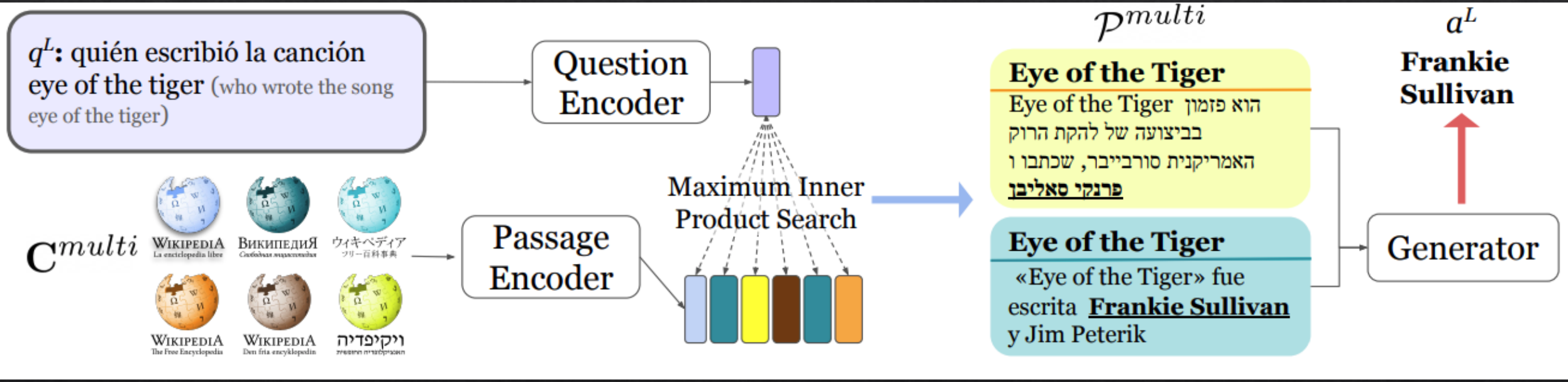
Multilingual Question Answering

601.764

2/23/23

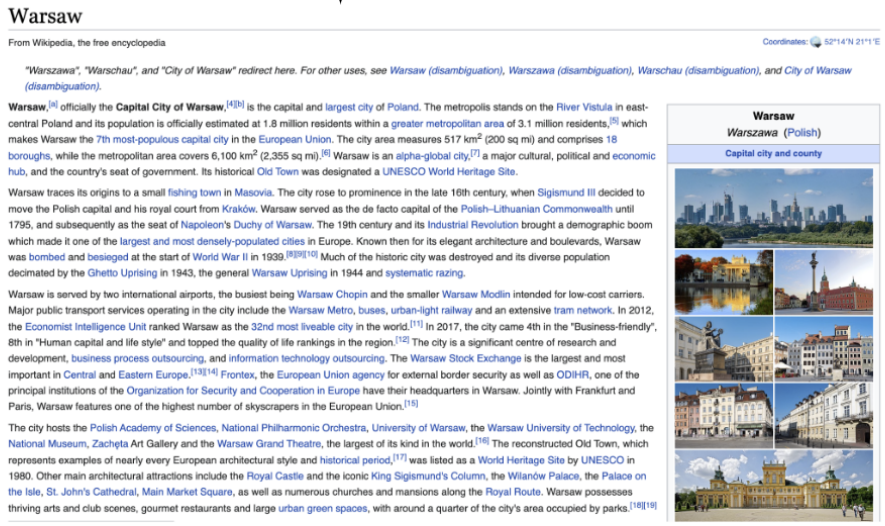
Much of this lecture is inspired by (and slides copied from):
EMNLP 2021 tutorial on Multi-domain Multilingual Question Answering.


q^L: quién escribió la canción eye of the tiger (who wrote the song eye of the tiger)
Question Encoder
C^{multi}
Passage Encoder
Maximum Inner Product Search
$\mathcal{P}^{multi}$
Eye of the Tiger
Eye of the Tiger הוא פזמון בביצועה של להקת הרוק האמריקנית סורבייבר, שכתבו ו פרנקי סאליבן
Eye of the Tiger
«Eye of the Tiger» fue escrita Frankie Sullivan y Jim Peterik
Generator
a^L
Frankie Sullivan

What is QA? Reading Comprehension vs Open-Retrieval QA

Document Retriever

Q: How many of Warsaw's inhabitants spoke Polish in 1933?



