禁止德国议员探视德国驻军。

Germany has deployed about 250 troops to the Turkish Air Force base in its multinational operations against extremist groups. The Turkish government has previously accused Germany of providing political asylum for those who participated in last July's attempted coup in Turkey. In retaliation, **the Earth** forbids German MPs to visit the German garrison.



“土方” (contextualized acronym of 土耳其 /Turkey) is wrongly translated

Unknown named entities

日前闭幕的“**一带一路**”国际合作高峰论坛收获了丰硕成果，达成多个合作项目，提出了一系列合作举措，赢得广泛赞誉。

The recent closing of the "area all the way" International Cooperation Summit Forum Harvest fruitful results, reached a number of cooperation projects, put forward a series of cooperation initiatives, won wide acclaim.

国务院总理李克强19日在中南海**紫光阁**会见菲律宾**众议长**阿尔瓦雷兹。

Chinese Premier Li Keqiang met with Philippine **Chancellor** Alvarez in Zhongnanhai 19th.

“一带一路” (one belt one road) is wrongly translated

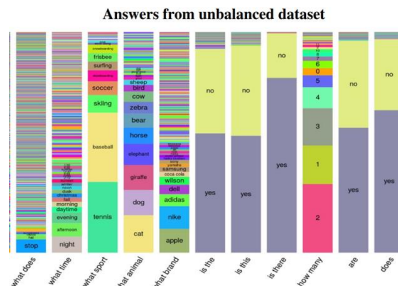
“众议长” (house speaker) is wrongly translated perhaps due to noises of training data

“紫光阁” (Ziguangge) is wrongly missed

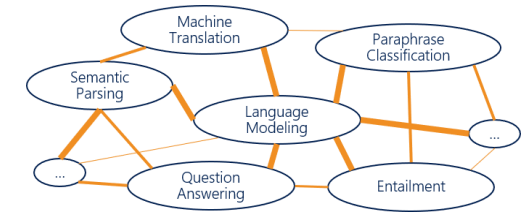
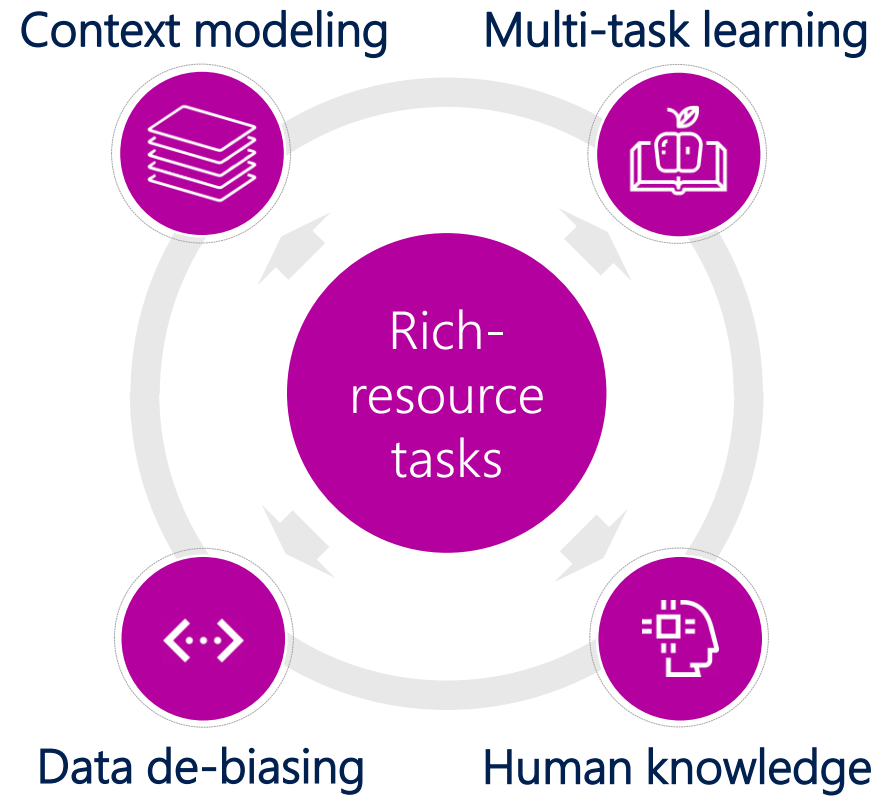
Important topics for rich-resource tasks



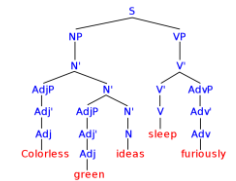
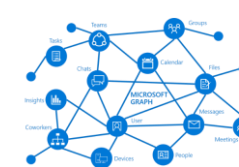
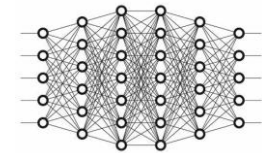
Model longer context for document MT, cross-document summarization dialogue system and chatbot.



Alleviate bias issues of training and evaluation datasets for robust models.



Further strengthen models with multi-task learning



Leverage linguistic knowledge and domain knowledge in modelling

Analysis on typical tasks



Rich-Resource



Low-Resource



Multi-Turn

Low-resource scenarios



A task with little training data but highly related to other rich resource tasks

Transfer Learning: learn from other tasks



A task with little training data in one language but with rich training data in other languages

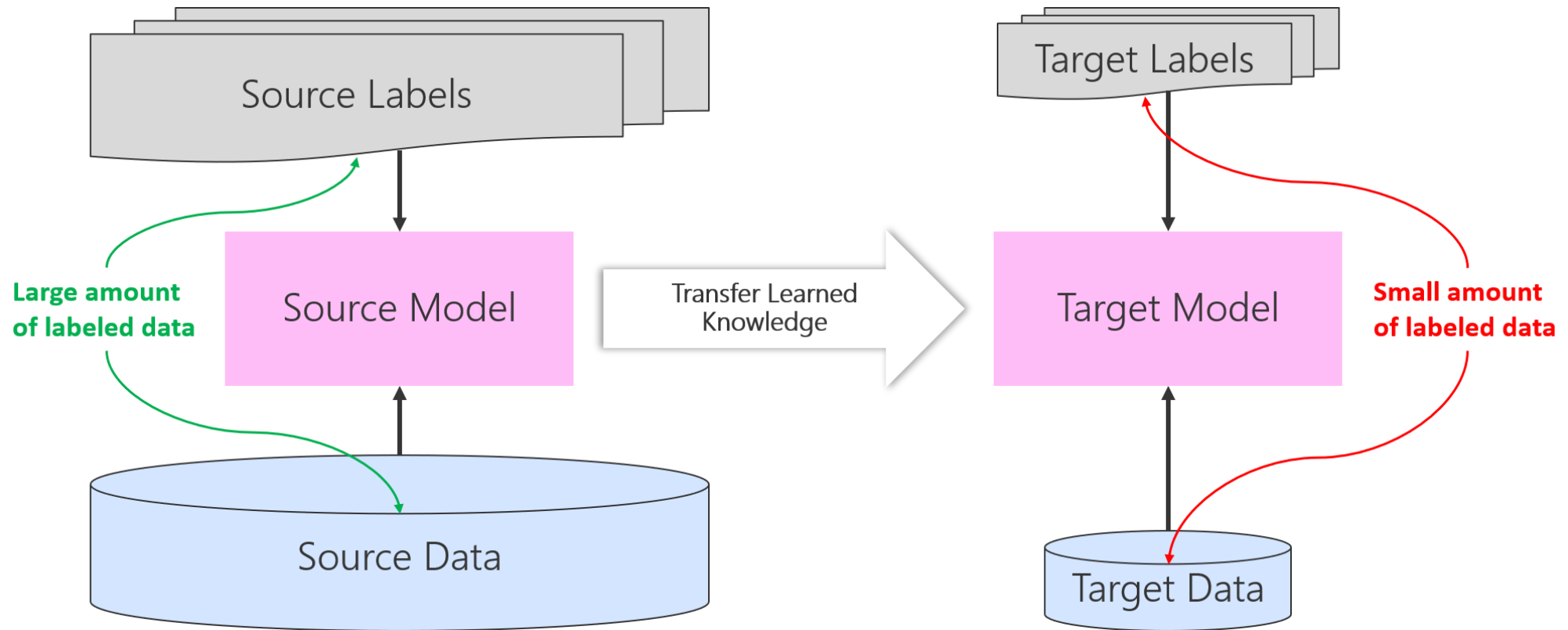
Cross-lingual Learning: learn from other languages



