

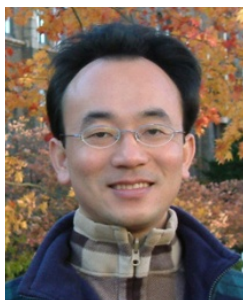


問答系統與對話系統 (Question Answering and Dialogue Systems)

Time: 2020/06/19 (Fri) (9:10 -12:00)

Place: 國立臺北護理健康大學 (台北市明德路365號) G210

Host: 祝國忠 院長 (健康科技學院院長)



Min-Yuh Day

戴敏育

Associate Professor

副教授

Dept. of Information Management, Tamkang University

淡江大學 資訊管理學系

<http://mail.tku.edu.tw/myday/>

2020-06-19



Topics

1. 自然語言處理核心技術與文字探勘

(Core Technologies of Natural Language Processing and Text Mining)

2. 人工智慧文本分析基礎與應用

(Artificial Intelligence for Text Analytics: Foundations and Applications)

3. 文本表達特徵工程

(Feature Engineering for Text Representation)

4. 語意分析和命名實體識別

(Semantic Analysis and Named Entity Recognition; NER)

5. 深度學習和通用句子嵌入模型

(Deep Learning and Universal Sentence-Embedding Models)

6. 問答系統與對話系統

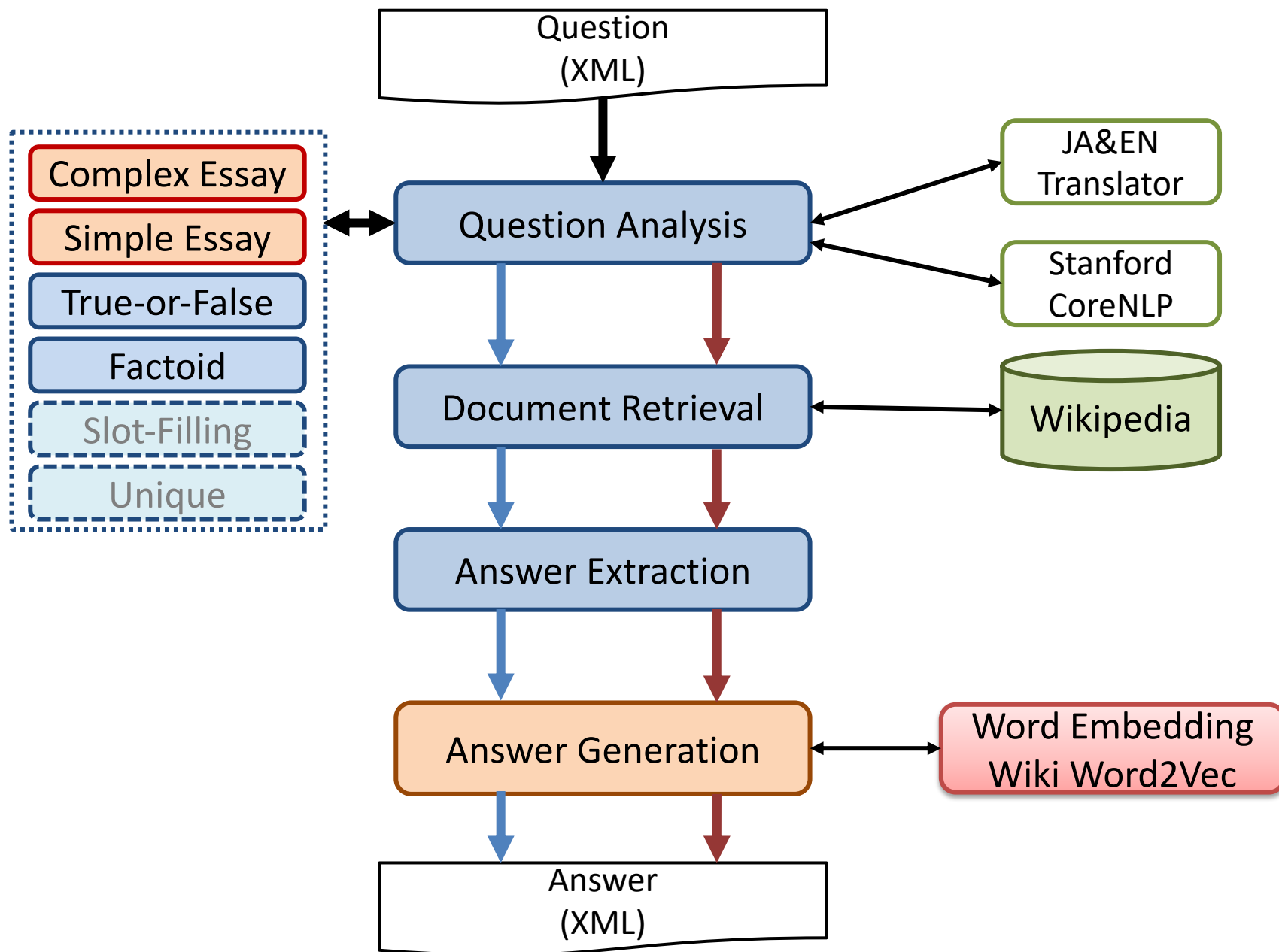
(Question Answering and Dialogue Systems)

Question Answering and Dialogue Systems

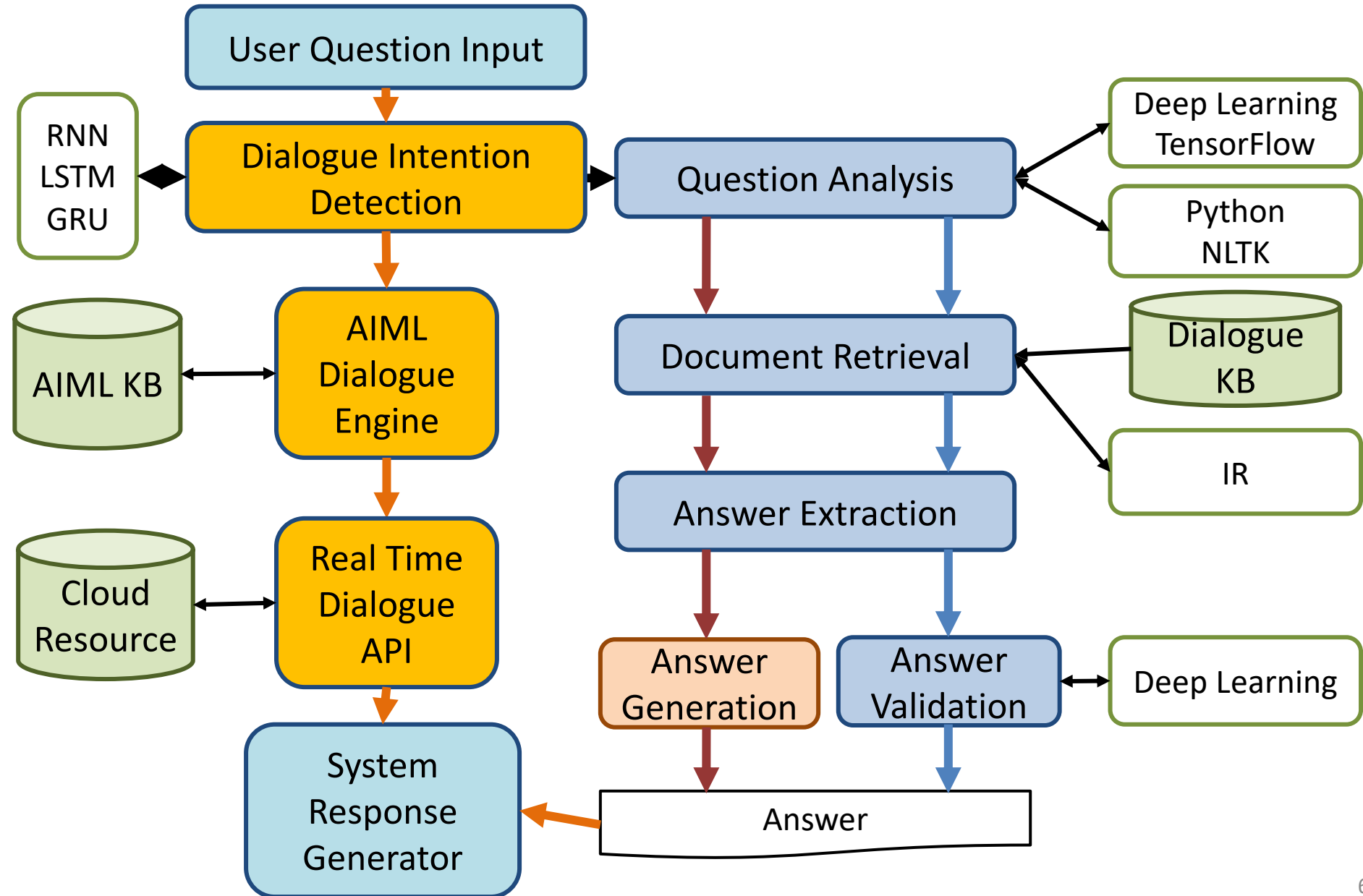
Outline

- Question Answering
- Dialogue Systems

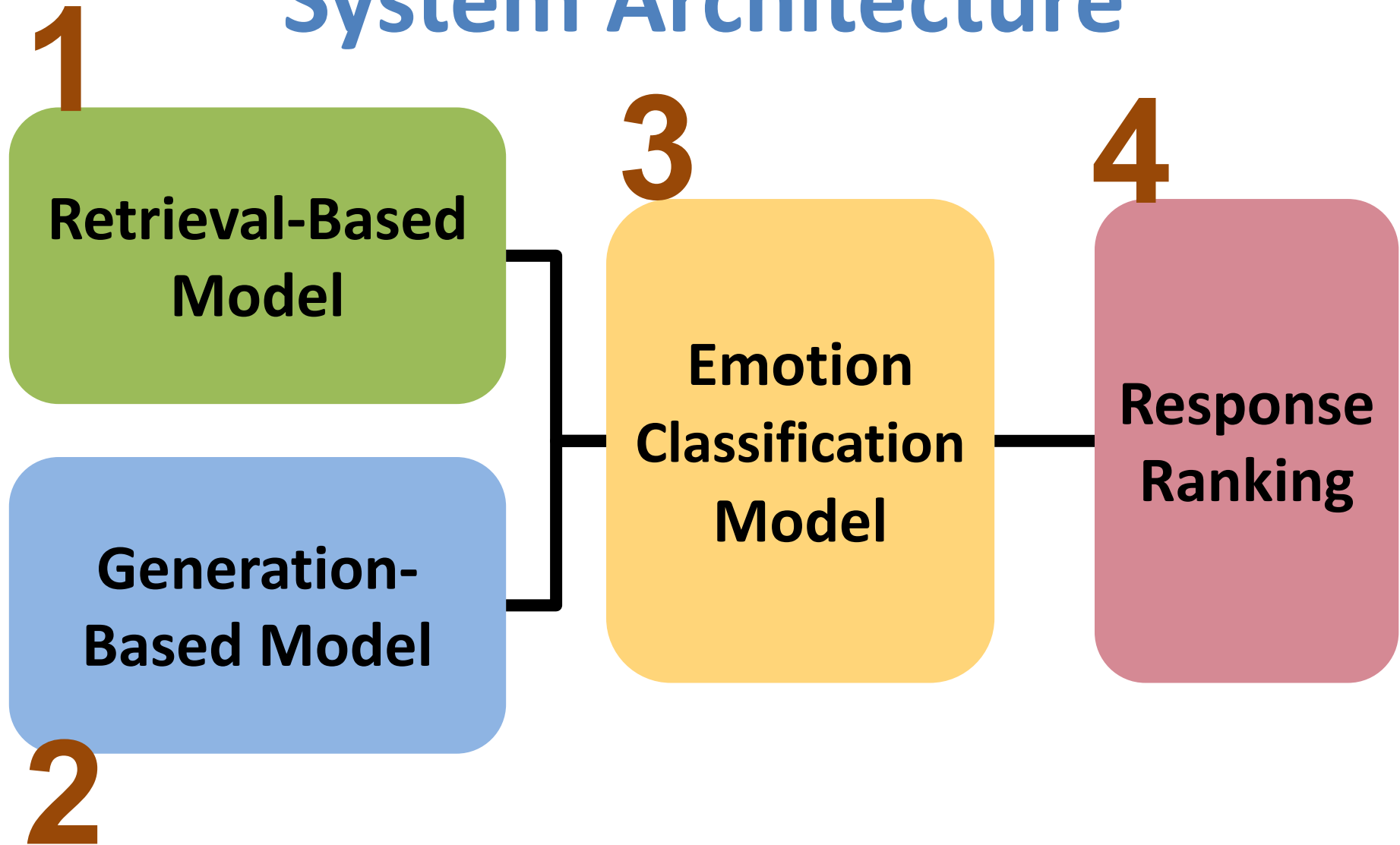
IMTKU System Architecture for NTCIR-13 QALab-3



System Architecture of Intelligent Dialogue and Question Answering System



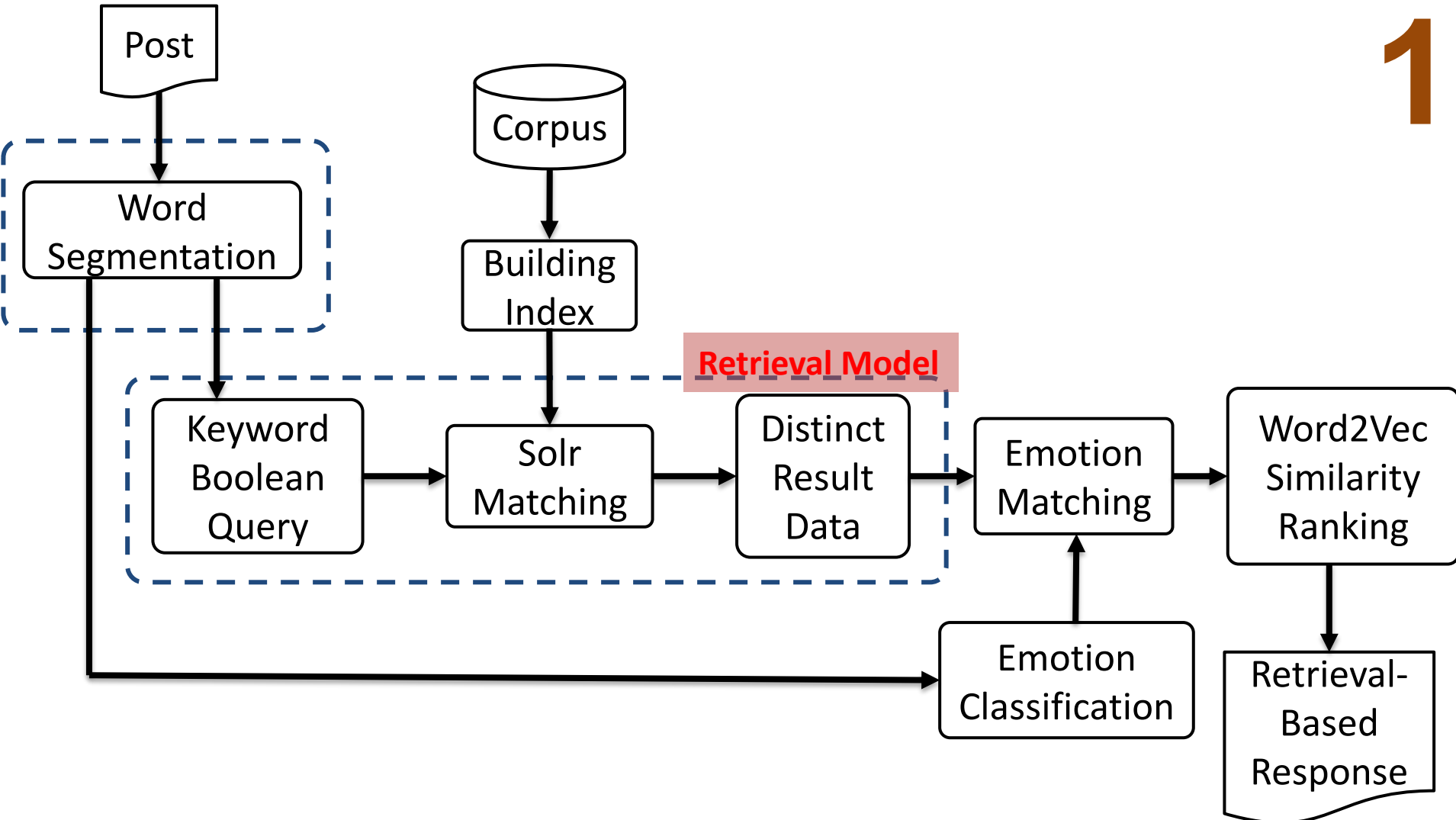
IMTKU Emotional Dialogue System Architecture