Warsaw

From Wikipedia, the free encyclopedia

Coordinates: 52°14′N 21°1′E

"Warszawa", "Warschau", and "City of Warsaw" redirect here. For other uses, see Warsaw (disambiguation), Warszawa (disambiguation), Warschau (disambiguation), and City of Warsaw (disambiguation).

Warsaw,[a] officially the **Capital City of Warsaw**,[4][b] is the capital and largest city of Poland. The metropolis stands on the River Vistula in east-central Poland and its population is officially estimated at 1.8 million residents within a greater metropolitan area of 3.1 million residents,[5] which makes Warsaw the 7th most-populous capital city in the European Union. The city area measures 517 km² (200 sq mi) and comprises 18 boroughs, while the metropolitan area covers 6,100 km² (2,355 sq mi).[6] Warsaw is an alpha-global city,[7] a major cultural, political and economic hub, and the country's seat of government. Its historical Old Town was designated a UNESCO World Heritage Site.

Warsaw traces its origins to a small fishing town in Masovia. The city rose to prominence in the late 16th century, when Sigismund III decided to move the Polish capital and his royal court from Kraków. Warsaw served as the de facto capital of the Polish–Lithuanian Commonwealth until 1795, and subsequently as the seat of Napoleon's Duchy of Warsaw. The 19th century and its Industrial Revolution brought a demographic boom which made it one of the largest and most densely-populated cities in Europe. Known then for its elegant architecture and boulevards, Warsaw was bombed and besieged at the start of World War II in 1939.[8][9][10] Much of the historic city was destroyed and its diverse population decimated by the Ghetto Uprising in 1943, the general Warsaw Uprising in 1944 and systematic razing.

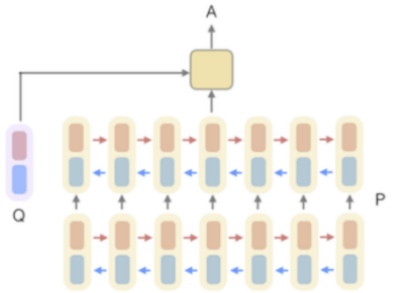
Warsaw is served by two international airports, the busiest being Warsaw Chopin and the smaller Warsaw Modlin intended for low-cost carriers. Major public transport services operating in the city include the Warsaw Metro, buses, urban-light railway and an extensive tram network. In 2012, the Economist Intelligence Unit ranked Warsaw as the 32nd most liveable city in the world.[11] In 2017, the city came 4th in the "Business-friendly", 8th in "Human capital and life style" and topped the quality of life rankings in the region.[12] The city is a significant centre of research and development, business process outsourcing, and information technology outsourcing. The Warsaw Stock Exchange is the largest and most important in Central and Eastern Europe.[13][14] Frontex, the European Union agency for external border security as well as ODIHR, one of the principal institutions of the Organization for Security and Cooperation in Europe have their headquarters in Warsaw. Jointly with Frankfurt and Paris, Warsaw features one of the highest number of skyscrapers in the European Union.[15]

The city hosts the Polish Academy of Sciences, National Philharmonic Orchestra, University of Warsaw, the Warsaw University of Technology, the National Museum, Zachęta Art Gallery and the Warsaw Grand Theatre, the largest of its kind in the world.[16] The reconstructed Old Town, which represents examples of nearly every European architectural style and historical period,[17] was listed as a World Heritage Site by UNESCO in 1980. Other main architectural attractions include the Royal Castle and the iconic King Sigismund's Column, the Wilanów Palace, the Palace on the Isle, St. John's Cathedral, Main Market Square, as well as numerous churches and mansions along the Royal Route. Warsaw possesses thriving arts and club scenes, gourmet restaurants and large urban green spaces, with around a quarter of the city's area occupied by parks.[18][19]