A task with little training data, without related tasks, without rich training data in other languages

Less or unsupervised Learning: learn from seeds/dictionaries/rules/...

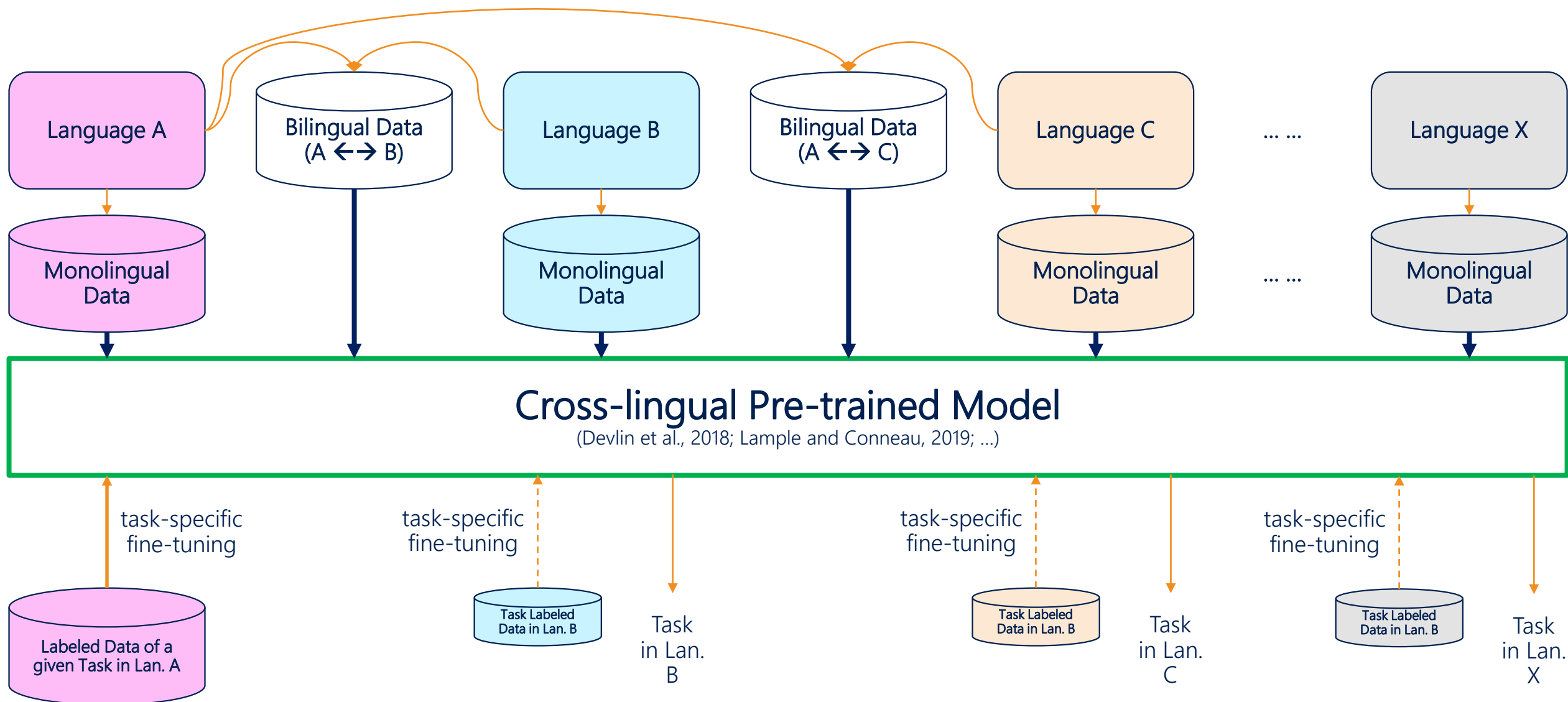
Transfer learning: learn from other tasks