The system architecture of IMTKU retrieval-based model for NTCIR-14 STC-3

Retrieval-Based Model

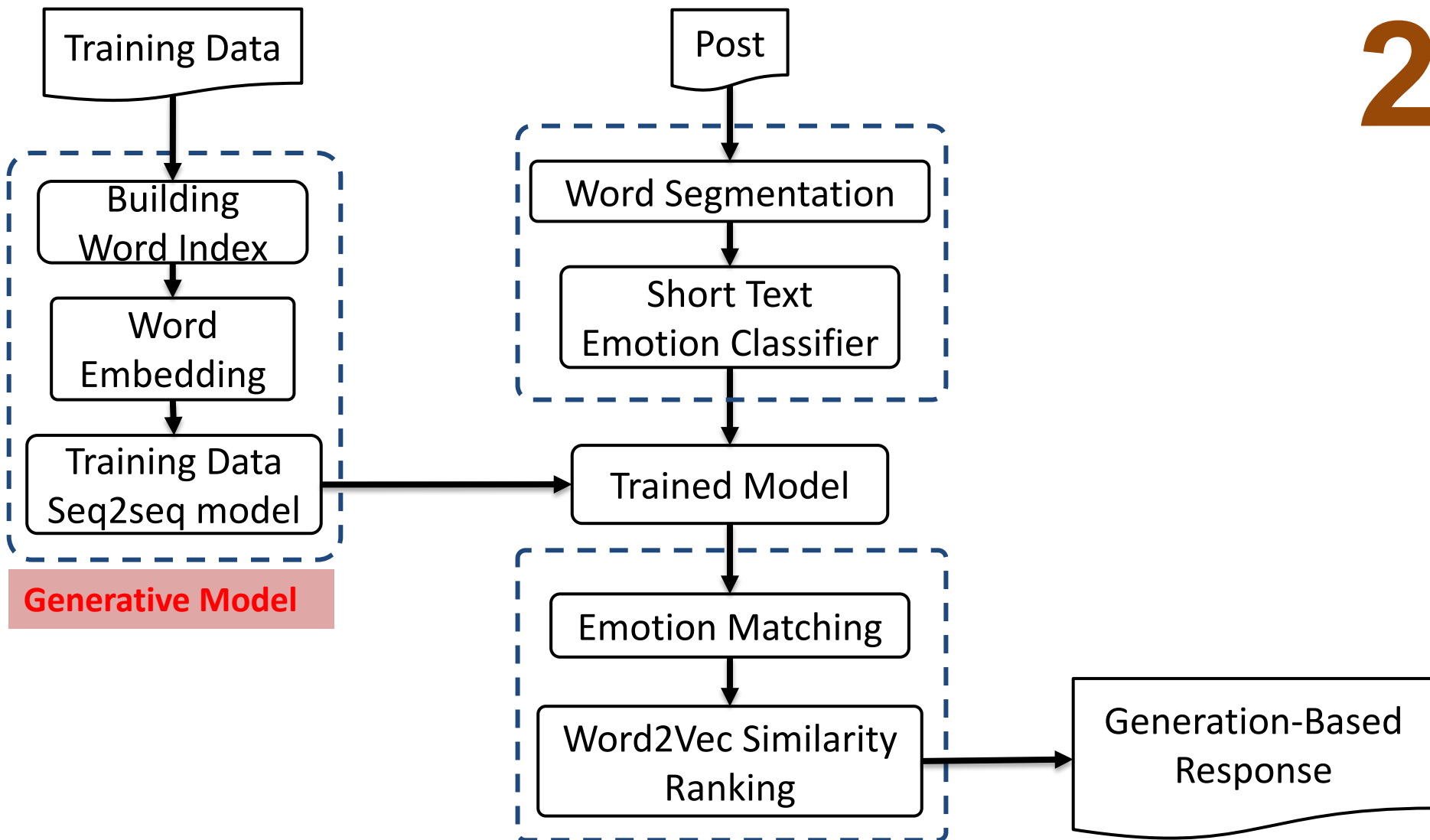
1



The system architecture of IMTKU generation-based model for NTCIR-14 STC-3

Generation-Based Model

2

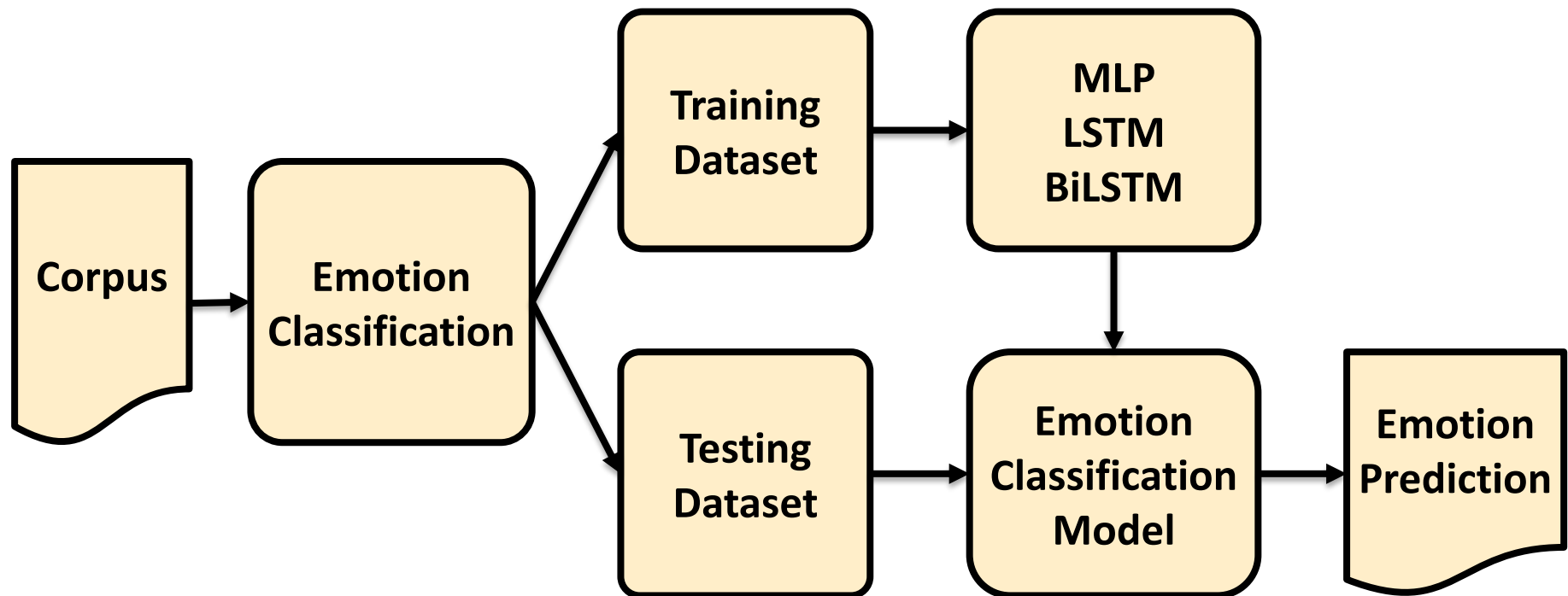


The system architecture of IMTKU emotion classification model for NTCIR-14 STC-3



Emotion Classification Model

3

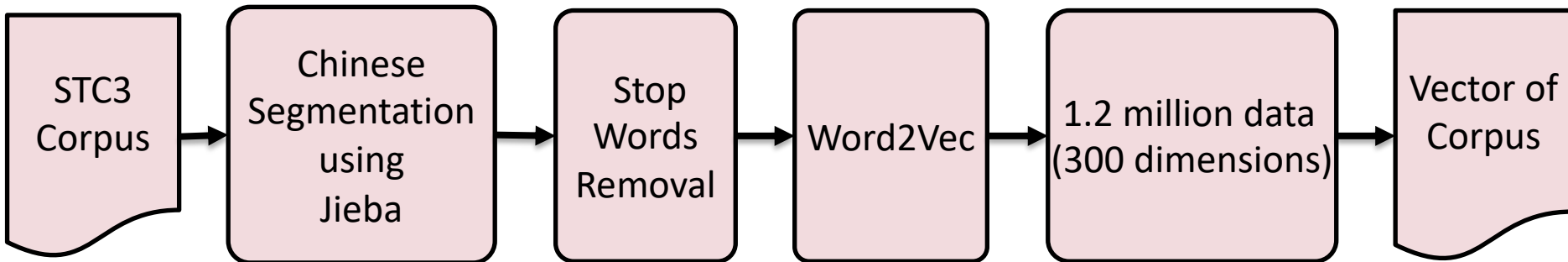


The system architecture of IMTKU Response Ranking for NTCIR-14 STC-3



Response Ranking

4





**Short Text Conversation Task
(STC-3)
Chinese Emotional Conversation
Generation (CECG) Subtask**

NTCIR Short Text Conversation

STC-1, STC-2, STC-3

	Japanese	Chinese	English	
NTCIR-12 STC-1 22 active participants	Twitter, Retrieval	Weibo, Retrieval		Single-turn, Non task-oriented
NTCIR-13 STC-2 27 active participants	Yahoo! News, Retrieval+ Generation	Weibo, Retrieval+ Generation		
NTCIR-14 STC-3		Weibo, Generation for given emotion categories		Multi-turn, task-oriented (helpdesk)
Chinese Emotional Conversation Generation (CECG) subtask				
Dialogue Quality (DQ) and Nugget Detection (ND) subtasks		Weibo+English translations, distribution estimation for subjective annotations		

Source: <https://waseda.app.box.com/v/STC3atNTCIR-14>

Chatbots: Evolution of UI/UX

Paradigm	mid - 80s PC	mid - 90s Web	mid - 00s Smartphone	mid - 10s Messaging
Platform Examples	Desktop DOS, Windows, Mac OS	Browser Mosaic, Explorer, Chrome	Mobile OS iOS, Android	Messaging Apps WhatsApp, Messenger, Slack
Applications Examples	Clients Excel, PPT, Lotus	Website Yahoo, Amazon	Apps Angry Birds, Instagram	Bots Weather, Travel
UI/UX	Native Screens	Web Pages	Native Mobile Screens	Message
S/w Dev	Client-side	Server-side	Client-side	Server-side

AI Dialogue System

Dialogue Subtasks

[Browse](#) > [Natural Language Processing](#) > Dialogue

Dialogue subtasks

Dialogue Generation

Dialogue Generation

🏆 9 leaderboards

35 papers with code



Dialogue State Tracking

🏆 2 leaderboards

30 papers with code



Visual Dialog

🏆 3 leaderboards

28 papers with code

Task-Oriented Dialogue Systems

Task-Oriented Dialogue Systems

20 papers with code



Goal-Oriented Dialog

15 papers with code

Short-Text Conversation

Dialogue Management

10 papers with code



Dialogue Understanding

6 papers with code

Short-Text Conversation

5 papers with code

Goal-Oriented Dialogue Systems

3 papers with code

Task-Completion Dialogue Policy Learning

2 papers with code

Chatbot

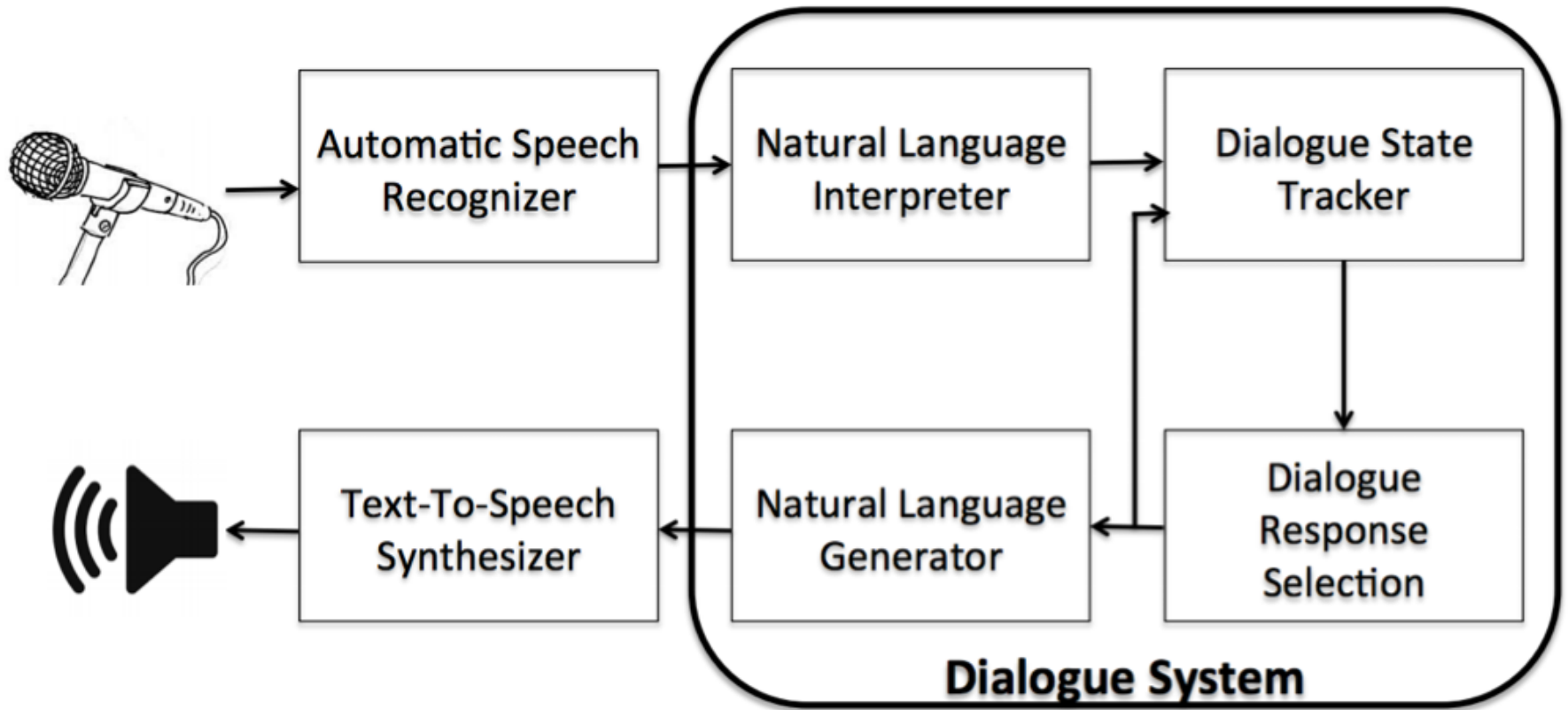
Dialogue System

Intelligent Agent

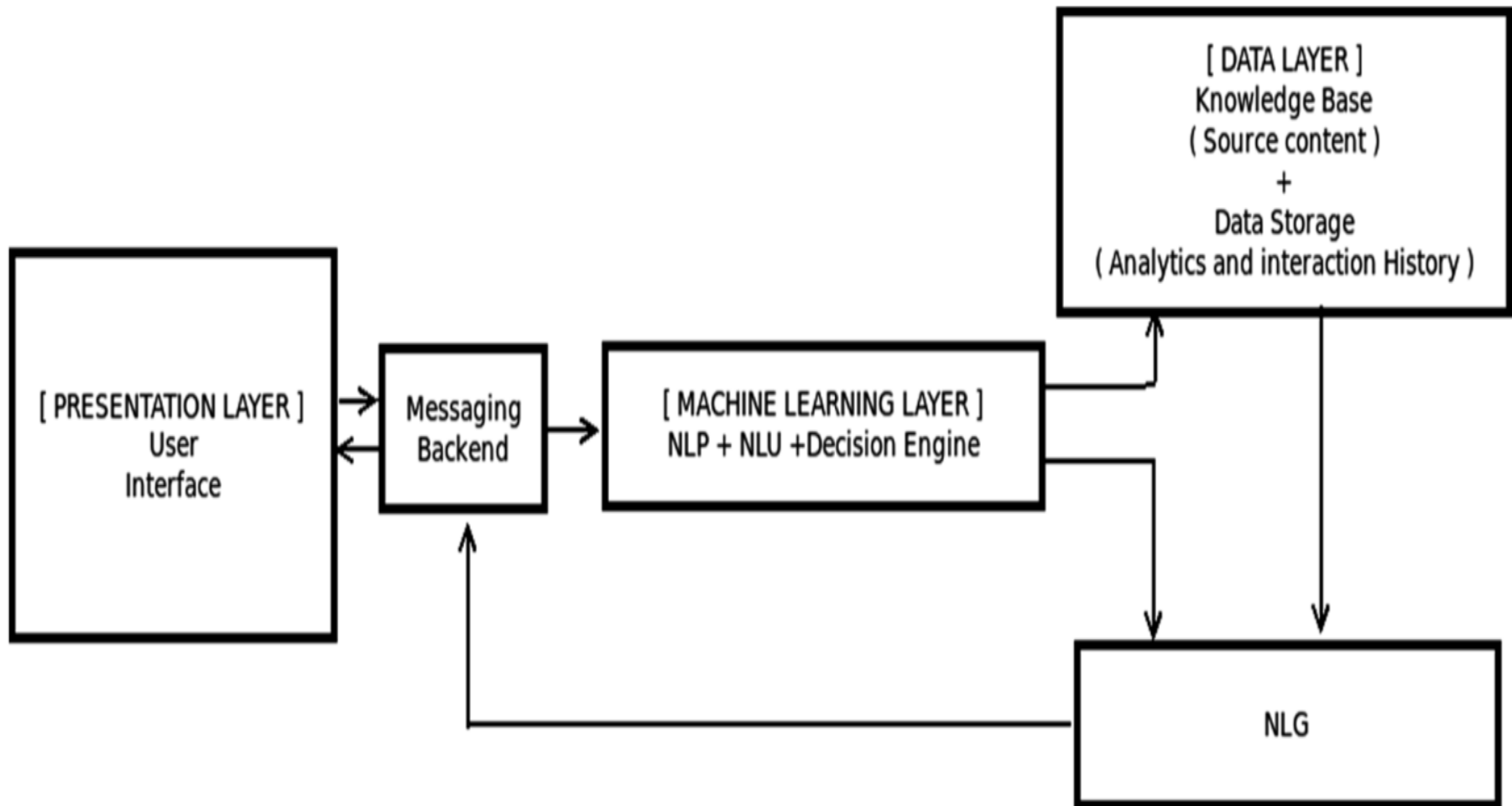
Chatbot



Dialogue System



Overall Architecture of Intelligent Chatbot



Can machines think?

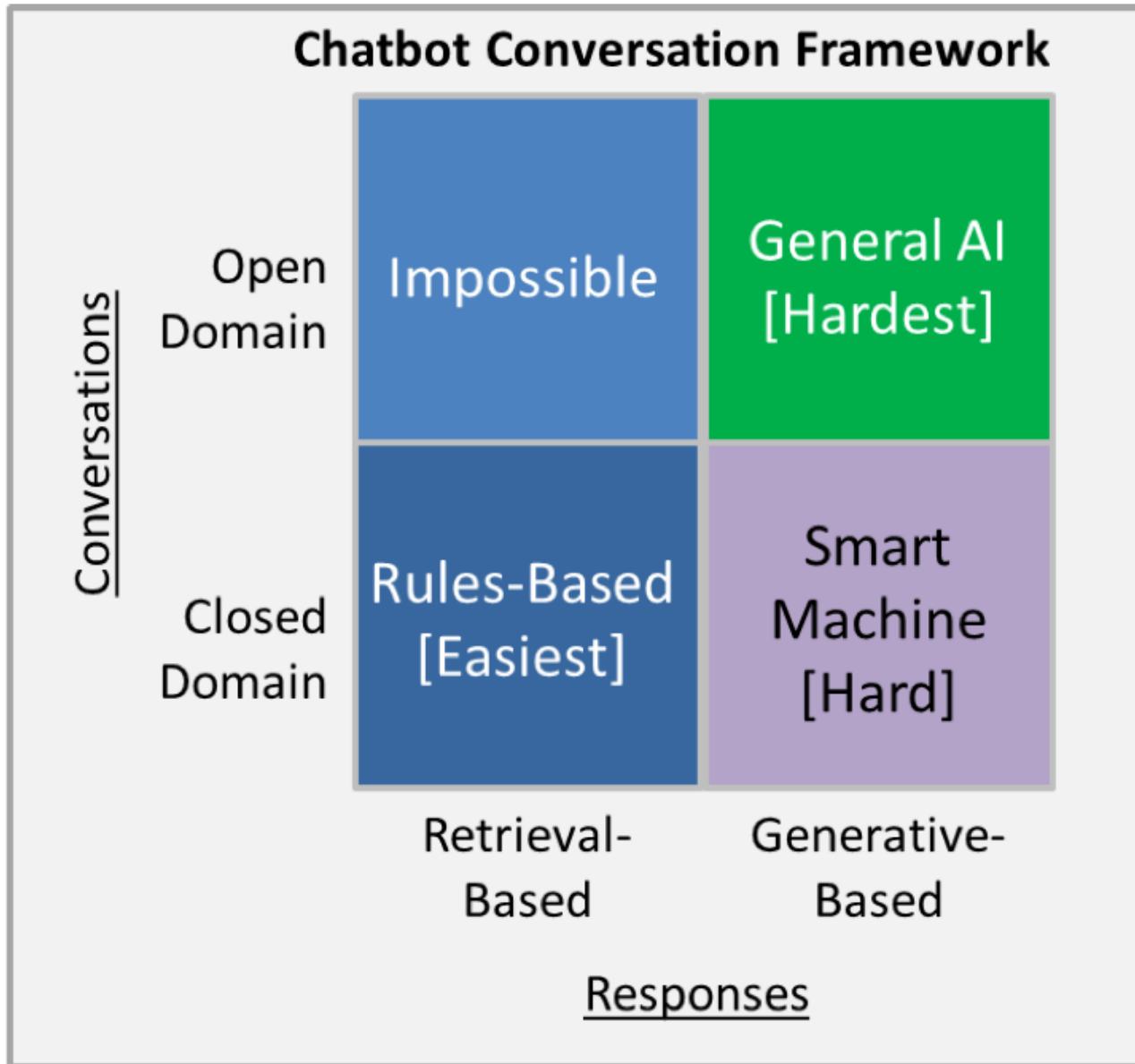
(Alan Turing ,1950)

Source: Cahn, Jack. "CHATBOT: Architecture, Design, & Development."
PhD diss., University of Pennsylvania, 2017.