Warsaw
Warszawa (Polish)
Capital city and county

Document Reader

833,500



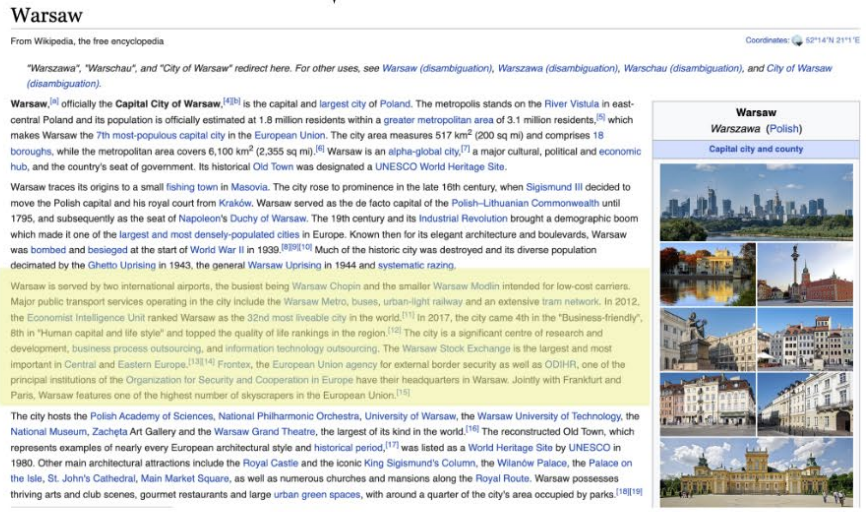
Open-Retrieval QA (ORQA)

What is QA? Reading Comprehension vs Open-Retrieval QA

Document Retriever

Reading comprehension

Q: How many of Warsaw's inhabitants spoke Polish in 1933?



Warsaw

From Wikipedia, the free encyclopedia

Coordinates: 52°14′N 21°1′E

"Warszawa", "Warschau", and "City of Warsaw" redirect here. For other uses, see Warsaw (disambiguation), Warszawa (disambiguation), Warschau (disambiguation), and City of Warsaw (disambiguation).

Warsaw,[a] officially the **Capital City of Warsaw**,[d][b] is the capital and largest city of Poland. The metropolis stands on the River Vistula in east-central Poland and its population is officially estimated at 1.8 million residents within a greater metropolitan area of 3.1 million residents,[5] which makes Warsaw the 7th most-populous capital city in the European Union. The city area measures 517 km^2 (200 sq mi) and comprises 18 boroughs, while the metropolitan area covers 6,100 km^2 (2,355 sq mi).[6] Warsaw is an alpha-global city,[7] a major cultural, political and economic hub, and the country's seat of government. Its historical Old Town was designated a UNESCO World Heritage Site.

Warsaw traces its origins to a small fishing town in Masovia. The city rose to prominence in the late 16th century, when Sigismund III decided to move the Polish capital and his royal court from Kraków. Warsaw served as the de facto capital of the Polish–Lithuanian Commonwealth until 1795, and subsequently as the seat of Napoleon's Duchy of Warsaw. The 19th century and its Industrial Revolution brought a demographic boom which made it one of the largest and most densely-populated cities in Europe. Known then for its elegant architecture and boulevards, Warsaw was bombed and besieged at the start of World War II in 1939.[8][9][10] Much of the historic city was destroyed and its diverse population decimated by the Ghetto Uprising in 1943, the general Warsaw Uprising in 1944 and systematic razing.