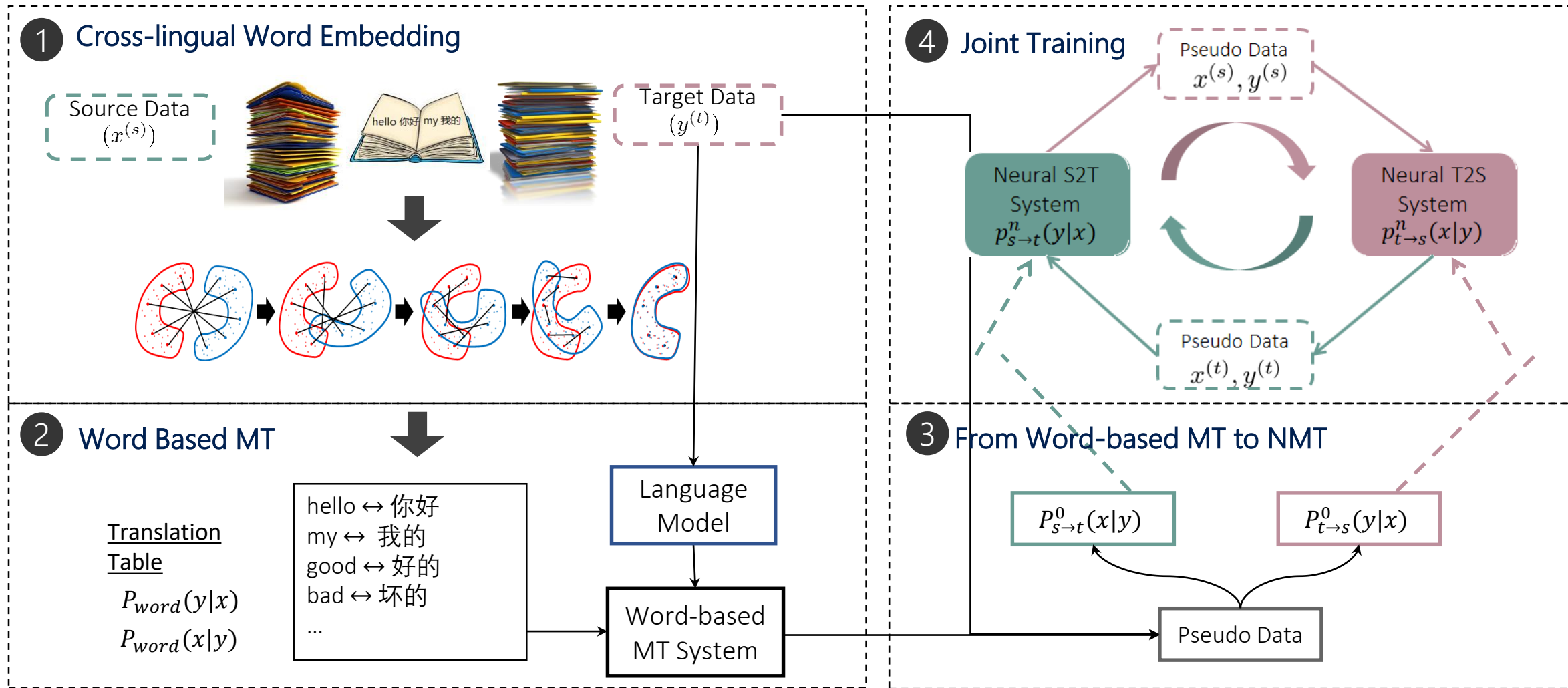
Successful Cases

- LM pre-training (BERT, GPT, XLNet) → various NLP tasks such as QA/MRC, summarization, paraphrase classification, etc.
- ImageNet pre-training (VGGNet, ResNet) → various CV tasks such as visual QA, object detection, scene graph generation, etc.
- ...

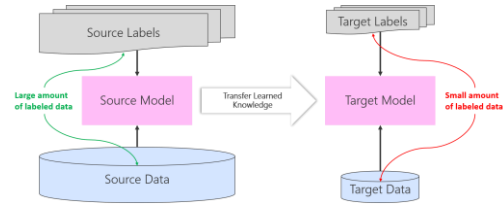
Cross-lingual learning: learn from other languages



Learning with seeds (lexicon, rules, small annotated data)



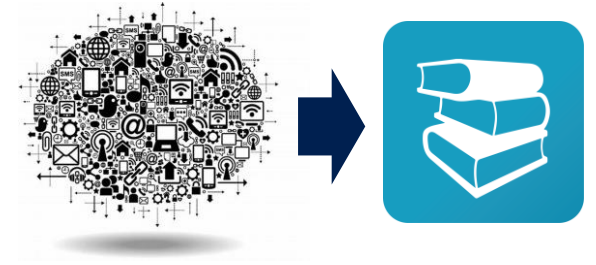
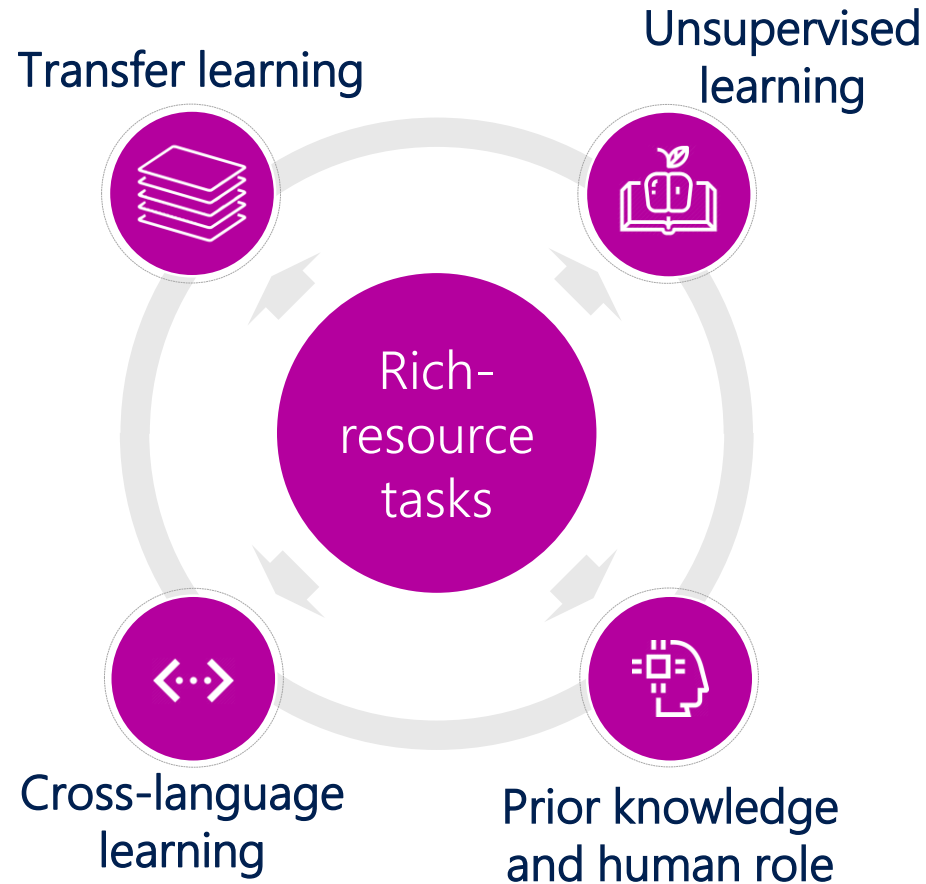
Important topics for low-resource tasks



Transfer knowledge learnt from rich-resource tasks to low-resource tasks, such as BERT and ResNet.



Learn mappings and relationships among languages for cross-lingual NLP tasks.



Discover knowledge from unannotated data based on distribution and patterns.