Chatbot

**“online human-computer
dialog system
with
natural language.”**

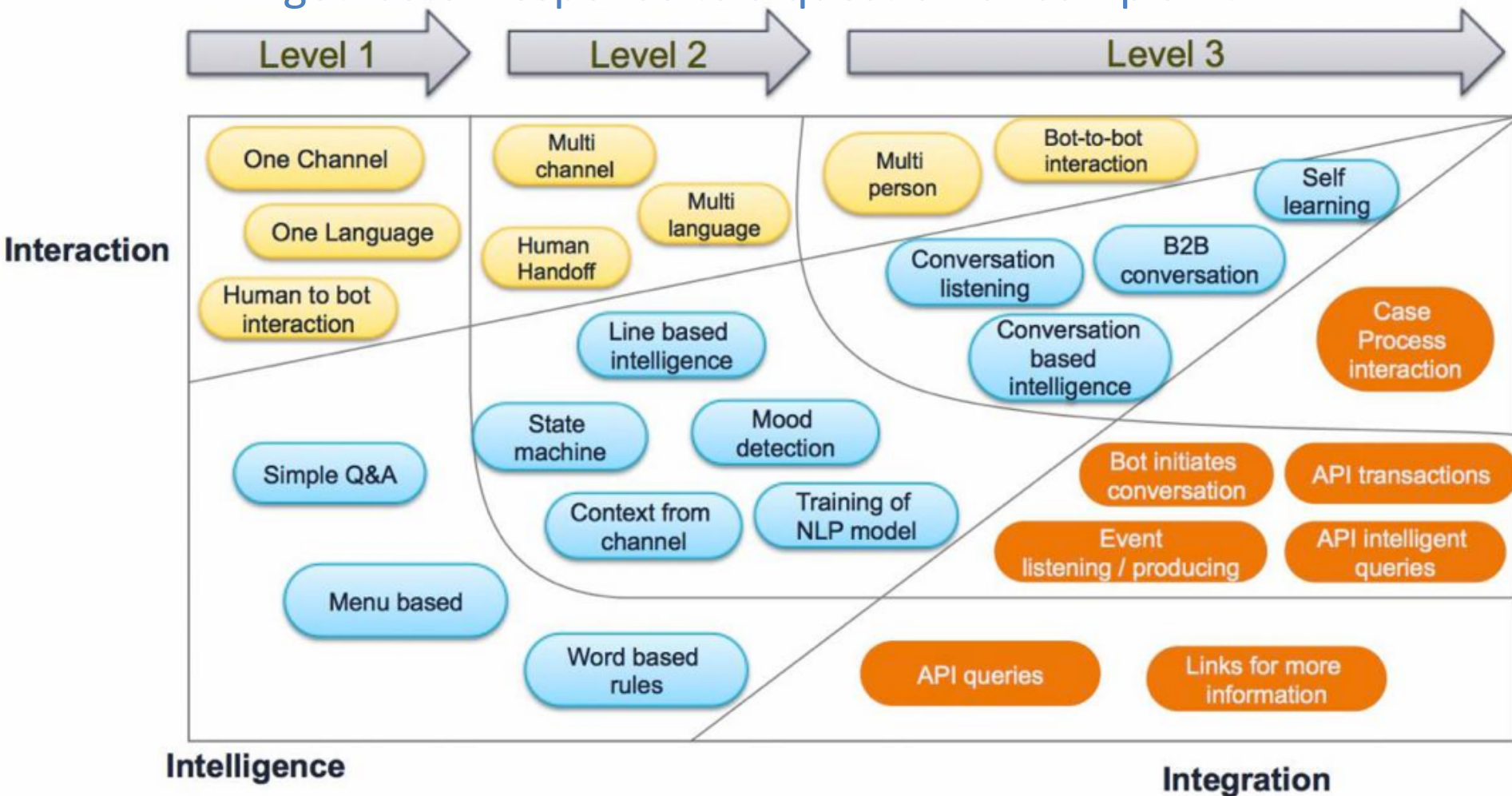
Chatbot Conversation Framework



Chatbots

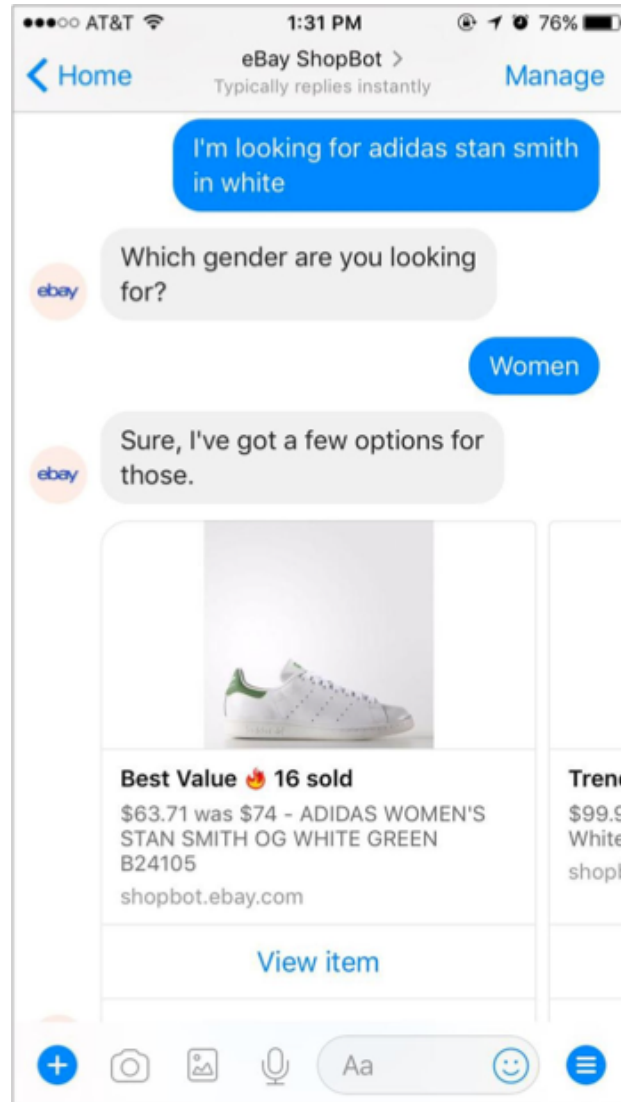
Bot Maturity Model

Customers want to have simpler means to interact with businesses and get faster response to a question or complaint.

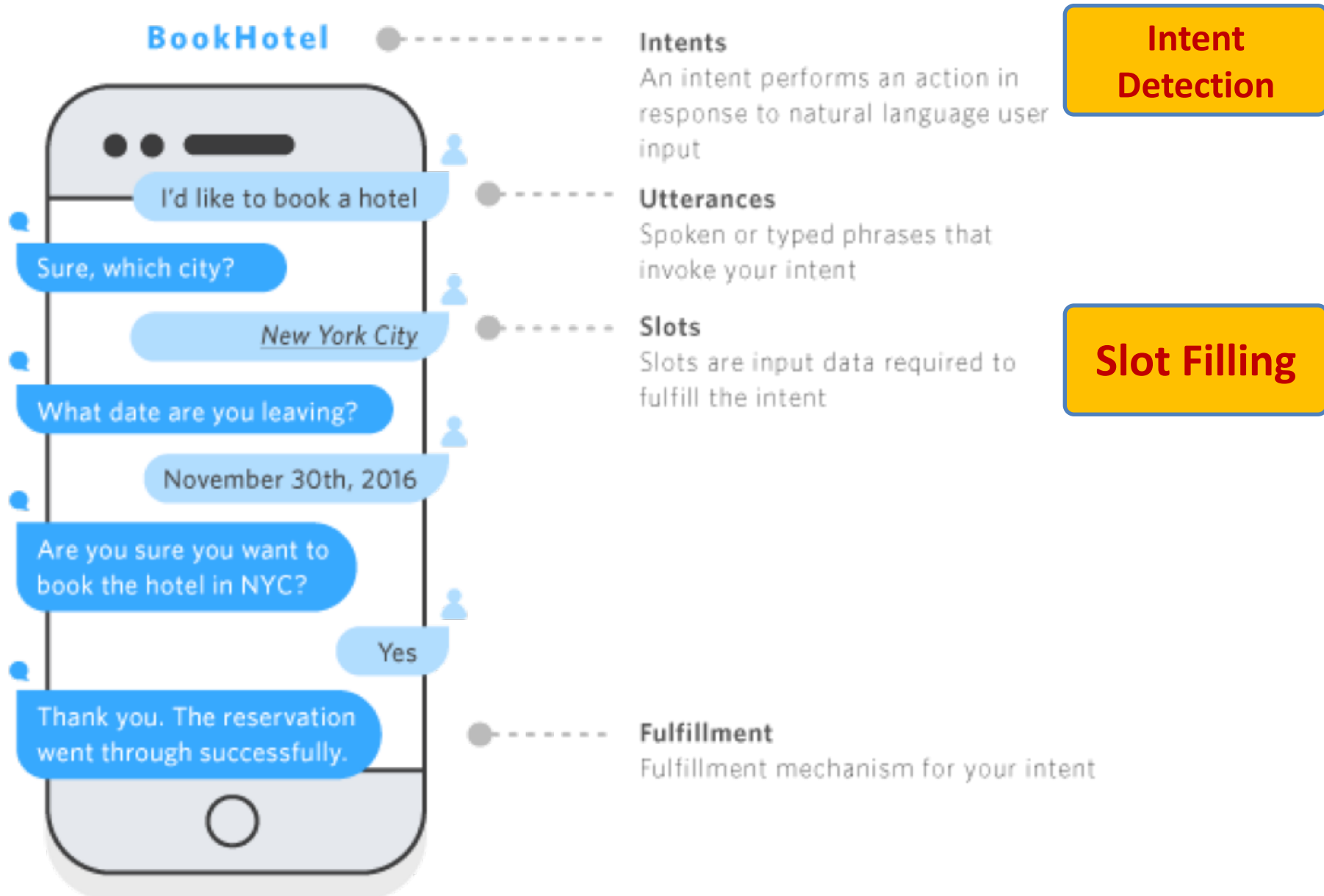


**From
E-Commerce
to
Conversational Commerce:
Chatbots
and
Virtual Assistants**

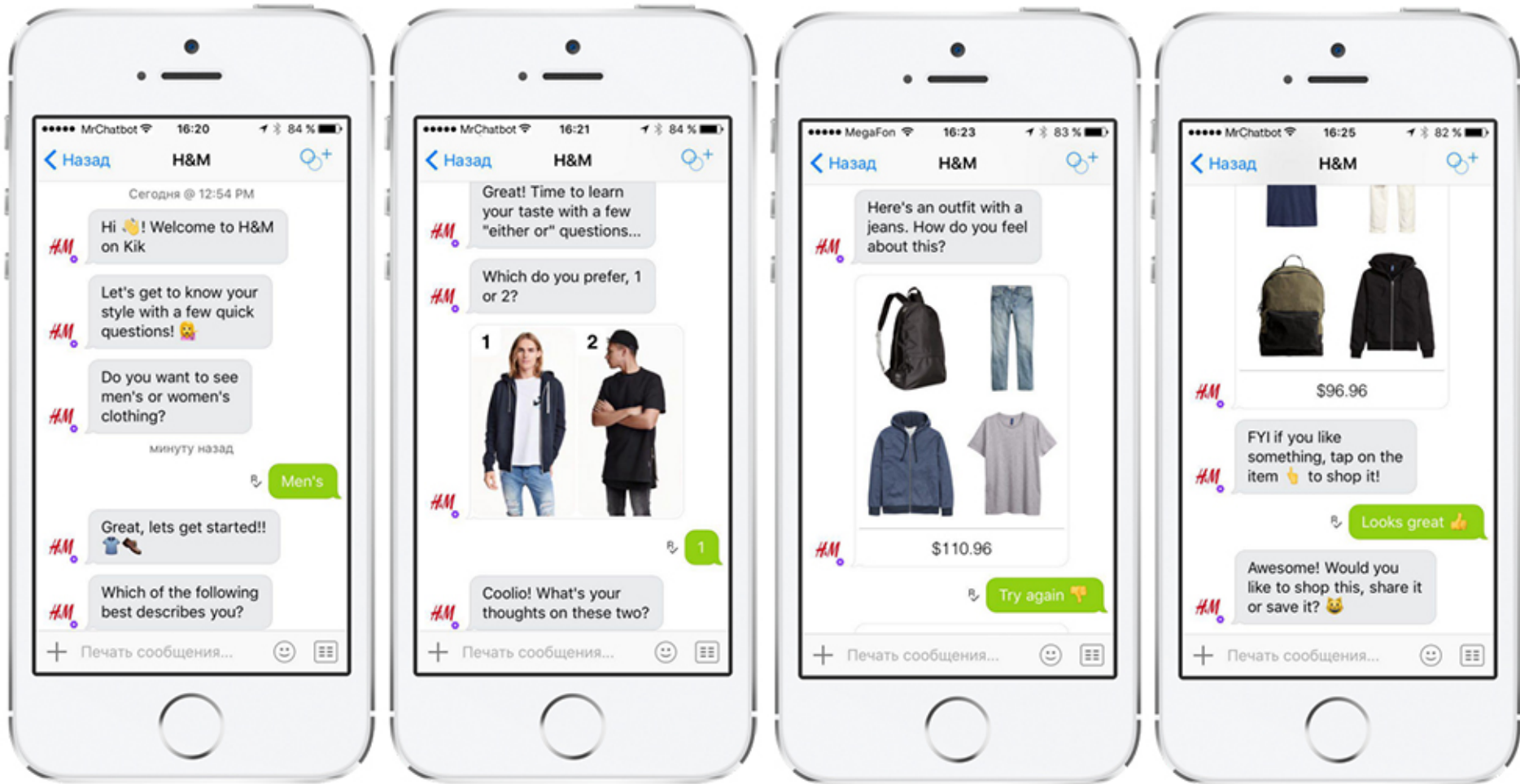
Conversational Commerce: eBay AI Chatbots