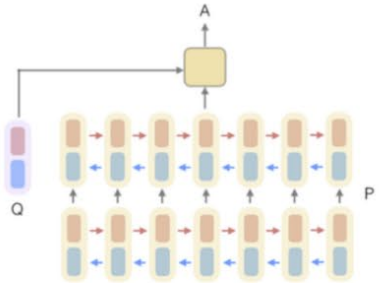
Warsaw is served by two international airports, the busiest being Warsaw Chopin and the smaller Warsaw Modlin intended for low-cost carriers. Major public transport services operating in the city include the Warsaw Metro, buses, urban-light railway and an extensive tram network. In 2012, the Economist Intelligence Unit ranked Warsaw as the 32nd most liveable city in the world.[11] In 2017, the city came 4th in the "Business-friendly", 8th in "Human capital and life style" and topped the quality of life rankings in the region.[12] The city is a significant centre of research and development, business process outsourcing, and information technology outsourcing. The Warsaw Stock Exchange is the largest and most important in Central and Eastern Europe.[13][14] Frontex, the European Union agency for external border security as well as ODIHR, one of the principal institutions of the Organization for Security and Cooperation in Europe have their headquarters in Warsaw. Jointly with Frankfurt and Paris, Warsaw features one of the highest number of skyscrapers in the European Union.[15]

The city hosts the Polish Academy of Sciences, National Philharmonic Orchestra, University of Warsaw, the Warsaw University of Technology, the National Museum, Zachęta Art Gallery and the Warsaw Grand Theatre, the largest of its kind in the world.[16] The reconstructed Old Town, which represents examples of nearly every European architectural style and historical period,[17] was listed as a World Heritage Site by UNESCO in 1980. Other main architectural attractions include the Royal Castle and the iconic King Sigismund's Column, the Wilanów Palace, the Palace on the Isle, St. John's Cathedral, Main Market Square, as well as numerous churches and mansions along the Royal Route. Warsaw possesses thriving arts and club scenes, gourmet restaurants and large urban green spaces, with around a quarter of the city's area occupied by parks.[18][19]