Cold-start with seeds such as rules and dictionary, active learning, reinforcement learning

Analysis on typical tasks



Rich-Resource

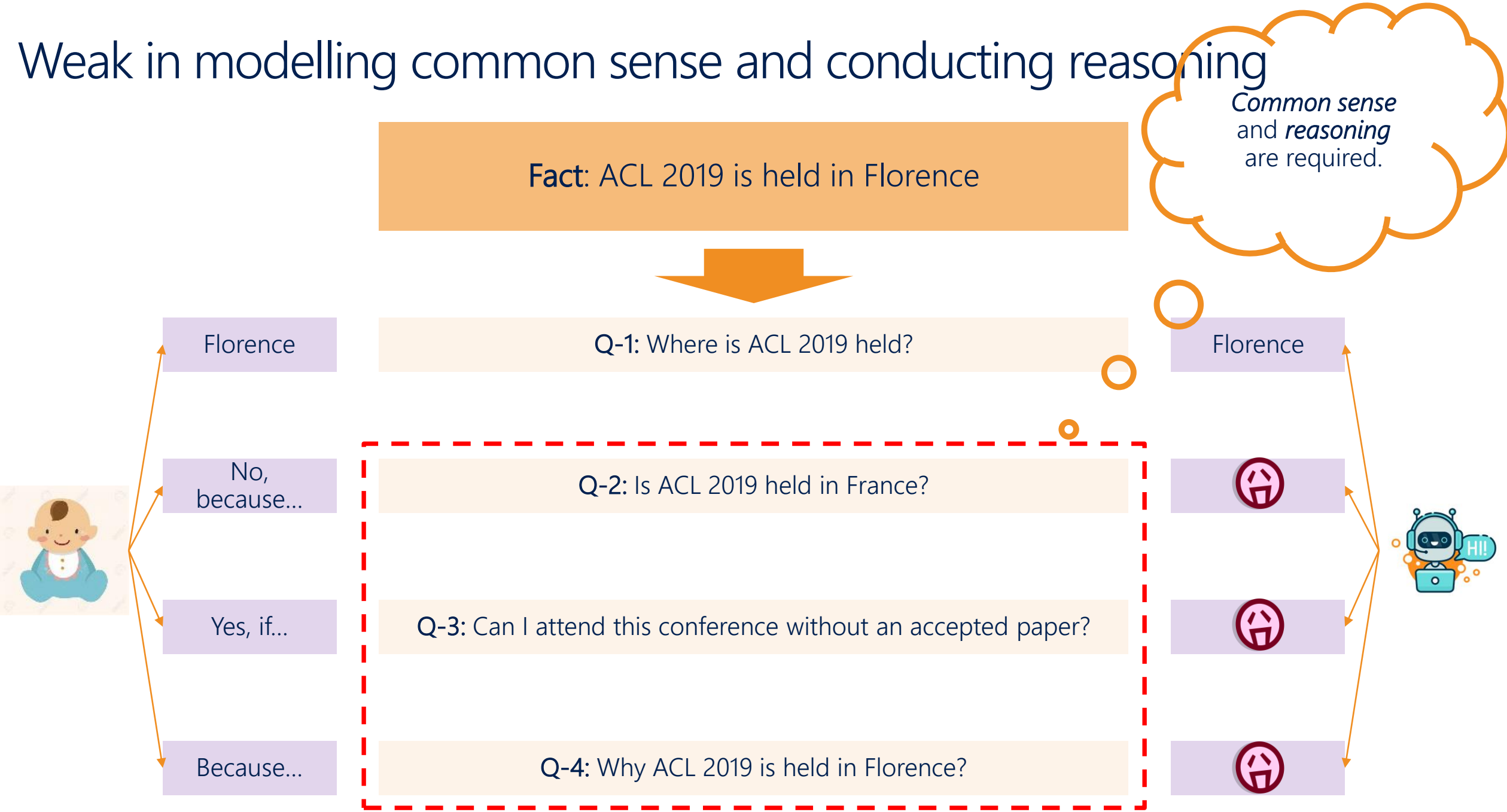


Low-Resource



Multi-Turn

Weak in modelling common sense and conducting reasoning

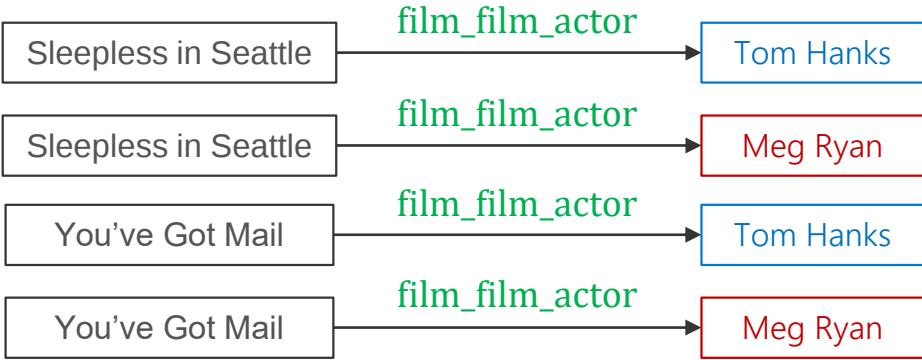


What kind reasoning is needed?

 Tell me the movies with Tom Hanks and Meg Ryan



Reasoning by semantic parsing using open domain knowledge

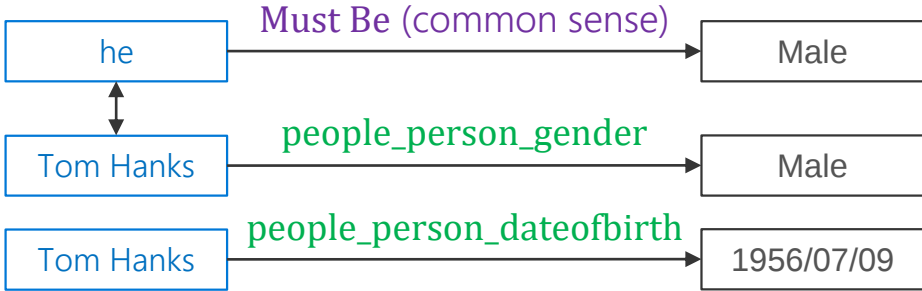


 When was he born ?



Sleepless in Seattle, You've Got Mail, ... 

Reasoning by semantic parsing and coreference resolution using common sense, open domain knowledge and context

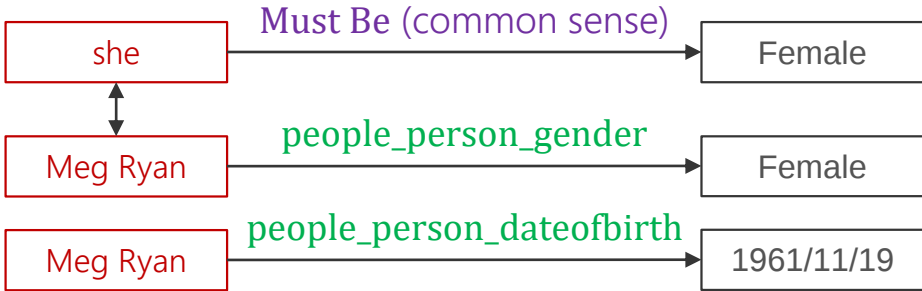


 How about her ?