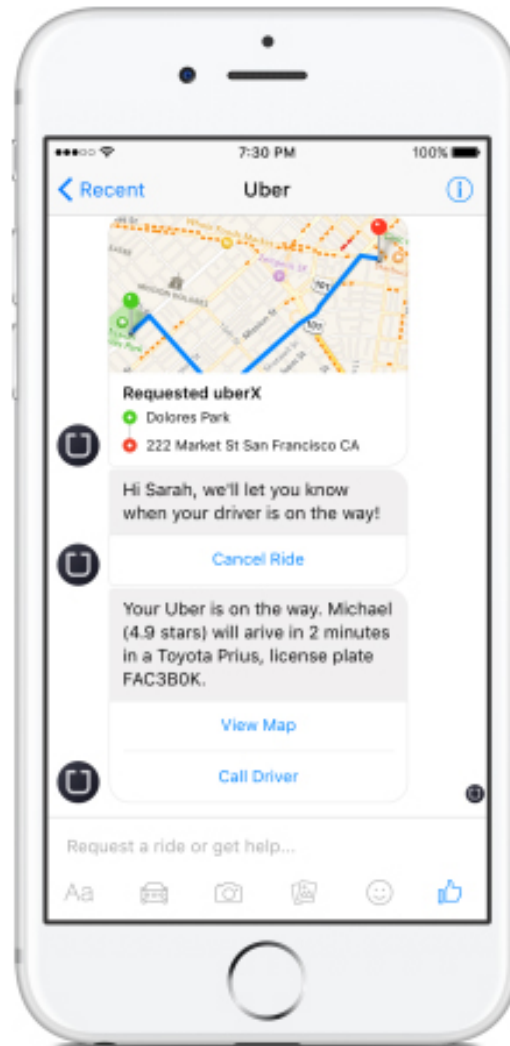
Hotel Chatbot



H&M's Chatbot on Kik



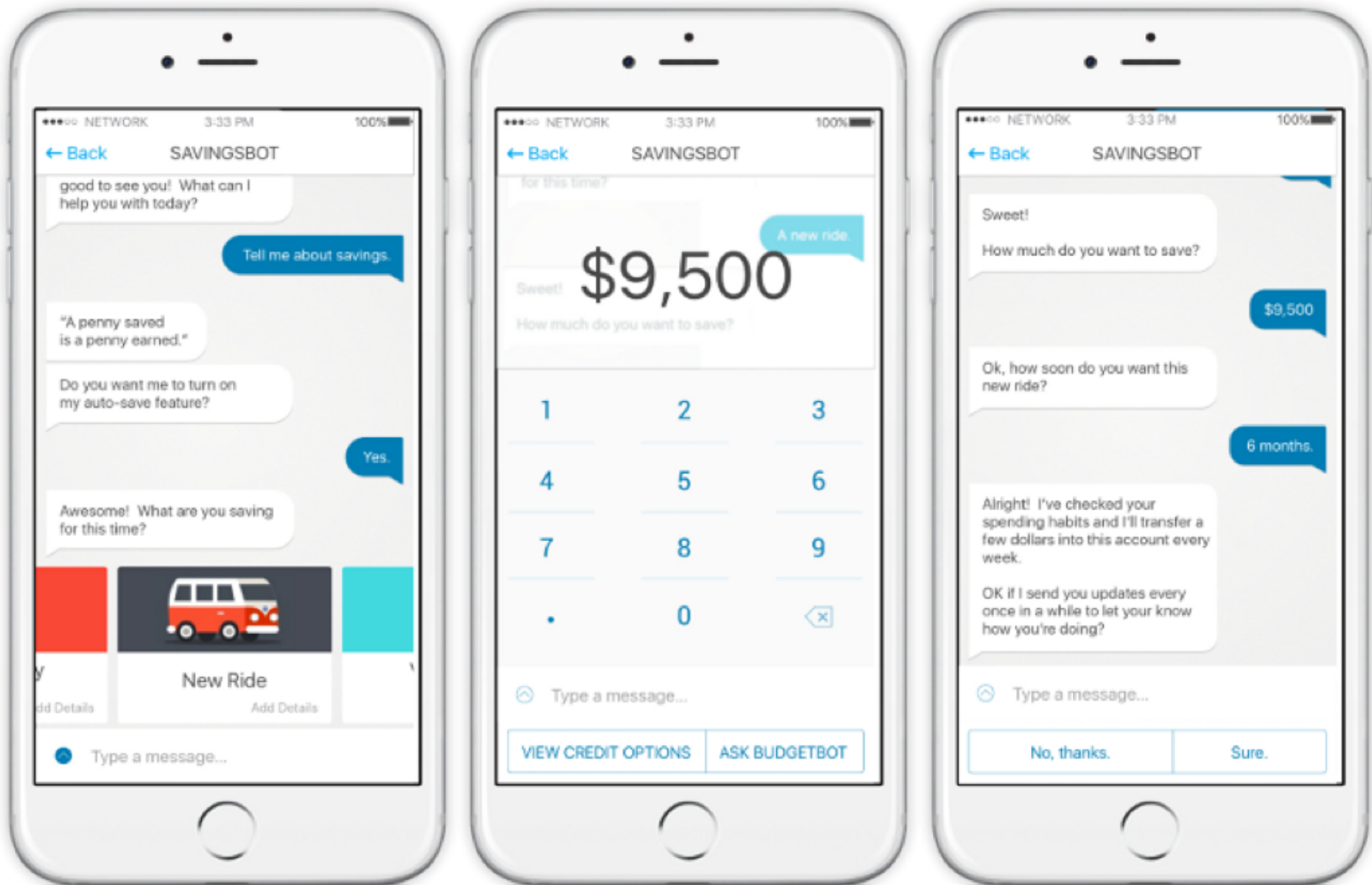
Uber's Chatbot on Facebook's Messenger



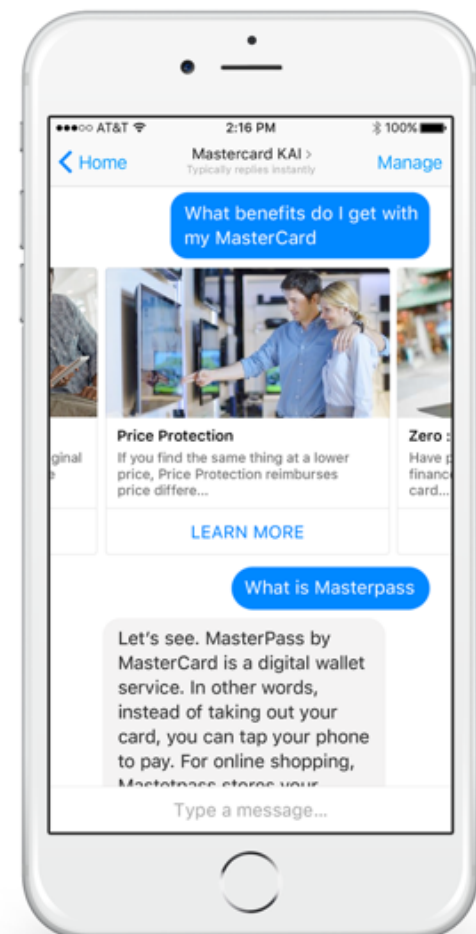
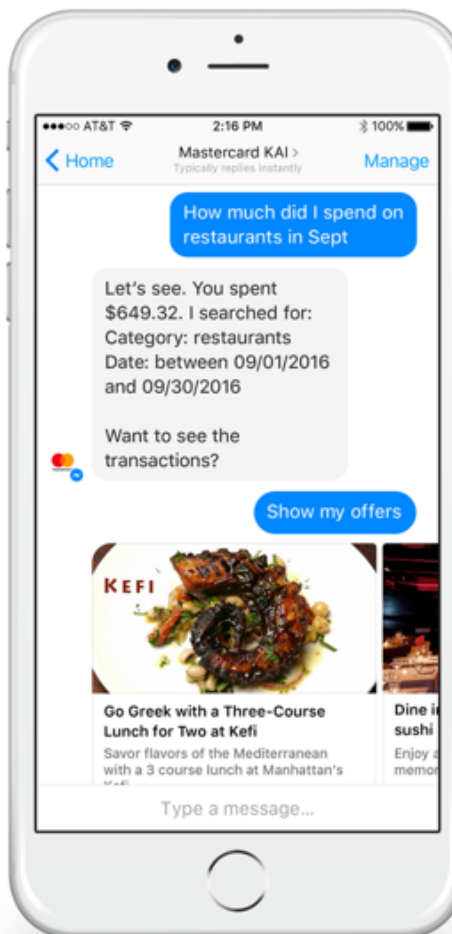
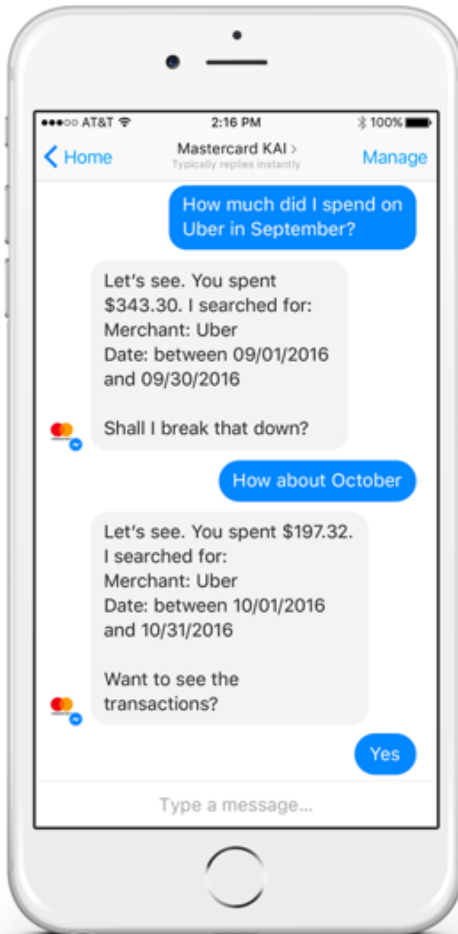
- Uber's chatbot on Facebook's messenger
- one main benefit: it loads much faster than the Uber app

Source: <http://www.guided-selling.org/from-e-commerce-to-conversational-commerce/>

Savings Bot



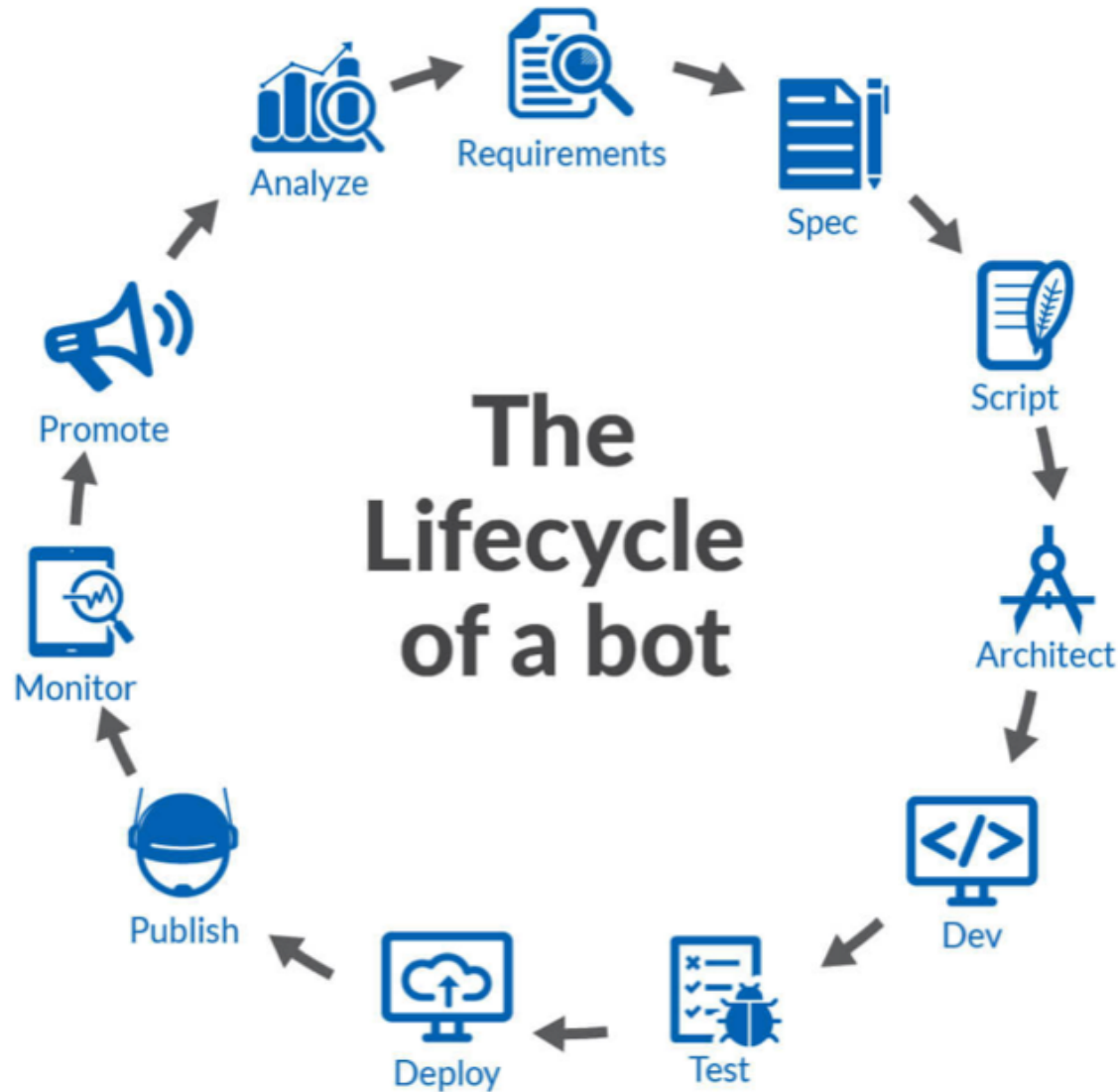
Mastercard Makes Commerce More Conversational



POWERED BY
Kasisto

Bot Life Cycle and Platform Ecosystem

The Bot Lifecycle



The bot platform ecosystem and the emerging giants

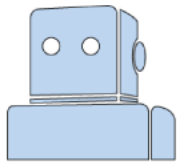
Nearly every large software company has announced some sort of bot strategy in the last year. Here's a look at a handful of leading platforms that developers might use to send messages, interpret natural language, and deploy bots, with the emerging bot-ecosystem giants highlighted.

General AI agents with platforms

Developer access available now or announced



Bot frameworks and deployment platforms



Wit.ai
Facebook



BotKit
Howdy



Chatfuel

AUTOMAT

Automat



Bot Framework
Microsoft



Api.ai
Google



Pandorabots



MindMeld



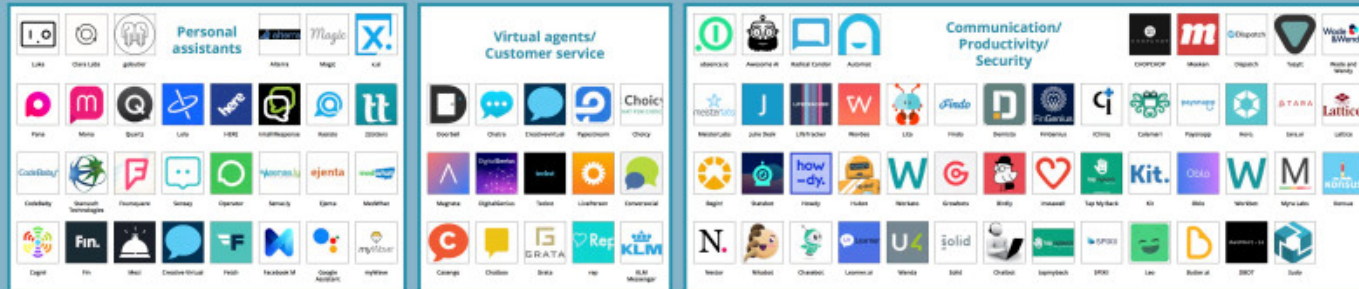
Gupshup



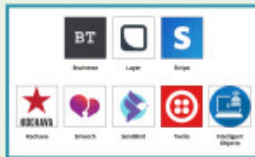
Sequel

Bots Landscape

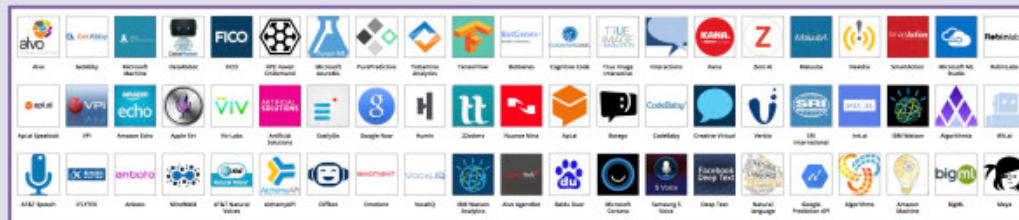
Bots with traction



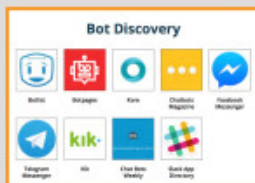
Connectors/ Shared Services



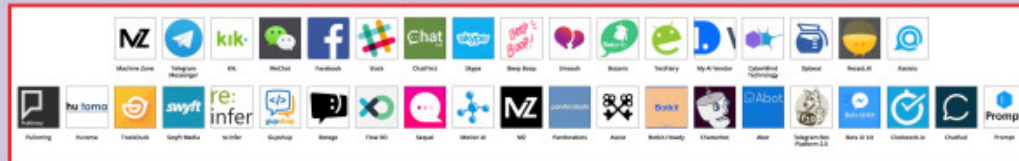
AI Tools: Natural Language Processing, Machine Learning, Speech & Voice Recognition



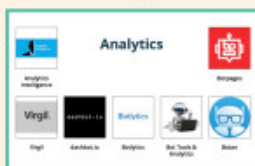
Bot Discovery



Bot developer frameworks and tools



Analytics



Messaging



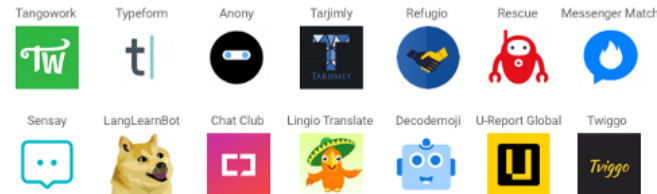
RECAST.AI Messenger Bot Landscape

May 2017

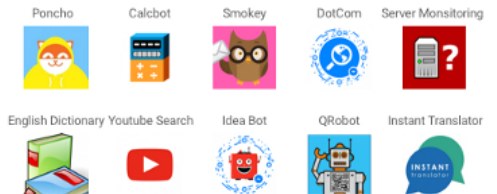
Food



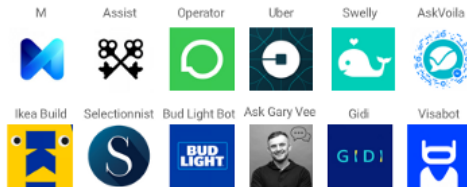
Communication



Utilities



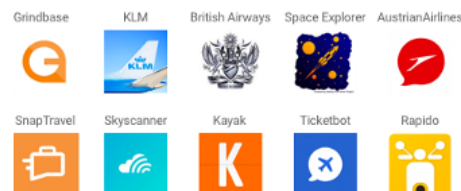
Personal



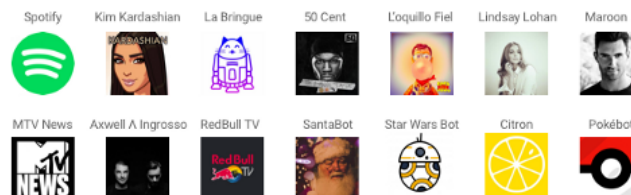
Analytics



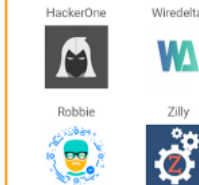
Travel



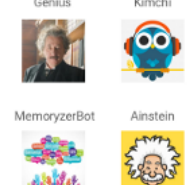
Entertainment



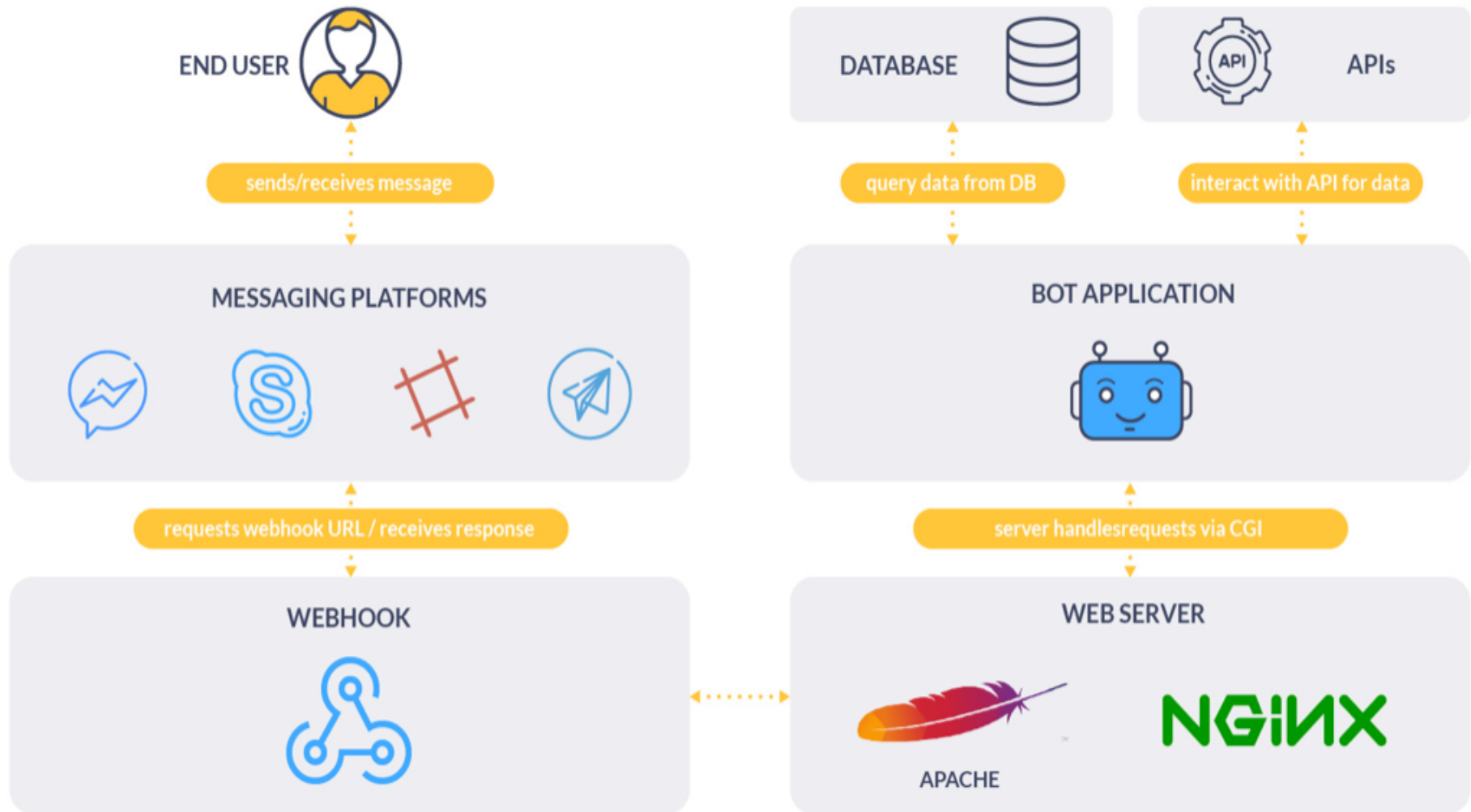
Developer Tools



Education



How to Build Chatbots



Chatbot Frameworks and AI Services

- Bot Frameworks
 - Botkit
 - Microsoft Bot Framework
 - Rasa NLU
- AI Services
 - Wit.ai
 - api.ai
 - LUIS.ai
 - IBM Watson

Chatbot Frameworks

Comparison Table of Most Prominent Bot Frameworks



Botkit



Microsoft Bot Framework



Built-In Integration with messaging platforms	✓	✓	✗
NLP support	✗ but possible to integrate with middlewares	✗ but have close bonds with LUIS.ai	✓
Out-of-box bots ready to be deployed	✓	✗	✗
Programming Language	JavaScript (Node)	JavaScript (Node), C#	Python