Warsaw
Warszawa (Polish)
Capital city and county

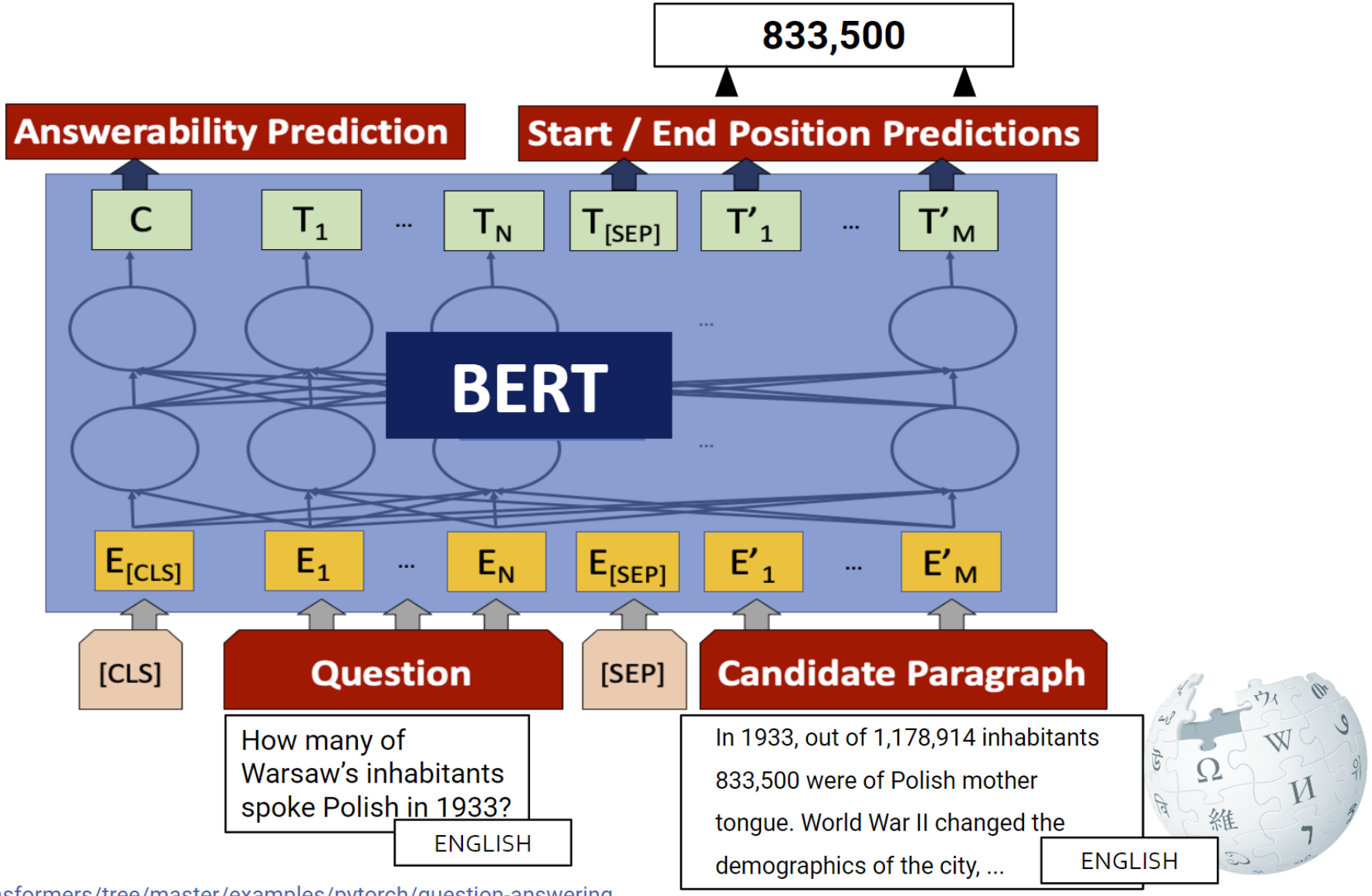
Document Reader

833,500



Machine Reading Comprehension (MRC)

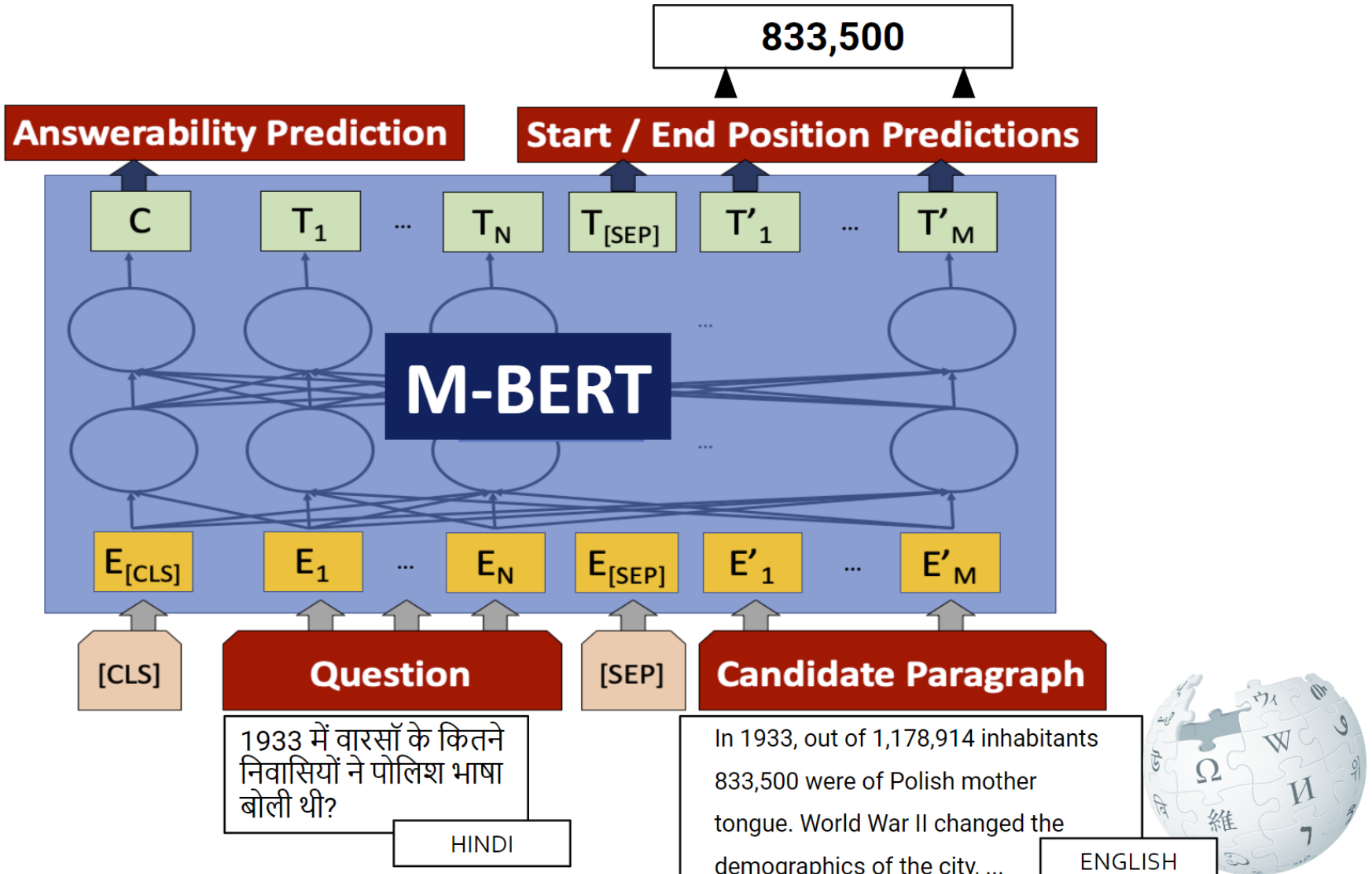
- Popular choice: Add a fine-tuning layer on top of BERT [Devlin et al., 2019]