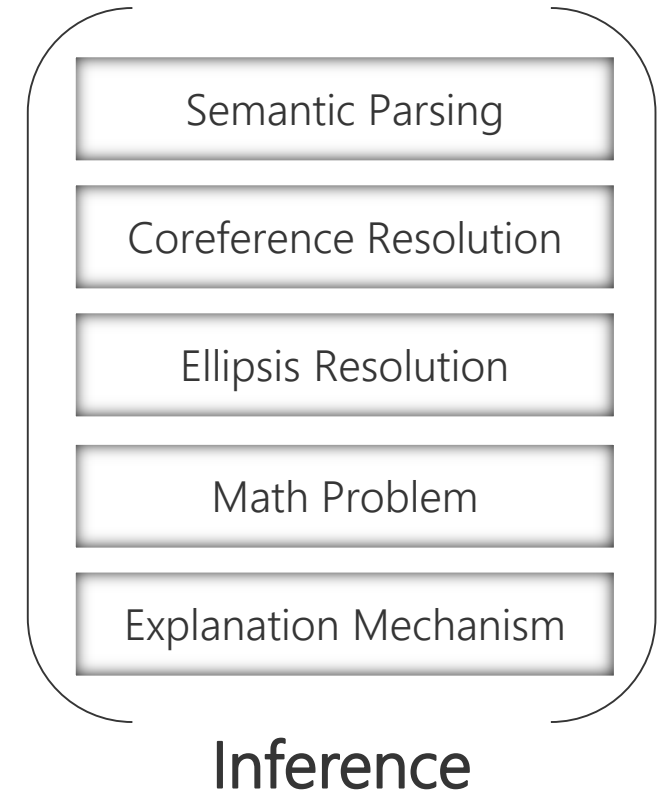
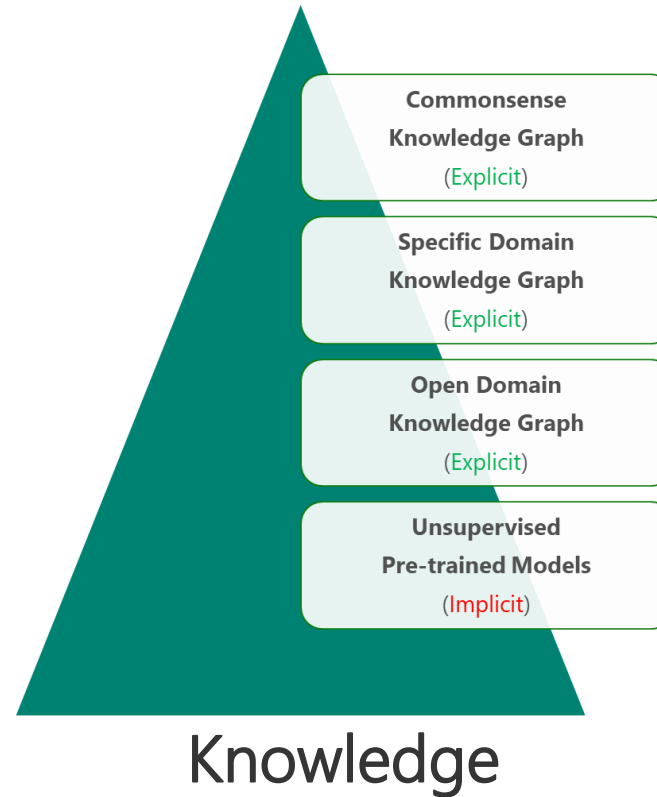
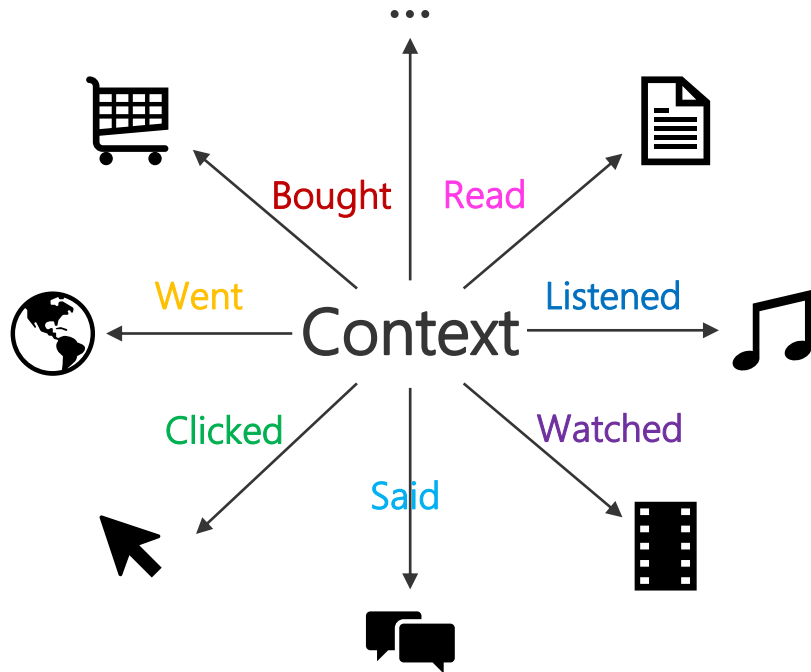


1956/07/09 

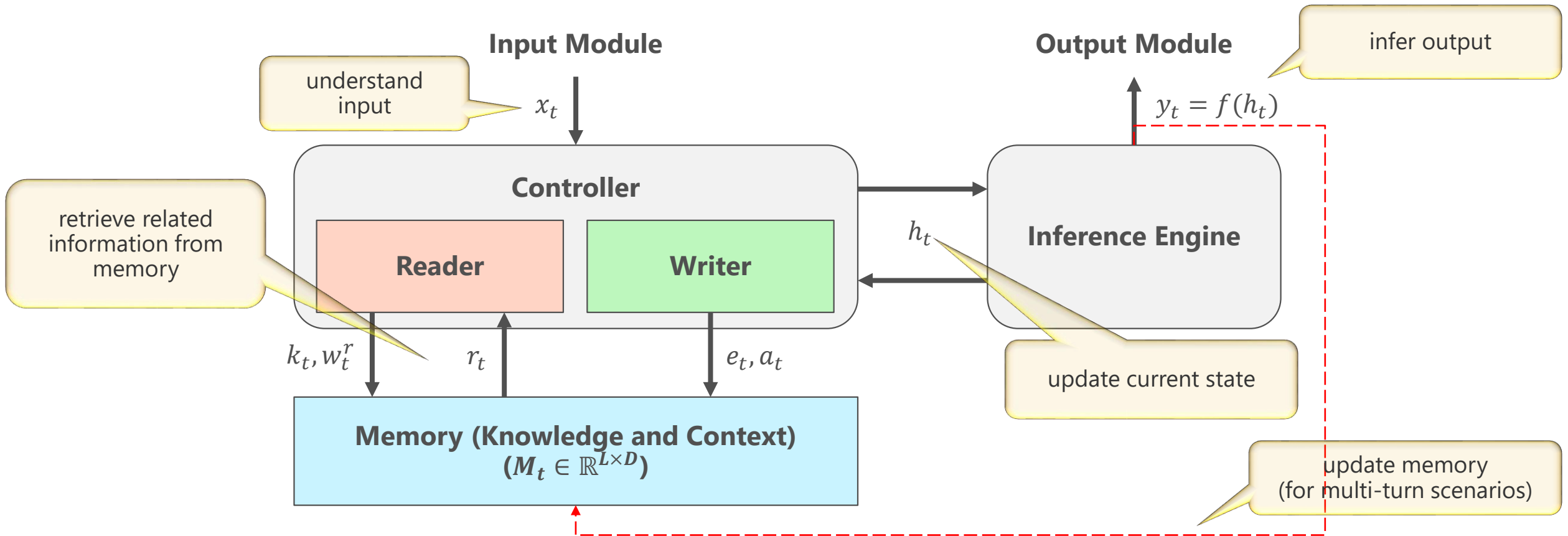
Reasoning by semantic parsing, coreference resolution and ellipsis resolution using common sense, open domain knowledge and context