Created by  ActiveWizards

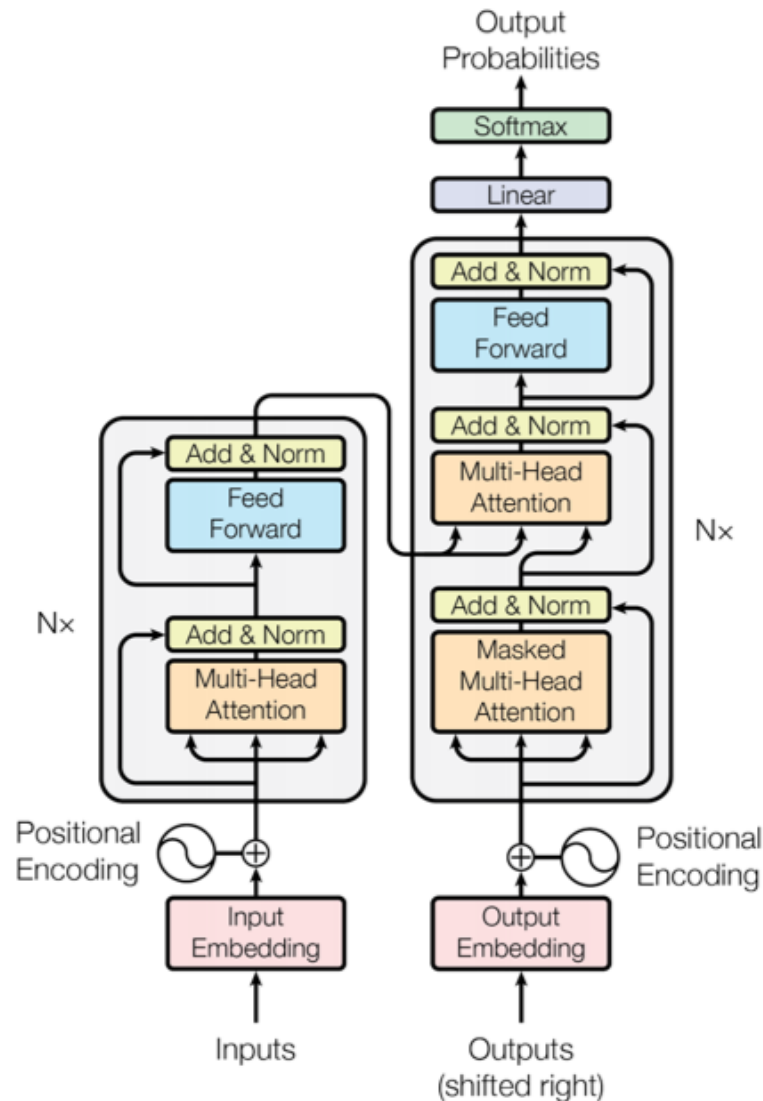
Comparison of Most Prominent AI Services

	wit.ai	api.ai	LUIS.ai	IBM Watson
Free of charge	✓	✓ but has paid enterprise version	✓ it is in beta and has transaction limits	30 days trial then priced for enterprise use
Text and Speech processing	✓	✓	✓ with use of Cortana	✓
Machine Learning Modeling	✓	✓	✓	✓
Support for Intents, Entities, Actions	✓ Intents used as trait entities, actions are combined operations	✓ Intents is the main prediction mechanism. Domains of entities, intents and actions	✓	✓
Pre-build entities for easy parsing of numbers, temperature, date, etc.	✓	✓	✓	✓
Integration to messaging platforms	✗ web service API	✓ also has facility for deploying to heroku. Paid environment	✓ integrated to Azure	✓ possible via API
Support of SDKs	✓ includes SDKs for Python, Node.js, Rust, C, Ruby, iOS, Android, Windows Phone	✓ C#, Xamarin, Python, Node.js, iOS, Android, Windows Phone	✓ enables building with Web Service API, Microsoft Bot Framework integration	Proprietary language "AlchemyLanguage"

Created by **ActiveWizards**

Transformer (Attention is All You Need)

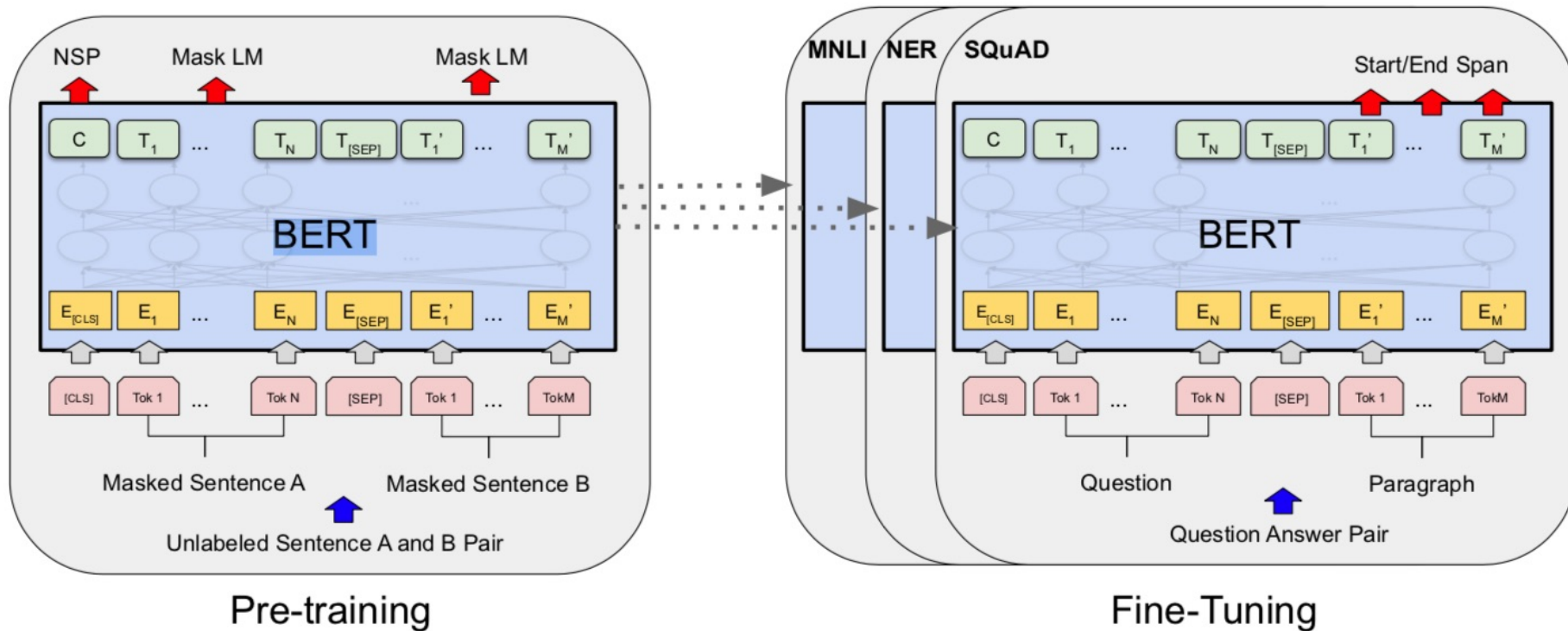
(Vaswani et al., 2017)



BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding

BERT (Bidirectional Encoder Representations from Transformers)

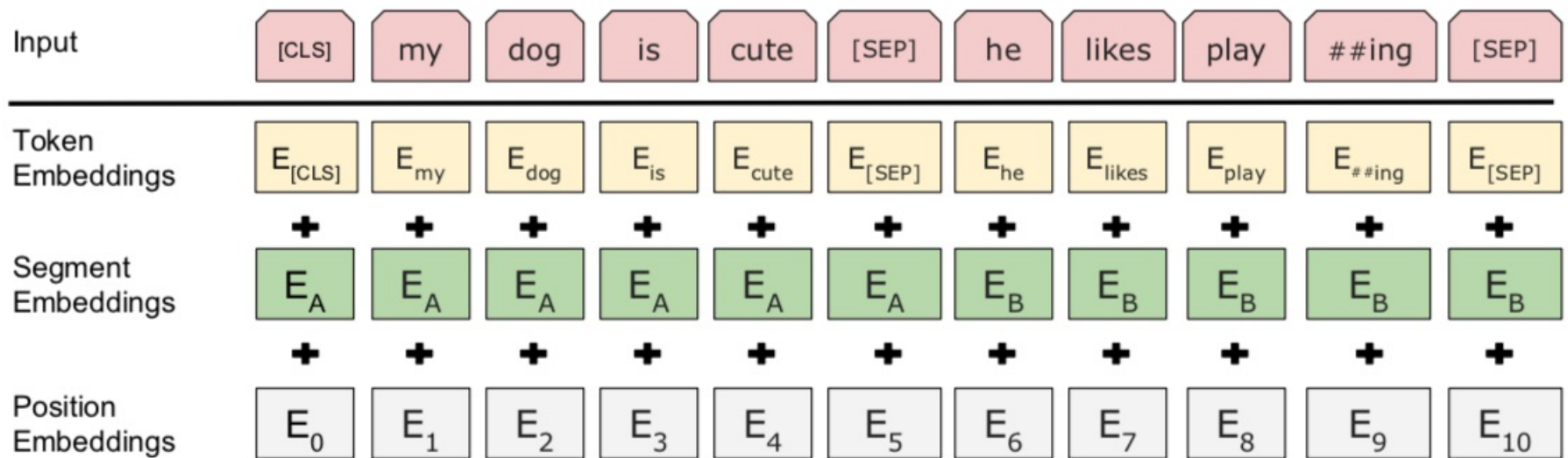
Overall pre-training and fine-tuning procedures for BERT



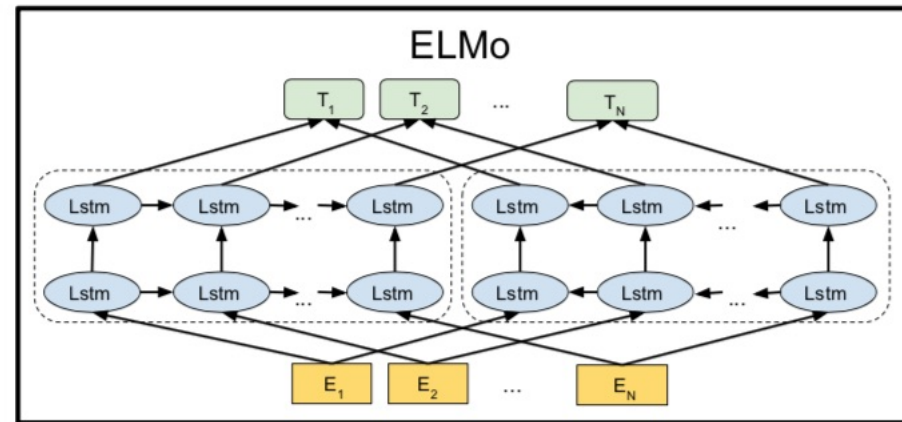
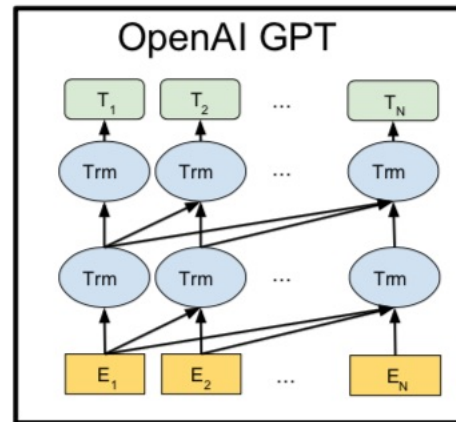
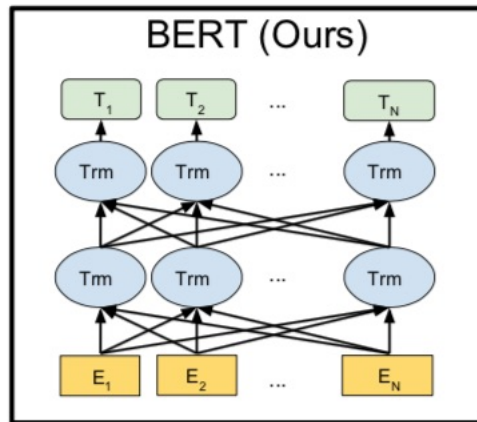
BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding

BERT (Bidirectional Encoder Representations from Transformers)

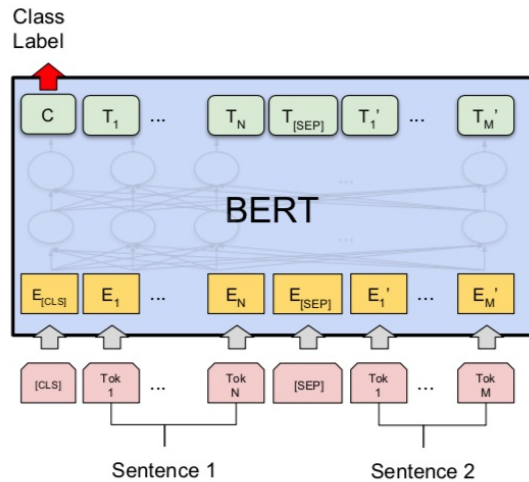
BERT input representation



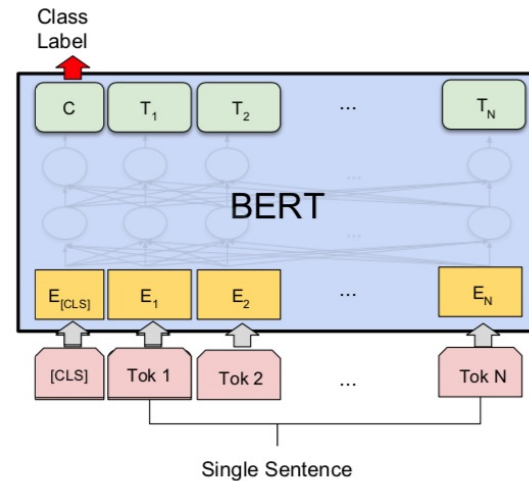
BERT, OpenAI GPT, ELMo



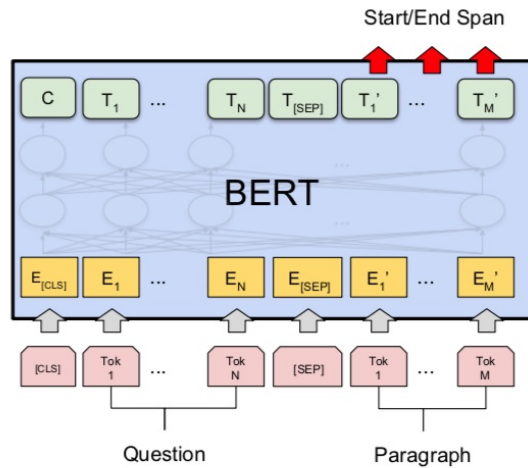
Fine-tuning BERT on Different Tasks



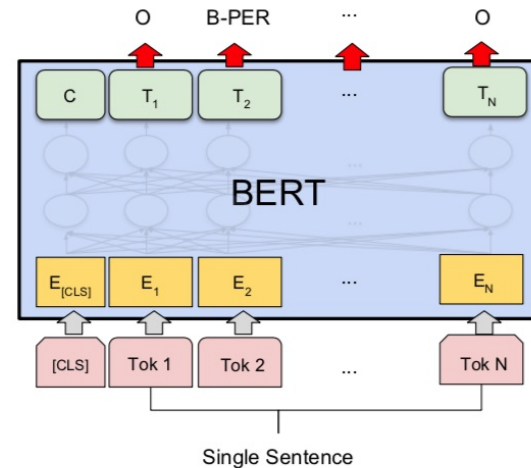
(a) Sentence Pair Classification Tasks:
MNLI, QQP, QNLI, STS-B, MRPC,
RTE, SWAG



(b) Single Sentence Classification Tasks:
SST-2, CoLA



(c) Question Answering Tasks:
SQuAD v1.1

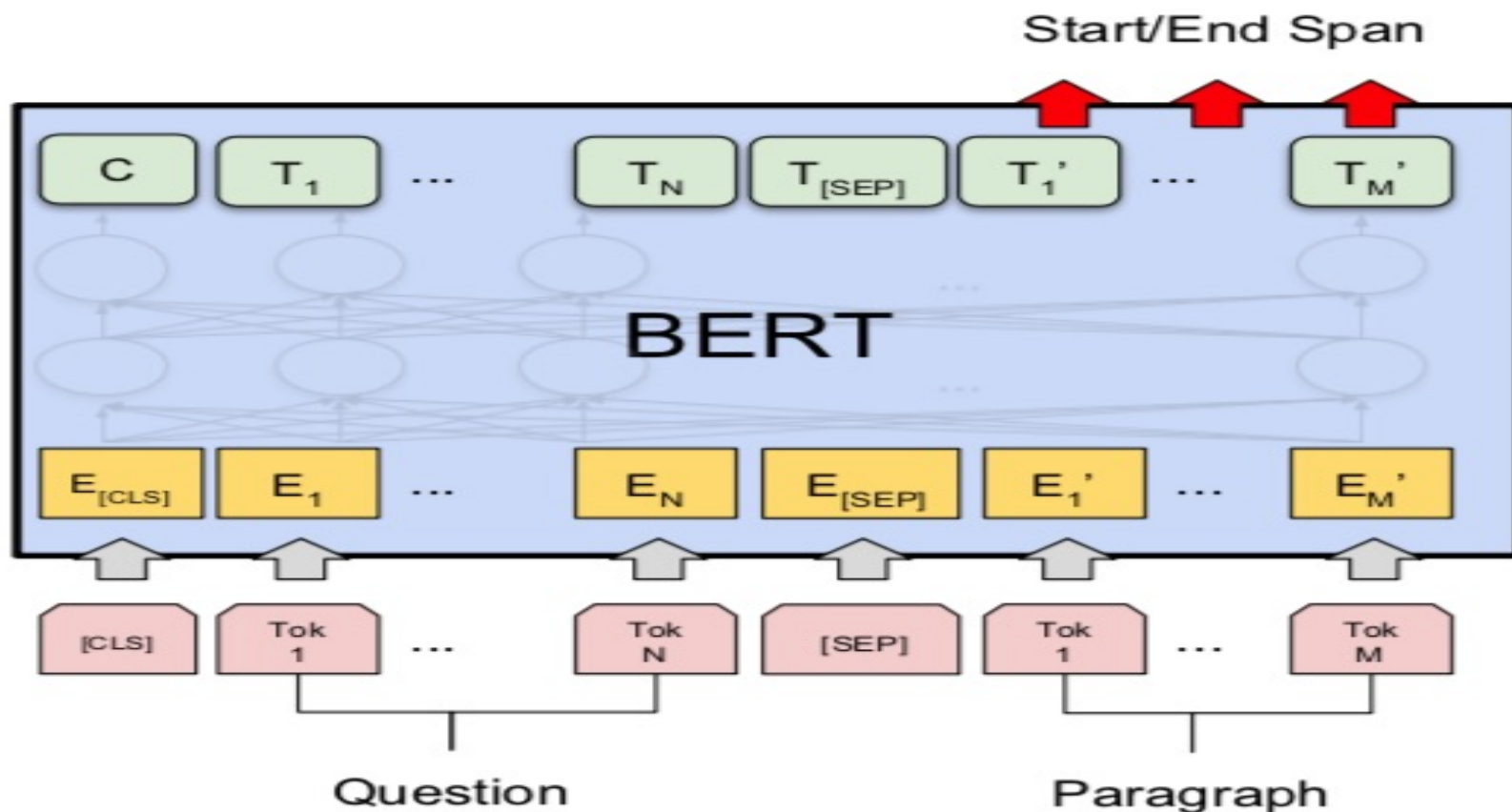


(d) Single Sentence Tagging Tasks:
CoNLL-2003 NER

Source: Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova (2018).