Code: https://github.com/huggingface/transformers/tree/master/examples/pytorch/question-answering

Multilingual Machine Reading Comprehension (MRC)

- Popular choice: Add a fine-tuning layer on top of M-BERT [Bornea et al., 2021]



Code: https://github.com/huggingface/transformers/tree/master/examples/pytorch/question-answering

Why is QA in this part of the course?

- After IR
- NLU/SLU was just a scheduling detour ☺

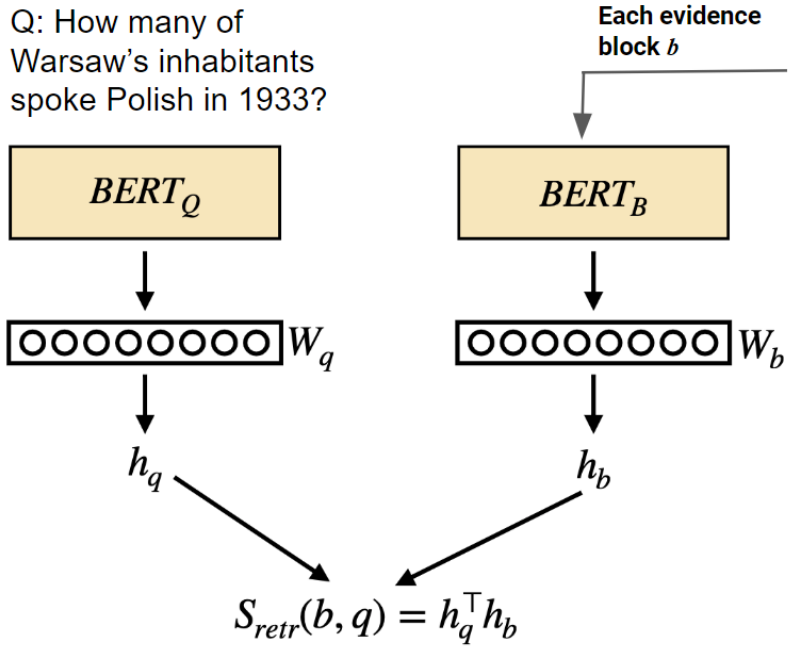
Open Retrieval QA (ORQA)

Q: How many of Warsaw's inhabitants spoke Polish in 1933?