Key components in reasoning

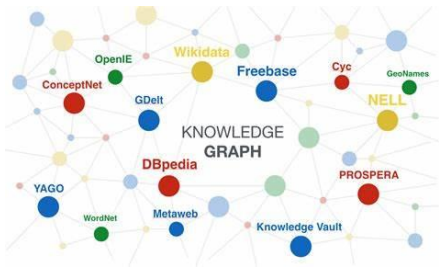


Concept model of reasoning with memory-augmented network

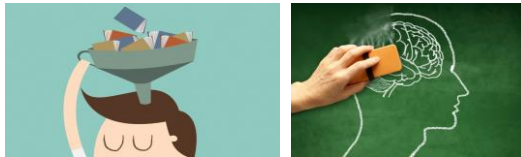


- Alex Graves, Greg Wayne, Ivo Danihelka. **Neural Turing Machines**. arXiv, 2014.
- Sainbayar Sukhbaatar, Arthur Szlam, Jason Weston, Rob Fergus. **End-To-End Memory Networks**. NeurIPS, 2015.
- Alexander Miller, Adam Fisch, Jesse Dodge, Amir-Hossein Karimi, Antoine Bordes, Jason Weston. **Key-Value Memory Networks for Directly Reading Documents**. EMNLP, 2016.
- Adam Santoro, Sergey Bartunov, Matthew Botvinick, Daan Wierstra, Timothy Lillicrap. **Meta-Learning with Memory-Augmented Neural Networks**. ICML, 2016.
- Drew A. Hudson and Christopher D. Manning. **Compositional Attention Networks for Machine Reasoning**. ICLR, 2018.
- ...

Important topics for multi-turn tasks



Extract, represent, conflate and use different types of knowledge and common sense..



Represent, memorize and forget context information in reasoning.

Knowledge and
Common sense

Inference
mechanism

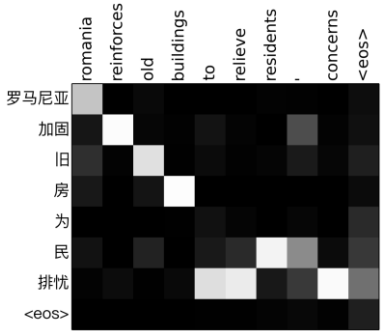
Context modeling

Explainability

Rich-
resource
tasks

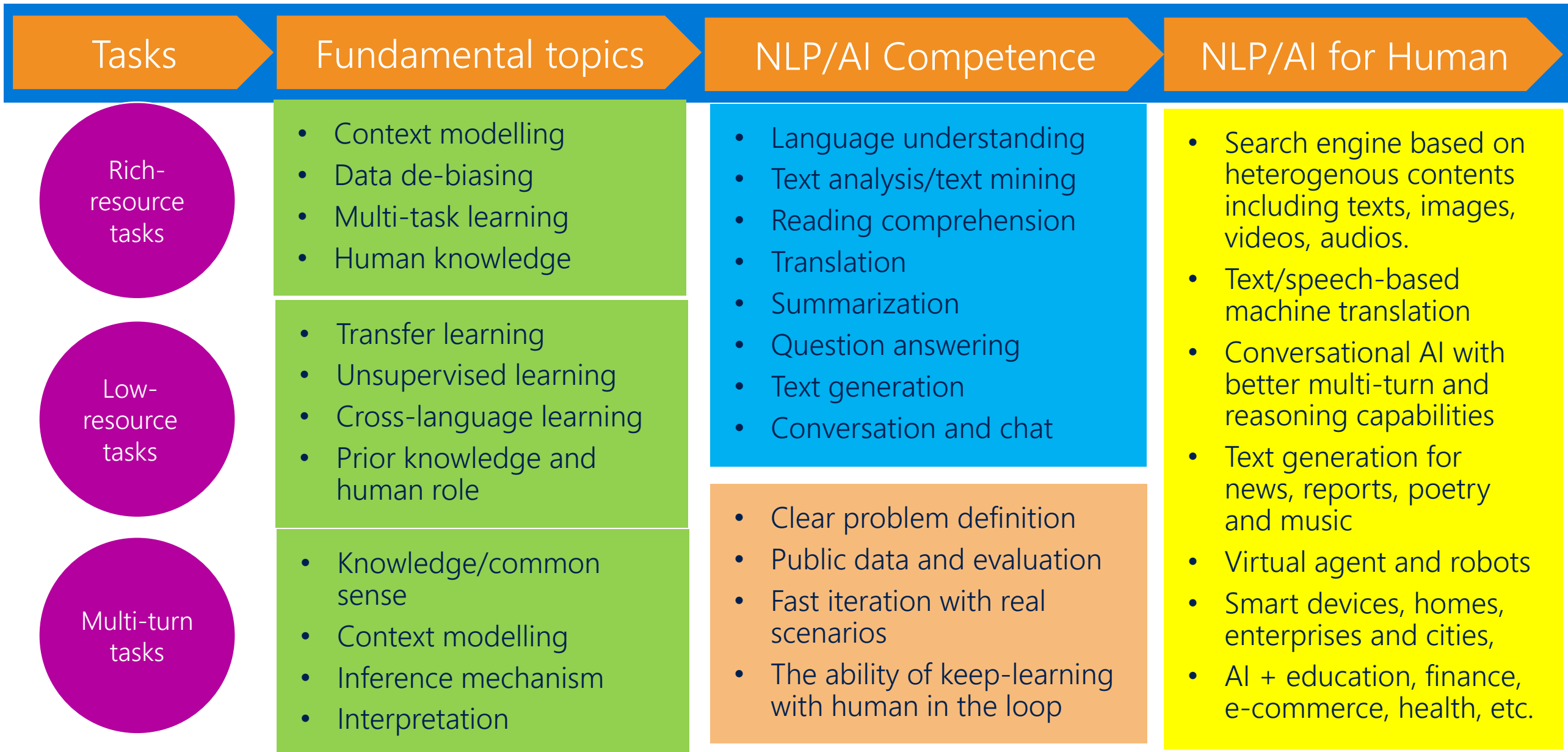


Annotate, model and evaluate the inference procedure.