"Bert: Pre-training of deep bidirectional transformers for language understanding." arXiv preprint arXiv:1810.04805.

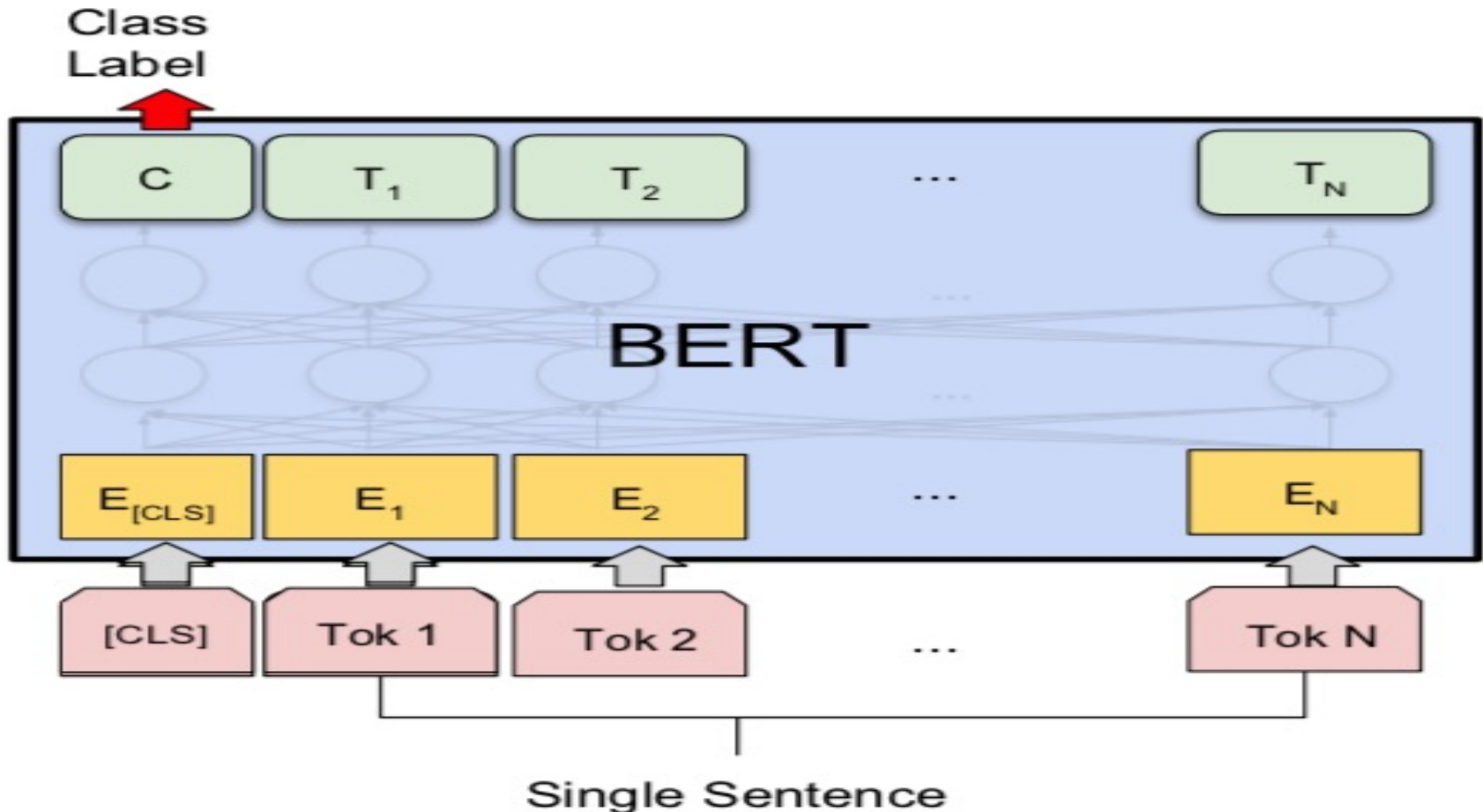
Fine-tuning BERT on Question Answering (QA)



(c) Question Answering Tasks:
SQuAD v1.1

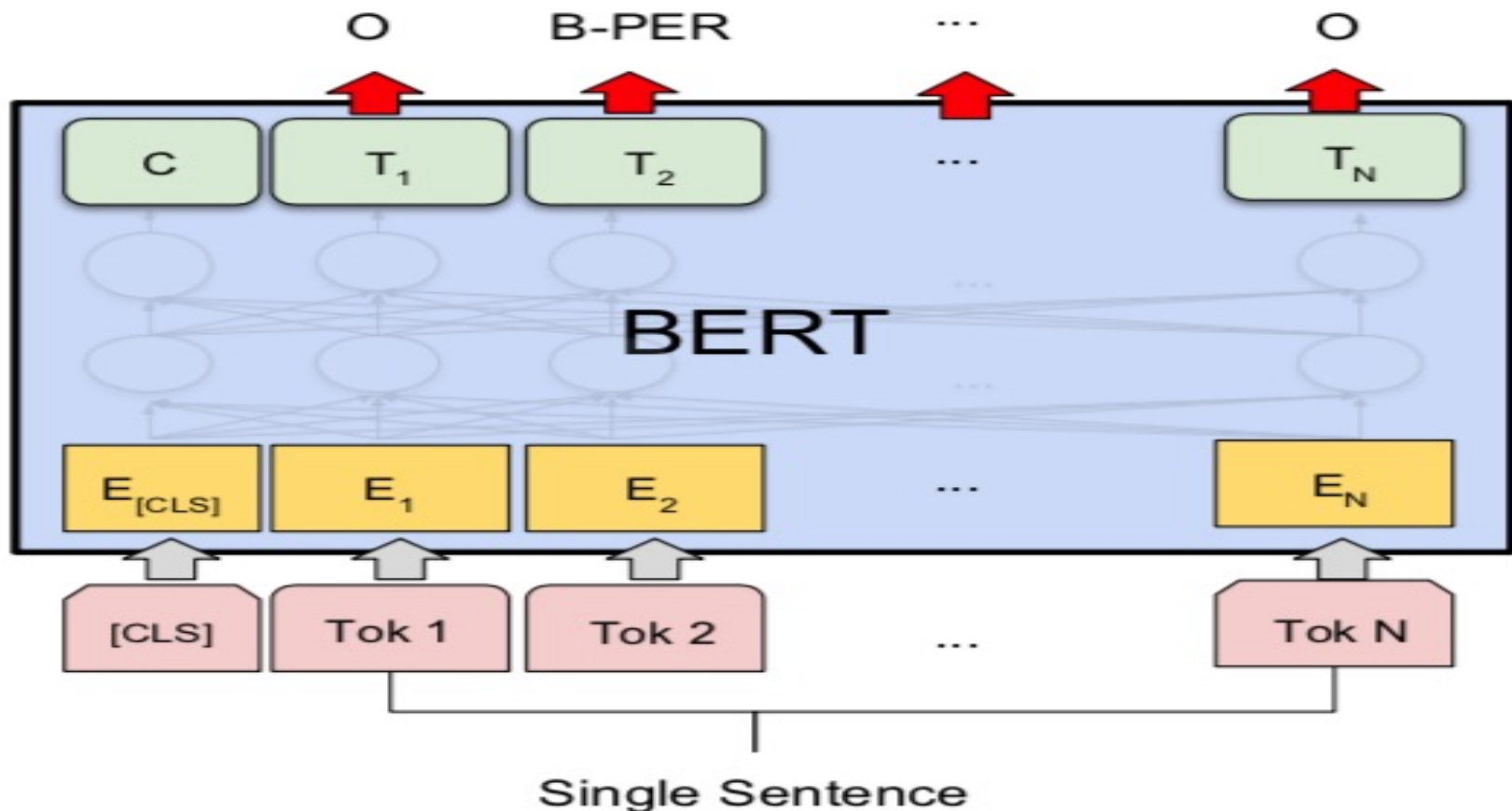
Fine-tuning BERT on Dialogue

Intent Detection (ID; Classification)



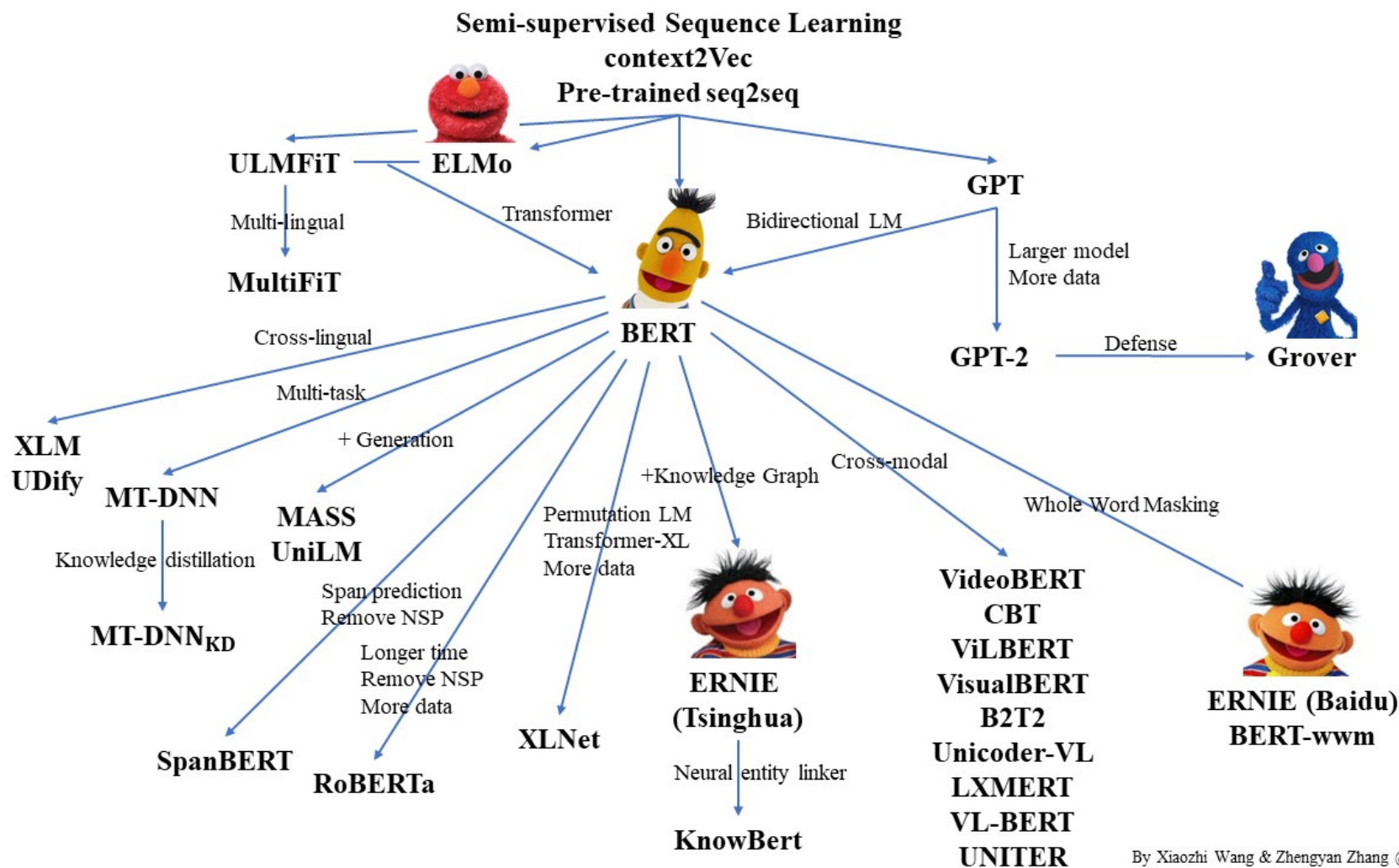
(b) Single Sentence Classification Tasks:
SST-2, CoLA

Fine-tuning BERT on Dialogue Slot Filling (SF)



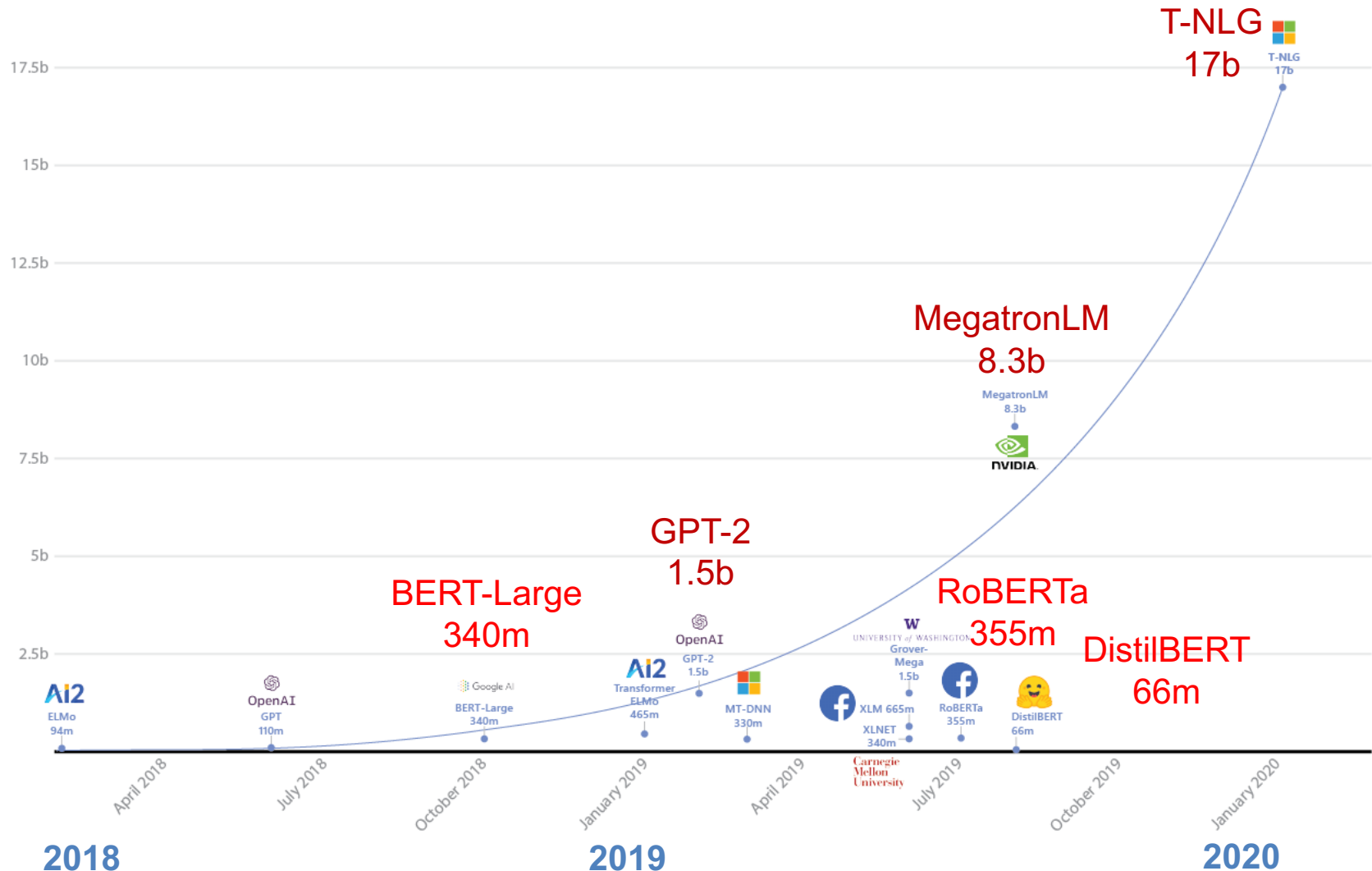
(d) Single Sentence Tagging Tasks:
CoNLL-2003 NER

Pre-trained Language Model (PLM)



By Xiaozhi Wang & Zhengyan Zhang @THUNLP

Turing Natural Language Generation (T-NLG)



Source: <https://www.microsoft.com/en-us/research/blog/turing-nlg-a-17-billion-parameter-language-model-by-microsoft/>

Transformers Transformers

State-of-the-art Natural Language Processing for TensorFlow 2.0 and PyTorch

- Transformers
 - pytorch-transformers
 - pytorch-pretrained-bert
- provides state-of-the-art general-purpose architectures
 - (BERT, GPT-2, RoBERTa, XLM, DistilBert, XLNet, CTRL...)
 - for Natural Language Understanding (NLU) and Natural Language Generation (NLG)
with over 32+ pretrained models
in 100+ languages
and deep interoperability between TensorFlow 2.0 and PyTorch.

Transfer Learning in Natural Language Processing

Source: Sebastian Ruder, Matthew E. Peters, Swabha Swayamdipta, and Thomas Wolf (2019), "Transfer learning in natural language processing." In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Tutorials, pp. 15-18.

NLP Benchmark Datasets

Task	Dataset	Link
Machine Translation	WMT 2014 EN-DE WMT 2014 EN-FR	http://www-lium.univ-lemans.fr/~schwenk/cs1m_joint_paper/
Text Summarization	CNN/DM Newsroom DUC Gigaword	https://cs.nyu.edu/~kcho/DMQA/ https://summari.es/ https://www-nlpir.nist.gov/projects/duc/data.html https://catalog.ldc.upenn.edu/LDC2012T21
Reading Comprehension Question Answering Question Generation	ARC CliCR CNN/DM NewsQA RACE SQuAD Story Cloze Test NarrativeQA Quasar SearchQA	http://data.allenai.org/arc/ http://aclweb.org/anthology/N18-1140 https://cs.nyu.edu/~kcho/DMQA/ https://datasets.maluuba.com/NewsQA http://www.qizhexie.com/data/RACE_leaderboard https://rajpurkar.github.io/SQuAD-explorer/ http://aclweb.org/anthology/W17-0906.pdf https://github.com/deepmind/narrativeqa https://github.com/bdhingra/quasar https://github.com/nyu-dl/SearchQA
Semantic Parsing	AMR parsing ATIS (SQL Parsing) WikiSQL (SQL Parsing)	https://amr.isi.edu/index.html https://github.com/jkkummerfeld/text2sql-data/tree/master/data https://github.com/salesforce/WikiSQL
Sentiment Analysis	IMDB Reviews SST Yelp Reviews Subjectivity Dataset	http://ai.stanford.edu/~amaas/data/sentiment/ https://nlp.stanford.edu/sentiment/index.html https://www.yelp.com/dataset/challenge http://www.cs.cornell.edu/people/pabo/movie-review-data/
Text Classification	AG News DBpedia TREC 20 NewsGroup	http://www.di.unipi.it/~gulli/AG_corpus_of_news_articles.html https://wiki.dbpedia.org/Datasets https://trec.nist.gov/data.html http://qwone.com/~jason/20Newsgroups/
Natural Language Inference	SNLI Corpus MultiNLI SciTail	https://nlp.stanford.edu/projects/snli/ https://www.nyu.edu/projects/bowman/multinli/ http://data.allenai.org/scitail/
Semantic Role Labeling	Proposition Bank OneNotes	http://propbank.github.io/ https://catalog.ldc.upenn.edu/LDC2013T19

Question Answering (QA) SQuAD

Stanford Question Answering Dataset

SQuAD2.0

The Stanford Question Answering Dataset

What is SQuAD?