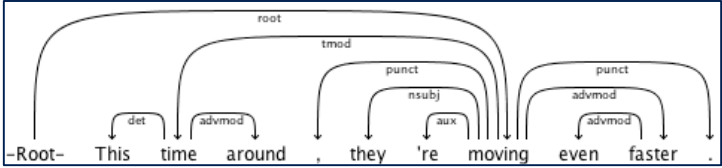
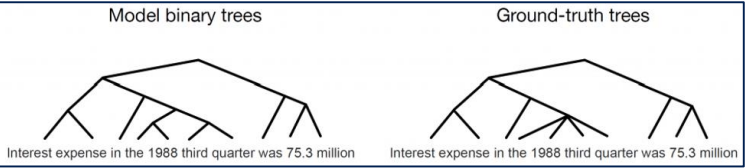


Mechanism, debugging, evaluation, visualization

Towards interpretable, knowledgeable, ethical, economical and non-stop-learnable NLP



Deep learning and linguistics boost each other

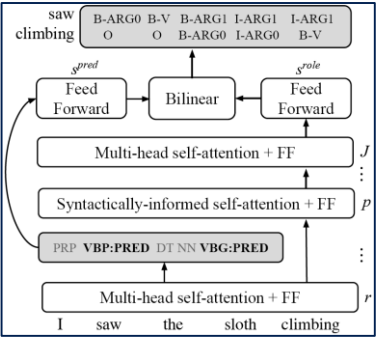
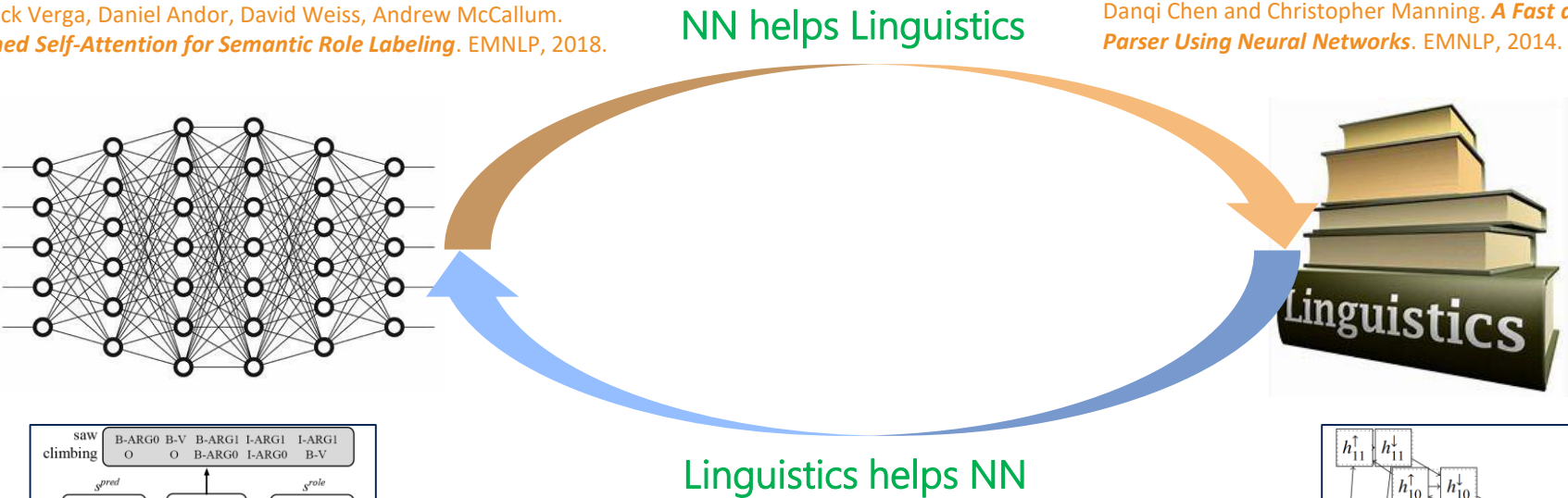


Deep learning models can find hidden syntactic tree structures of natural language sentences in an unsupervised way.

Emma Strubell, Patrick Verga, Daniel Andor, David Weiss, Andrew McCallum. *Linguistically-Informed Self-Attention for Semantic Role Labeling*. EMNLP, 2018.

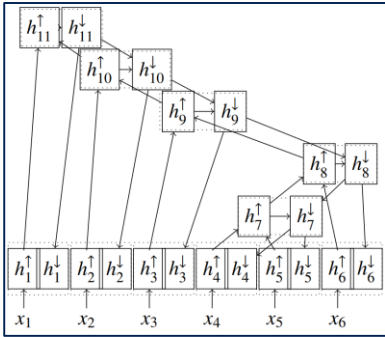
Deep learning models can predict better syntactic tree structures of natural language sentences in a supervised way.

Danqi Chen and Christopher Manning. *A Fast and Accurate Dependency Parser Using Neural Networks*. EMNLP, 2014.



Linguistic information can improve NLP tasks as input signals.

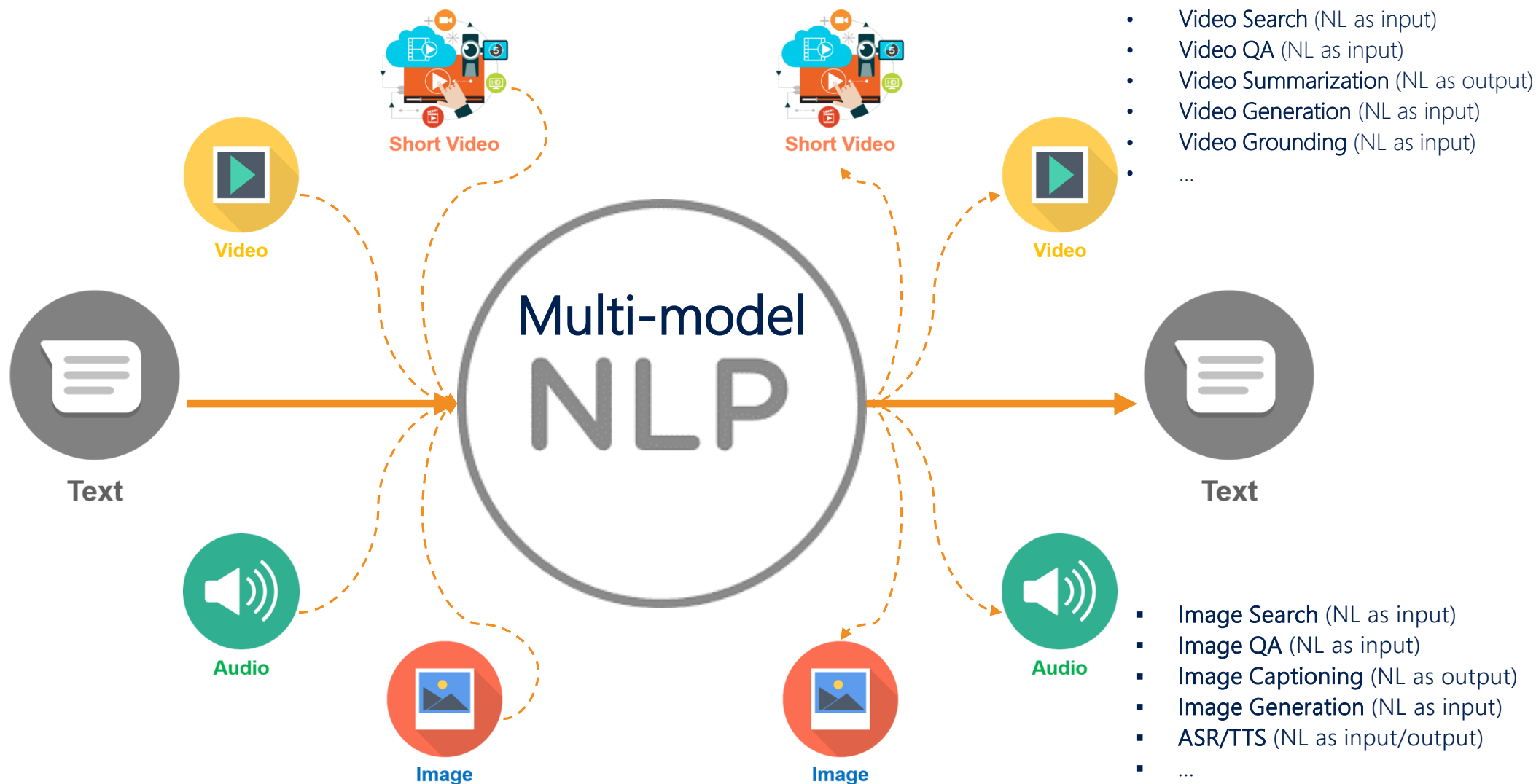
Yikang Shen, Shawn Tan, Alessandro Sordoni, Aaron Courville. *Ordered Neurons: Integrating Tree Structures into Recurrent Neural Networks*. ICLR, 2019.



Linguistic information can improve NLP tasks by designing syntactic-aware neural network structures..

Huadong Chen, Shujian Huang, David Chiang, Jiajun Chen. *Improved Neural Machine Translation with a Syntax-Aware Encoder and Decoder*. ACL, 2017.

Multi-modality processing to enrich input and output



Embrace the bright future with efforts from the whole society

Computing power

- Advanced chip and machine
- Powerful architecture and cloud computing
- Efficient resources management
- Model compression and acceleration

Talent

- Reform the curriculum
- Emphasize the system building capability
- Balance on following the trends and challenging the trends
- International view

Data

- Open-source data and shared tasks
- Efficient collection and annotation
- Data de-biasing and de-noising
- Privacy preserved learning

Collaboration

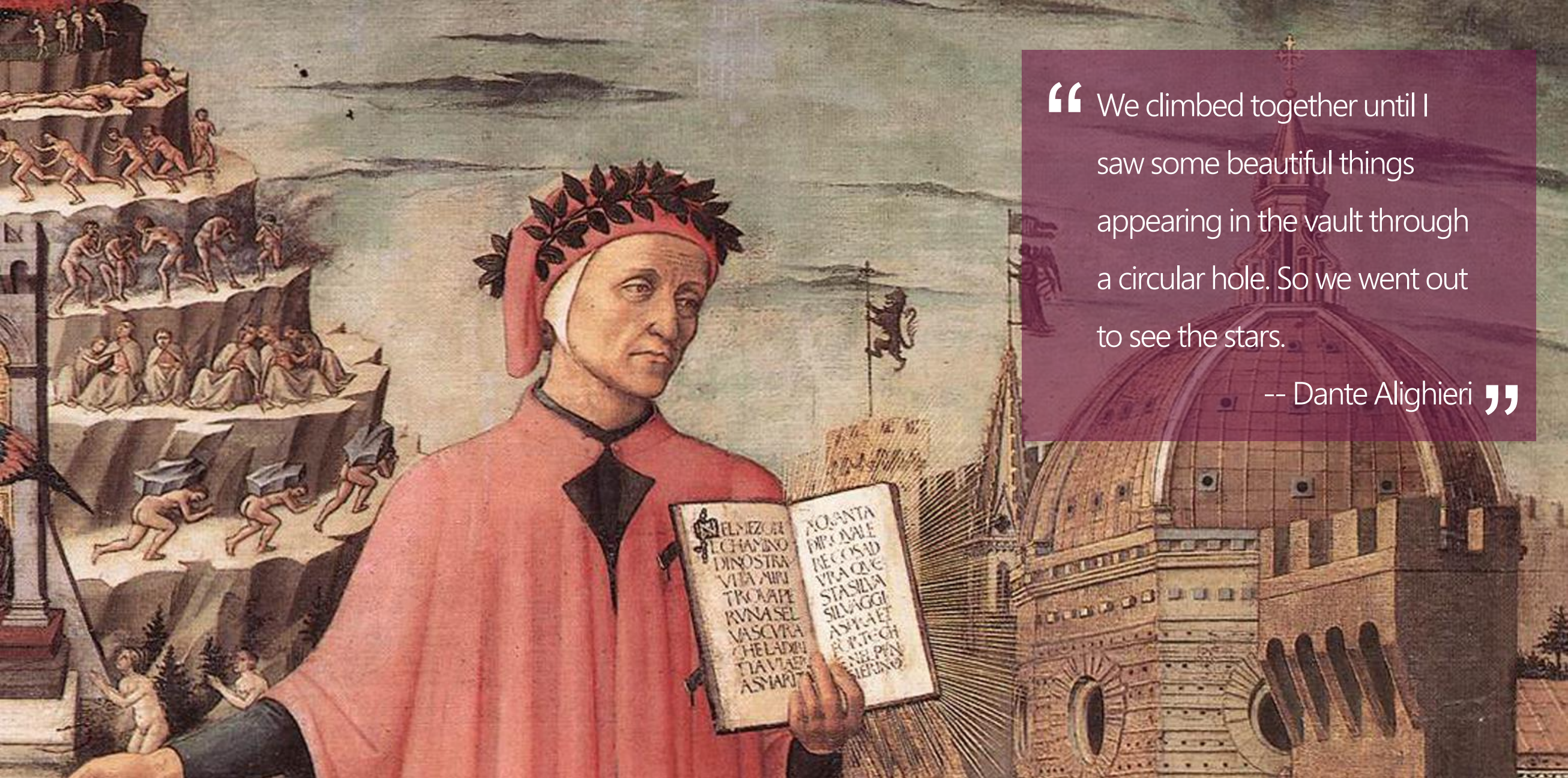
- University-enterprise
- Multi-domain and disciplinary(multi-modal processing, linguistics, brain science, ethics, big data,...)
- International partnership
- Eco-system with technical provider and users

Models

- New methods of supervised, less supervised and unsupervised
- Further development of pre-trained models
- Incorporating NN + Knowledge
- Reasoning and interpretability

Application

- Understand the needs of real scenarios of various verticals
- Result-oriented problem solving
- Human in the loop
- Market analysis and business model



“ We climbed together until I
saw some beautiful things
appearing in the vault through
a circular hole. So we went out
to see the stars.

-- Dante Alighieri ”

Thank you ~