Stanford **Q**uestion **A**nswering **D**ataset (SQuAD) is a reading comprehension dataset, consisting of questions posed by crowdworkers on a set of Wikipedia articles, where the answer to every question is a segment of text, or *span*, from the corresponding reading passage, or the question might be unanswerable.

SQuAD2.0 combines the 100,000 questions in SQuAD1.1 with over 50,000 unanswerable questions written adversarially by crowdworkers to look similar to answerable ones. To do well on SQuAD2.0, systems must not only answer questions when possible, but also determine when no answer is supported by the paragraph and abstain from answering.

Leaderboard

SQuAD2.0 tests the ability of a system to not only answer reading comprehension questions, but also abstain when presented with a question that cannot be answered based on the provided paragraph.

Rank	Model	EM	F1
	Human Performance Stanford University (Rajpurkar & Jia et al. '18)	86.831	89.452
1 Apr 06, 2020	SA-Net on Albert (ensemble) QIANXIN	90.724	93.011
2 May 05, 2020	SA-Net-V2 (ensemble) QIANXIN	90.679	92.948
?	Retro-Reader (ensemble)	90.578	92.978

SQuAD

SQuAD: 100,000+ Questions for Machine Comprehension of Text

Pranav Rajpurkar and **Jian Zhang** and **Konstantin Lopyrev** and **Percy Liang**

{pranavs, zjian, klopyrev, pliang}@cs.stanford.edu

Computer Science Department

Stanford University

Abstract

We present the Stanford Question Answering Dataset (SQuAD), a new reading comprehension dataset consisting of 100,000+ questions posed by crowdworkers on a set of Wikipedia articles, where the answer to each question is a segment of text from the corresponding reading passage. We analyze the dataset to understand the types of reasoning required to answer the questions, leaning heavily on dependency and constituency trees. We build a strong logistic regression model, which achieves an F1 score of 51.0%, a significant improvement over a simple baseline (20%). However, human performance (86.8%) is much higher, indicating that the dataset presents a good challenge problem for future research. The dataset is freely available at <https://stanford-qa.com>.

In meteorology, precipitation is any product of the condensation of atmospheric water vapor that falls under **gravity**. The main forms of precipitation include drizzle, rain, sleet, snow, **grau-pel** and hail... Precipitation forms as smaller droplets coalesce via collision with other rain drops or ice crystals **within a cloud**. Short, intense periods of rain in scattered locations are called "showers".

What causes precipitation to fall?

gravity

What is another main form of precipitation besides drizzle, rain, snow, sleet and hail?

grau-pel

Where do water droplets collide with ice crystals to form precipitation?

within a cloud

Figure 1: Question-answer pairs for a sample passage in the

Source: Rajpurkar, Pranav, Jian Zhang, Konstantin Lopyrev, and Percy Liang.

"Squad: 100,000+ questions for machine comprehension of text." arXiv preprint arXiv:1606.05250 (2016).

SQuAD (Question Answering)

Q: What causes precipitation to fall?

Precipitation

From Wikipedia, the free encyclopedia

For other uses, see [Precipitation \(disambiguation\)](#).

In meteorology, **precipitation** is any product of the condensation of atmospheric water vapor that falls under gravity from clouds.^[2] The main forms of precipitation include drizzle, rain, sleet, snow, ice pellets, graupel and hail. Precipitation occurs when a portion of the atmosphere becomes saturated with water vapor (reaching 100% [relative humidity](#)), so that the water condenses and "precipitates". Thus, fog and mist are not precipitation but suspensions, because the water vapor does not condense sufficiently to precipitate. Two processes, possibly acting together, can lead to air becoming saturated: cooling the air or adding water vapor to the air. Precipitation forms as smaller droplets coalesce via collision with other rain drops or ice crystals within a cloud. **Short, intense periods of rain in scattered locations are called "showers."**^[3]

SQuAD (Question Answering)

Paragraph

In meteorology, precipitation is any product of the condensation of atmospheric water vapor that falls under gravity. The main forms of precipitation include drizzle, rain, sleet, snow, graupel and hail... Precipitation forms as smaller droplets coalesce via collision with other rain drops or ice crystals within a cloud. Short, intense periods of rain in scattered locations are called “showers”.

Q: What causes precipitation to fall?

SQuAD (Question Answering)

In meteorology, precipitation is any product of the condensation of atmospheric water vapor that falls under gravity. The main forms of precipitation include drizzle, rain, sleet, snow, graupel and hail... Precipitation forms as smaller droplets coalesce via collision with other rain drops or ice crystals within a cloud. Short, intense periods of rain in scattered locations are called “showers”.

Q: What causes precipitation to fall?

A: gravity

SQuAD (Question Answering)

In meteorology, precipitation is any product of the condensation of atmospheric water vapor that falls under gravity. The main forms of precipitation include drizzle, rain, sleet, snow, graupel and hail... Precipitation forms as smaller droplets coalesce via collision with other rain drops or ice crystals within a cloud. Short, intense periods of rain in scattered locations are called “showers”.

Q: What is another main form of precipitation besides drizzle, rain, snow, sleet and hail?

A: graupel

SQuAD (Question Answering)

In meteorology, precipitation is any product of the condensation of atmospheric water vapor that falls under gravity. The main forms of precipitation include drizzle, rain, sleet, snow, graupel and hail... Precipitation forms as smaller droplets coalesce via collision with other rain drops or ice crystals within a cloud. Short, intense periods of rain in scattered locations are called “showers”.

Q: Where do water droplets collide with ice crystals to form precipitation?

A: within a cloud

SQuAD (Question Answering)

In meteorology, precipitation is any product of the condensation of atmospheric water vapor that falls under gravity. The main forms of precipitation include drizzle, rain, sleet, snow, graupel and hail... Precipitation forms as smaller droplets coalesce via collision with other rain drops or ice crystals within a cloud. Short, intense periods of rain in scattered locations are called “showers”.

Q: What causes precipitation to fall?

A: gravity

Q: What is another main form of precipitation besides drizzle, rain, snow, sleet and hail?

A: graupel

Q: Where do water droplets collide with ice crystals to form precipitation?

A: within a cloud

SQuAD (Question Answering)

Super Bowl 50

From Wikipedia, the free encyclopedia

"2016 Super Bowl" redirects here. For the Super Bowl that was played at the completion of the 2016 season, see [Super Bowl LI](#).

"SB 50" redirects here. For the California transit-density bill, see [California Senate Bill 50](#).

Super Bowl 50 was an [American football](#) game to determine the champion of the [National Football League](#) (NFL) for the [2015 season](#). The [American Football Conference](#) (AFC) champion [Denver Broncos](#) defeated the [National Football Conference](#) (NFC) champion [Carolina Panthers](#), 24–10. The game was played on February 7, 2016, at [Levi's Stadium](#) in [Santa Clara, California](#), in the [San Francisco Bay Area](#). As this was the 50th Super Bowl game, the league emphasized the "golden anniversary" with various gold-themed initiatives during the 2015 season, as well as suspending the tradition of naming each Super Bowl game with [Roman numerals](#) (under which the game would have been known as "Super Bowl L"), so the logo could prominently feature the [Arabic numerals](#) 5 and 0.^{[5][6]}

The Panthers finished the regular season with a 15–1 record, racking up the league's top offense, and quarterback [Cam Newton](#) was named the [NFL Most Valuable Player](#) (MVP). They defeated the [Arizona Cardinals](#) 49–15 in the [NFC Championship Game](#) and advanced to their second Super Bowl appearance since the franchise began playing in 1995. The Broncos finished the regular season with a 12–4 record, bolstered by having the league's top defense. The Broncos defeated the defending Super Bowl champion [New England Patriots](#) 20–18 in the [AFC Championship Game](#) joining the [Patriots](#), [Dallas Cowboys](#), and [Pittsburgh Steelers](#) as one of four teams that have made [eight appearances in the Super Bowl](#). This record would later be broken the next season, in 2017, when the Patriots advanced to their ninth Super Bowl appearance in [Super Bowl LI](#).

Super Bowl 50



**Dialogue
on
Airline Travel
Information System
(ATIS)**

The ATIS (Airline Travel Information System) Dataset

<https://www.kaggle.com/siddhadev/atis-dataset-from-ms-cntk>

Sentence	what	flights	leave	from	phoenix
Slots	O	O	O	O	B-fromloc
Intent	atis_flight				

Training samples: 4978

Testing samples: 893

Vocab size: 943

Slot count: 129

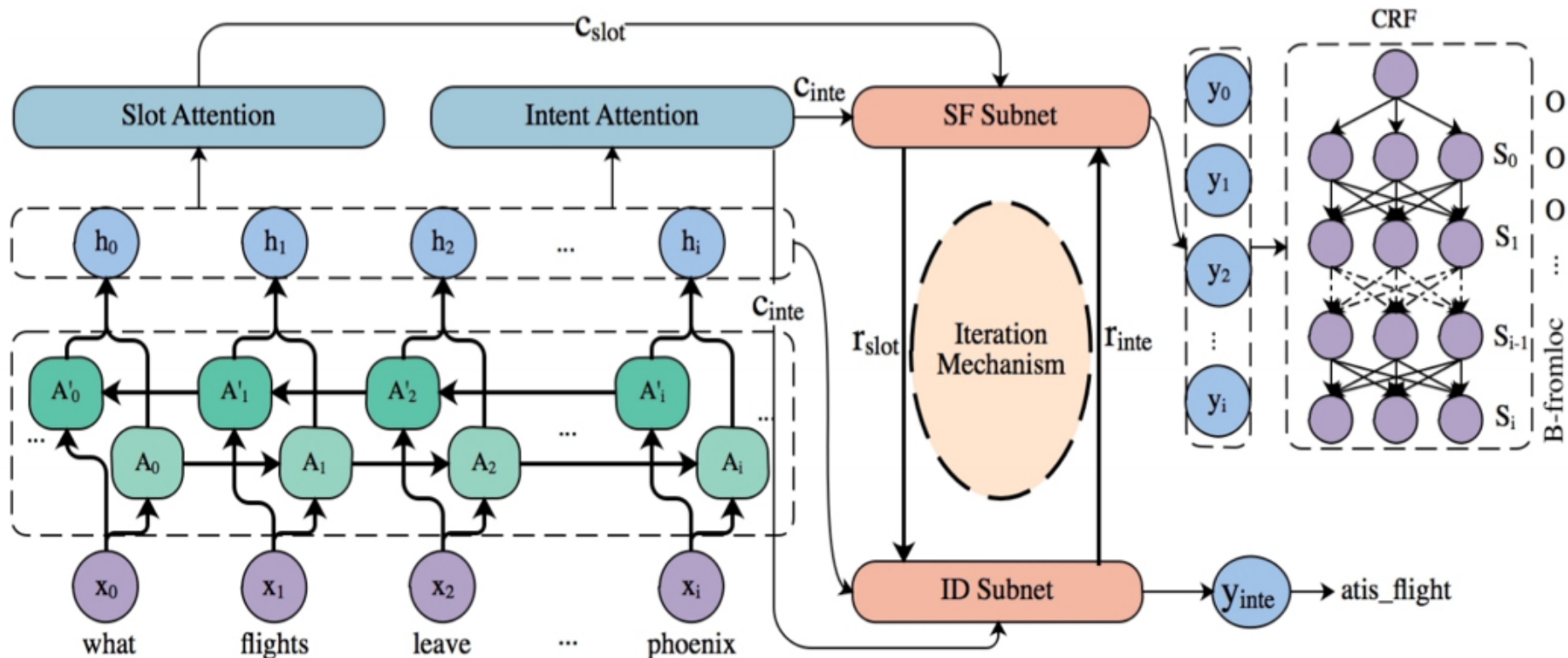
Intent count: 26

SF-ID Network (E et al., 2019)

Slot Filling (SF)

Intent Detection (ID)

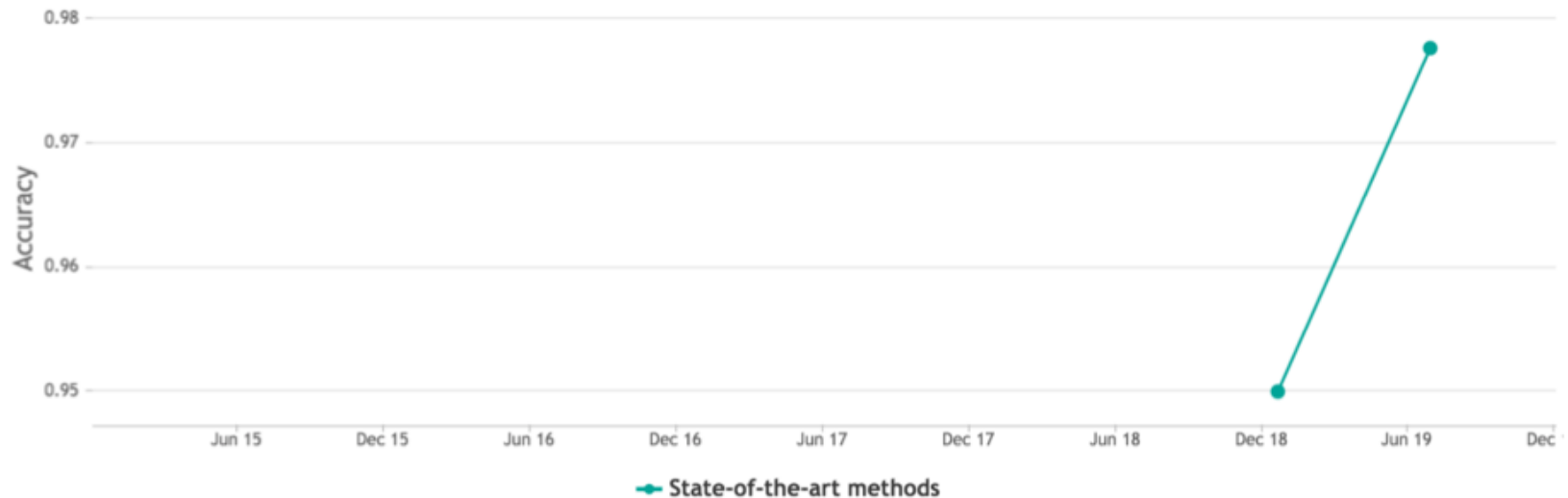
A Novel Bi-directional Interrelated Model for Joint Intent Detection and Slot Filling



Intent Detection on ATIS

State-of-the-art

Intent Detection on ATIS



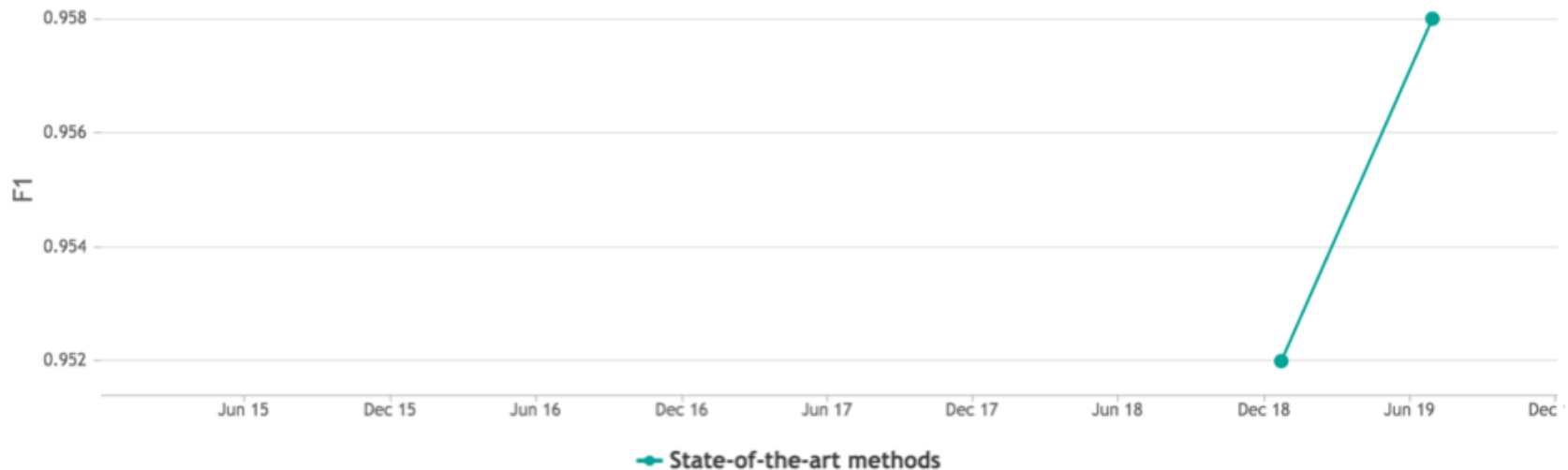
Edit





RANK	METHOD	ACCURACY	PAPER TITLE	YEAR	PAPER	CODE
1	SF-ID	0.9776	A Novel Bi-directional Interrelated Model for Joint Intent Detection and Slot Filling	2019		
2	Capsule-NLU	0.950	Joint Slot Filling and Intent Detection via Capsule Neural Networks	2018		

Slot Filling on ATIS

State-of-the-art

Slot Filling on ATIS



RANK	METHOD	F1	PAPER TITLE	YEAR	PAPER	CODE
1	SF-ID	0.958	A Novel Bi-directional Interrelated Model for Joint Intent Detection and Slot Filling	2019		
2	Capsule-NLU	0.952	Joint Slot Filling and Intent Detection via Capsule Neural Networks	2018		

Source: <https://paperswithcode.com/sota/slot-filling-on-atis>

Restaurants Dialogue Datasets

- MIT Restaurant Corpus
 - <https://groups.csail.mit.edu/sls/downloads/restaurant/>
- CamRest676
(Cambridge restaurant dialogue domain dataset)
 - <https://www.repository.cam.ac.uk/handle/1810/260970>
- DSTC2 (Dialog State Tracking Challenge 2 & 3)
 - <http://camdial.org/~mh521/dstc/>

CrossWOZ:

A Large-Scale Chinese Cross-Domain Task-Oriented Dialogue Dataset

MultiWOZ	CrossWOZ
usr: I'm looking for a college type attraction. ... usr: I would like to visit <u>in town centre</u> please. ... usr: Can you find an Indian restaurant for me that is also in the town centre ?	usr: 你好, 可以帮我推荐一个评分是4.5分以上的景点吗? Hello, could you recommend an attraction with a rating of 4.5 or higher? sys: 天安门城楼, 簋街小吃和 <u>北京欢乐谷</u> 都是很不错的地方呢。 Tiananmen, Gui Street, and <u>Beijing Happy Valley</u> are very nice places. usr: 我喜欢 <u>北京欢乐谷</u> , 你知道这个景点周边的酒店都是什么吗? I like <u>Beijing Happy Valley</u> . What hotels are around this attraction? sys: 那可多了, 有A酒店, B酒店, C酒店。 There are many, such as hotel A, hotel B, and hotel C. usr: 太好了, 我正打算在 <u>景点附近</u> 找个酒店住宿呢, 知道哪家评分是4分以上, 提供叫醒服务的? Great! I am planning to find a hotel to stay near the attraction . Which one has a rating of 4 or higher and offers wake-up call service?
Schema	
usr: I want a hotel in <u>San Diego</u> and I want to check out on Thursday next week. ... usr: I need a one way flight to go there .	

CrossWOZ:

A Large-Scale Chinese Cross-Domain Task-Oriented Dialogue Dataset

Type	Single-domain goal					Multi-domain goal		
Dataset	DSTC2	WOZ 2.0	Frames	KVRET	M2M	MultiWOZ	Schema	CrossWOZ
Language	EN	EN	EN	EN	EN	EN	EN	CN
Speakers	H2M	H2H	H2H	H2H	M2M	H2H	M2M	H2H
# Domains	1	1	1	3	2	7	16	5
# Dialogues	1,612	600	1,369	2,425	1,500	8,438	16,142	5,012
# Turns	23,354	4,472	19,986	12,732	14,796	115,424	329,964	84,692
Avg. domains	1	1	1	1	1	1.80	1.84	3.24
Avg. turns	14.5	7.5	14.6	5.3	9.9	13.7	20.4	16.9
# Slots	8	4	61	13	14	25	214	72
# Values	212	99	3,871	1363	138	4,510	14,139	7,871

Task-Oriented Dialogue

Initial user state (=user goal)

id=1(Attraction): fee=free,
name=?, nearby hotels=?

id=2(Hotel): **name=near (id=1)**,
wake-up call=yes, rating=?

id=3(Taxi): **from=(id=1), to=(id=2)**,
car type=? plate number=?

...

Final user state

id=1 (Attraction): name=Tiananmen Square,
fee=free, nearby hotels=[Beijing Capital
Hotel, Guidu Hotel Beijing]

id=2 (Hotel): **name=Beijing Capital Hotel**,
wake-up call=yes, rating=4.6

id=3 (Taxi): **from=Tiananmen Square**,
to=Beijing Capital Hotel,
car type=#CX, plate number=#CP



Source: Zhu, Qi, Kaili Huang, Zheng Zhang, Xiaoyan Zhu, and Minlie Huang. "Crosswoz: A large-scale chinese cross-domain task-oriented dialogue dataset." arXiv preprint arXiv:2002.11893 (2020).

任務型對話系統

The Evaluation of Chinese Human-Computer Dialogue Technology, SMP2019-ECDT

- 自然語言理解
Natural Language Understanding (NLU)
- 對話管理
Dialog Management (DM)
- 自然語言生成
Natural Language Generation (NLG)

Python in Google Colab (Python101)

<https://colab.research.google.com/drive/1FEG6DnGvwfUbeo4zJ1zTunjMqf2RkCrT>

The screenshot shows a Google Colab notebook titled "python101.ipynb". The top bar includes the Colab logo, the notebook name, a star icon, and a menu with options: File, Edit, View, Insert, Runtime, Tools, Help, and "All changes saved". On the right, there are icons for Comment, Share, a settings gear, and a user profile icon labeled 'A'. Below the top bar, a "Table of contents" sidebar is open on the left, listing various topics under "Semantic Analysis", "Sentiment Analysis", "Deep Learning and Universal Sentence-Embedding Models", and "Question Answering and Dialogue Systems". The "Question Answering and Dialogue Systems" section is highlighted. The main area of the notebook shows a code cell with the following content:

```
+ Code + Text
```

Question Answering and Dialogue Systems

Question Answering (QA)

BERT for Question Answering

Source: Apoorv Nandan (2020), BERT (from HuggingFace Transformers) for Text Extraction, https://keras.io/examples/nlp/text_extraction_with_bert/

Description: Fine tune pretrained BERT from HuggingFace Transformers on SQuAD.

Introduction

This demonstration uses SQuAD (Stanford Question-Answering Dataset). In SQuAD, an input consists of a question, and a paragraph for context. The goal is to find the span of text in the paragraph that answers the question. We evaluate our performance on this data with the "Exact Match" metric, which measures the percentage of predictions that exactly match any one of the ground-truth answers.

We fine-tune a BERT model to perform this task as follows:

1. Feed the context and the question as inputs to BERT.
2. Take two vectors S and T with dimensions equal to that of hidden states in BERT.
3. Compute the probability of each token being the start and end of the answer span. The probability of a token being the start of the answer is given by a dot product between S and the representation of the token in the last layer of BERT, followed by a softmax over all tokens. The probability of a token being the end of the answer is computed similarly with the vector T.
4. Fine-tune BERT and learn S and T along the way.


References:

- [BERT](#)
- [SQuAD](#)

<https://tinyurl.com/imtkupython101>

Python in Google Colab (Python101)

<https://colab.research.google.com/drive/1FEG6DnGvwfUbeo4zJ1zTunjMqf2RkCrT>

 python101.ipynb ☆

File Edit View Insert Runtime Tools Help All changes saved

Comment Share ⚙️ A

Table of contents ×

RandomizedSearchCV

Sentiment Analysis

Sentiment Analysis - Unsupervised Lexical

Sentiment Analysis - Supervised Machine Learning

Sentiment Analysis - Supervised Deep Learning Models

Sentiment Analysis - Advanced Deep Learning

Deep Learning and Universal Sentence-Embedding Models

Universal Sentence Encoder (USE)

Universal Sentence Encoder Multilingual (USEM)

Question Answering and Dialogue Systems

Question Answering (QA)

BERT for Question Answering

Dialogue Systems

Joint Intent Classification and Slot Filling with Transformers

Data Visualization

Section

+ Code + Text

✓ RAM Disk Editing ^

Downloading: 100% 433/433 [00:29<00:00, 14.5B/s]

Downloading: 100% 536M/536M [00:29<00:00, 18.3MB/s]

Model: "model"

Layer (type)	Output Shape	Param #	Connected to
input_1 (InputLayer)	[(None, 384)]	0	
input_3 (InputLayer)	[(None, 384)]	0	
input_2 (InputLayer)	[(None, 384)]	0	
tf_bert_model (TFBertModel)	((None, 384, 768), (109482240		input_1[0][0]
start_logits (Dense)	(None, 384, 1)	768	tf_bert_model[0][0]
end_logits (Dense)	(None, 384, 1)	768	tf_bert_model[0][0]
flatten (Flatten)	(None, 384)	0	start_logits[0][0]
flatten_1 (Flatten)	(None, 384)	0	end_logits[0][0]
activation_7 (Activation)	(None, 384)	0	flatten[0][0]
activation_8 (Activation)	(None, 384)	0	flatten_1[0][0]


Total params: 109,483,776
Trainable params: 109,483,776
Non-trainable params: 0

CPU times: user 20.8 s, sys: 7.75 s, total: 28.5 s
Wall time: 1min 42s

<https://tinyurl.com/imtkupython101>

Python in Google Colab (Python101)

<https://colab.research.google.com/drive/1FEG6DnGvwfUbeo4zJ1zTunjMqf2RkCrT>

 python101.ipynb ☆

File Edit View Insert Runtime Tools Help All changes saved

Comment Share ⚙️ A



RAM  Disk  Editing ^

Table of contents ✕

- RandomizedSearchCV
- Sentiment Analysis
 - Sentiment Analysis - Unsupervised Lexical
 - Sentiment Analysis - Supervised Machine Learning
 - Sentiment Analysis - Supervised Deep Learning Models
 - Sentiment Analysis - Advanced Deep Learning
- Deep Learning and Universal Sentence-Embedding Models
 - Universal Sentence Encoder (USE)
 - Universal Sentence Encoder Multilingual (USEM)
- Question Answering and Dialogue Systems
 - Question Answering (QA)
 - BERT for Question Answering
 - Dialogue Systems
 - Joint Intent Classification and Slot Filling with Transformers**
- Data Visualization
- + Section

+ Code + Text

▼ Dialogue Systems

```
[ ] 1 #Source: Olivier Grisel (2020), Transformers (BERT fine-tuning): Joint Intent Classification and S
    2 #https://github.com/m2dsupsdclass/lectures-labs/blob/master/labs/06_deep_nlp/Transformers_Joint_I
```

↑ ↓ ↻ 🗨️ ✎ 🗑️ ⋮

▼ Joint Intent Classification and Slot Filling with Transformers

The goal of this notebook is to fine-tune a pretrained transformer-based neural network model to convert a user query expressed in English into a representation that is structured enough to be processed by an automated service.

Here is an example of interpretation computed by such a Natural Language Understanding system:

```
>>> nlu("Book a table for two at Le Ritz for Friday night",
        tokenizer, joint_model, intent_names, slot_names)
```

```
{
  'intent': 'BookRestaurant',
  'slots': {
    'party_size_number': 'two',
    'restaurant_name': 'Le Ritz',
    'timeRange': 'Friday night'
  }
}
```

Intent classification is a simple sequence classification problem. The trick is to treat the structured knowledge extraction part ("Slot Filling") as token-level classification problem using BIO-annotations:

<https://tinyurl.com/imtkupython101>

Python in Google Colab (Python101)

<https://colab.research.google.com/drive/1FEG6DnGvwfUbeo4zJ1zTunjMqf2RkCrT>

The screenshot shows a Google Colab notebook titled "python101.ipynb". The left sidebar contains a "Table of contents" with the following items:

- RandomizedSearchCV
- Sentiment Analysis
 - Sentiment Analysis - Unsupervised Lexical
 - Sentiment Analysis - Supervised Machine Learning
 - Sentiment Analysis - Supervised Deep Learning Models
 - Sentiment Analysis - Advanced Deep Learning
- Deep Learning and Universal Sentence-Embedding Models
 - Universal Sentence Encoder (USE)
 - Universal Sentence Encoder Multilingual (USEM)
- Question Answering and Dialogue Systems
 - Question Answering (QA)
 - BERT for Question Answering
 - Dialogue Systems
 - Joint Intent Classification and Slot Filling with Transformers**
- Data Visualization

The main code cell contains the following Python code:

```
1 def show_predictions(text, tokenizer, model, intent_names, slot_names):
2     inputs = tf.constant(tokenizer.encode(text))[None, :] # batch_size = 1
3     outputs = model(inputs)
4     slot_logits, intent_logits = outputs
5     slot_ids = slot_logits.numpy().argmax(axis=-1)[0, 1:-1]
6     intent_id = intent_logits.numpy().argmax(axis=-1)[0]
7     print("Text:", text)
8     print("Intent:", intent_names[intent_id])
9     print("Slots:")
10    for token, slot_id in zip(tokenizer.tokenize(text), slot_ids):
11        print(f"{token}>10} : {slot_names[slot_id]}")
12
13 show_predictions("Book a table for two at Le Ritz for Friday night!",
14                 tokenizer, joint_model, intent_names, slot_names)
```


The output of the code is:

```
Text: Book a table for two at Le Ritz for Friday night!
Intent: BookRestaurant
Slots:
  Book : 0
    a : 0
  table : 0
    for : 0
    two : B-party_size_number
    at : 0
    Le : B-restaurant_name
    R : I-restaurant_name
    ##itz : I-restaurant_name
    for : 0
  Friday : B-timeRange
    night : 0
    ! : 0
```

<https://tinyurl.com/imtkupython101>

Python in Google Colab (Python101)

<https://colab.research.google.com/drive/1FEG6DnGvwfUbeo4zJ1zTunjMqf2RkCrT>

 python101.ipynb ☆

File Edit View Insert Runtime Tools Help All changes saved

Comment Share Settings A

RAM Disk Editing

Table of contents

- NER with CRF
- NER with CRF
- RandomizedSearchCV
- Sentiment Analysis
 - Sentiment Analysis - Unsupervised Lexical
 - Sentiment Analysis - Supervised Machine Learning
 - Sentiment Analysis - Supervised Deep Learning Models
 - Sentiment Analysis - Advanced Deep Learning
- Deep Learning and Universal Sentence-Embedding Models
 - Universal Sentence Encoder (USE)
 - Universal Sentence Encoder Multilingual (USEM)
- Question Answering and Dialogue Systems
 - Question Answering (QA)
 - BERT for Question Answering
 - Dialogue Systems
 - Joint Intent Classification and Slot Filling with Transformers**
- Data Visualization

```
19 # Naive BIOES-style handling: treat B- and I- the same...
20 new_slot_name = current_word_slot_name[2:]
21 if active_slot_name is None:
22     active_slot_words.append(word)
23     active_slot_name = new_slot_name
24 elif new_slot_name == active_slot_name:
25     active_slot_words.append(word)
26 else:
27     collected_slots[active_slot_name] = " ".join(active_slot_words)
28     active_slot_words = [word]
29     active_slot_name = new_slot_name
30 if active_slot_name:
31     collected_slots[active_slot_name] = " ".join(active_slot_words)
32 info["slots"] = collected_slots
33 return info
34
35 def nlu(text, tokenizer, model, intent_names, slot_names):
36     inputs = tf.constant(tokenizer.encode(text))[None, :] # batch_size = 1
37     outputs = model(inputs)
38     slot_logits, intent_logits = outputs
39     slot_ids = slot_logits.numpy().argmax(axis=-1)[0, 1:-1]
40     intent_id = intent_logits.numpy().argmax(axis=-1)[0]
41
42     return decode_predictions(text, tokenizer, intent_names, slot_names,
43                               intent_id, slot_ids)
44
45 nlu("Book a table for two at Le Ritz for Friday night",
46     tokenizer, joint_model, intent_names, slot_names)
```

```
{'intent': 'BookRestaurant',
 'slots': {'party_size_number': 'two',
           'restaurant_name': 'Le Ritz',
           'timeRange': 'Friday night'}}
```

<https://tinyurl.com/imtkupython101>

NLP Benchmark Datasets

Task	Dataset	Link
Machine Translation	WMT 2014 EN-DE WMT 2014 EN-FR	http://www-lium.univ-lemans.fr/~schwenk/csml_joint_paper/
Text Summarization	CNN/DM Newsroom DUC Gigaword	https://cs.nyu.edu/~kcho/DMQA/ https://summari.es/ https://www-nlpir.nist.gov/projects/duc/data.html https://catalog.ldc.upenn.edu/LDC2012T21
Reading Comprehension Question Answering Question Generation	ARC CliCR CNN/DM NewsQA RACE SQuAD Story Cloze Test NarrativeQA Quasar SearchQA	http://data.allenai.org/arc/ http://aclweb.org/anthology/N18-1140 https://cs.nyu.edu/~kcho/DMQA/ https://datasets.maluuba.com/NewsQA http://www.qizhexie.com/data/RACE_leaderboard https://rajpurkar.github.io/SQuAD-explorer/ http://aclweb.org/anthology/W17-0906.pdf https://github.com/deepmind/narrativeqa https://github.com/bdhingra/quasar https://github.com/nyu-dl/SearchQA
Semantic Parsing	AMR parsing ATIS (SQL Parsing) WikiSQL (SQL Parsing)	https://amr.isi.edu/index.html https://github.com/jkkummerfeld/text2sql-data/tree/master/data https://github.com/salesforce/WikiSQL
Sentiment Analysis	IMDB Reviews SST Yelp Reviews Subjectivity Dataset	http://ai.stanford.edu/~amaas/data/sentiment/ https://nlp.stanford.edu/sentiment/index.html https://www.yelp.com/dataset/challenge http://www.cs.cornell.edu/people/pabo/movie-review-data/
Text Classification	AG News DBpedia TREC 20 NewsGroup	http://www.di.unipi.it/~gulli/AG_corpus_of_news_articles.html https://wiki.dbpedia.org/Datasets https://trec.nist.gov/data.html http://qwone.com/~jason/20Newsgroups/
Natural Language Inference	SNLI Corpus MultiNLI SciTail	https://nlp.stanford.edu/projects/snli/ https://www.nyu.edu/projects/bowman/multinli/ http://data.allenai.org/scitail/
Semantic Role Labeling	Proposition Bank OneNotes	http://propbank.github.io/ https://catalog.ldc.upenn.edu/LDC2013T19

Summary

- Question Answering
- Dialogue Systems

References

- Day, Min-Yuh and Chi-Sheng Hung, "AI Affective Conversational Robot with Hybrid Generative-based and Retrieval-based Dialogue Models", in Proceedings of The 20th IEEE International Conference on Information Reuse and Integration for Data Science (IEEE IRI 2019), Los Angeles, CA, USA, July 30 - August 1, 2019.
- Day, Min-Yuh, Chi-Sheng Hung, Yi-Jun Xie, Jhih-Yi Chen, Yu-Ling Kuo and Jian-Ting Lin (2019), "IMTKU Emotional Dialogue System for Short Text Conversation at NTCIR-14 STC-3 (CECG) Task", The 14th NTCIR Conference on Evaluation of Information Access Technologies (NTCIR-14), Tokyo, Japan, June 10-13, 2019.
- Zhou, Hao, Minlie Huang, Tianyang Zhang, Xiaoyan Zhu, and Bing Liu. "Emotional chatting machine: emotional conversation generation with internal and external memory." arXiv preprint arXiv:1704.01074 (2017).
- Yu, Kai, Zijian Zhao, Xueyang Wu, Hongtao Lin, and Xuan Liu. "Rich Short Text Conversation Using Semantic Key Controlled Sequence Generation." IEEE/ACM Transactions on Audio, Speech, and Language Processing (2018).
- Borah, Bhiguraj, Dhrubajyoti Pathak, Priyankoo Sarmah, Bidisha Som, and Sukumar Nandi. "Survey of Textbased Chatbot in Perspective of Recent Technologies." In International Conference on Computational Intelligence, Communications, and Business Analytics, pp. 84-96. Springer, Singapore, 2018.
- Haihong, E., Peiqing Niu, Zhongfu Chen, and Meina Song. "A novel bi-directional interrelated model for joint intent detection and slot filling." In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pp. 5467-5471. 2019.
- Rajpurkar, Pranav, Jian Zhang, Konstantin Lopyrev, and Percy Liang. "Squad: 100,000+ questions for machine comprehension of text." arXiv preprint arXiv:1606.05250 (2016).
- Zhu, Qi, Kaili Huang, Zheng Zhang, Xiaoyan Zhu, and Minlie Huang. "Crosswoz: A large-scale chinese cross-domain task-oriented dialogue dataset." arXiv preprint arXiv:2002.11893 (2020).
- Apoorv Nandan (2020), BERT (from HuggingFace Transformers) for Text Extraction, https://keras.io/examples/nlp/text_extraction_with_bert/
- Olivier Grisel (2020), Transformers (BERT fine-tuning): Joint Intent Classification and Slot Filling, <https://m2dsupsdclass.github.io/lectures-labs/>
- Dipanjan Sarkar (2019), Text Analytics with Python: A Practitioner's Guide to Natural Language Processing, Second Edition. APress. <https://github.com/Apress/text-analytics-w-python-2e>
- Benjamin Bengfort, Rebecca Bilbro, and Tony Ojeda (2018), Applied Text Analysis with Python, O'Reilly Media. <https://www.oreilly.com/library/view/applied-text-analysis/9781491963036/>
- HuggingFace (2020), Transformers Notebook, <https://huggingface.co/transformers/notebooks.html>
- The Super Duper NLP Repo, <https://notebooks.quantumstat.com/>
- Min-Yuh Day (2020), Python 101, <https://tinyurl.com/imtkupython101>



Q & A

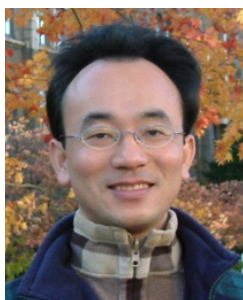


問答系統與對話系統 (Question Answering and Dialogue Systems)

Time: 2020/06/19 (Fri) (9:10 -12:00)

Place: 國立臺北護理健康大學 (台北市明德路365號) G210

Host: 祝國忠 院長 (健康科技學院院長)



Min-Yuh Day

戴敏育

Associate Professor

副教授

Dept. of Information Management, Tamkang University

淡江大學 資訊管理學系

<http://mail.tku.edu.tw/myday/>

2020-06-19