Retrieve top-k passages

Perform Information Retrieval (IR).
Methods:

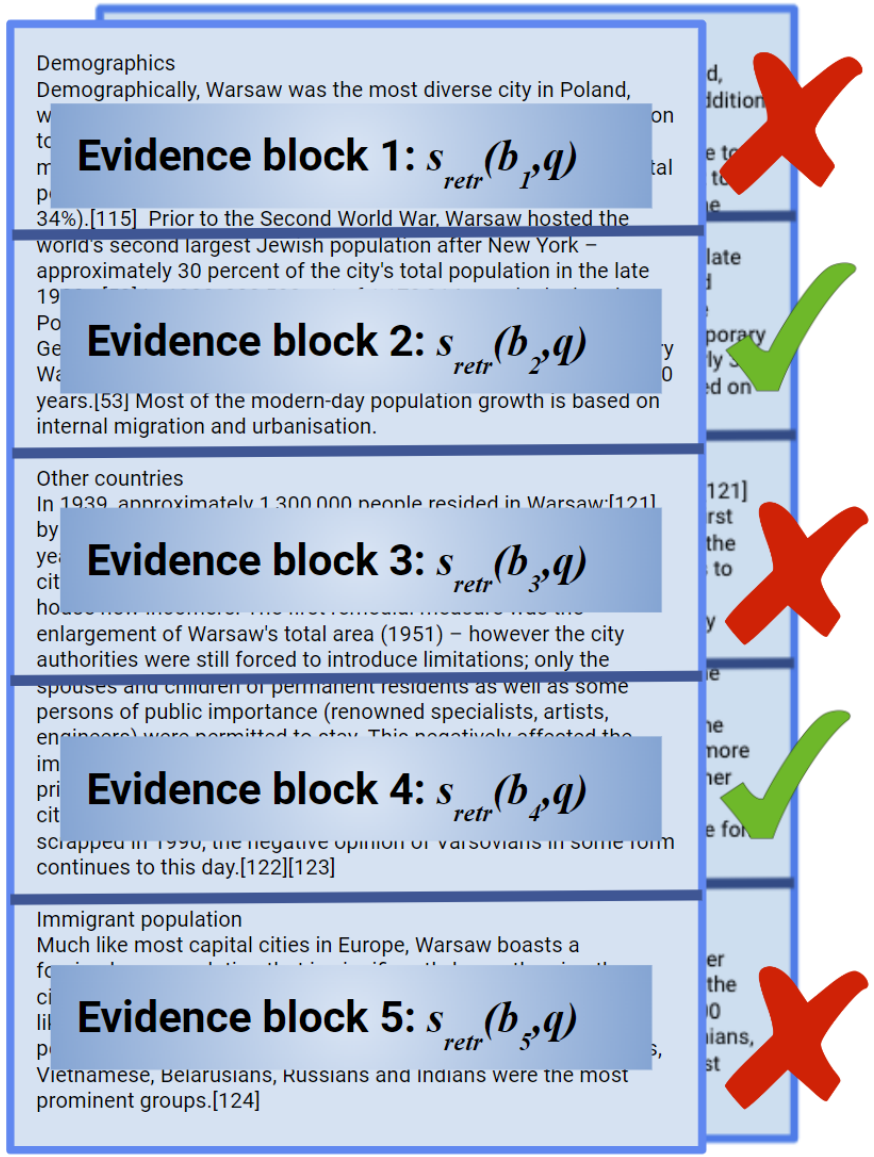
- Classic: BM25 [Robertson 2009]
- **Neural: Dense Passage Retrieval (DPR)** [Karpukhin et al., 2020]



Retriever score:

$$h_q = \mathbf{W_q}\text{BERT}_Q(q)[\text{CLS}]$$
$$h_b = \mathbf{W_b}\text{BERT}_B(b)[\text{CLS}]$$
$$S_{retr}(b, q) = h_q^\top h_b$$

Select top-k blocks from collection (e.g. Wikipedia



Open Retrieval QA (ORQA)

Q: How many of Warsaw's inhabitants spoke Polish in 1933?

Perform Information Retrieval (IR).
Methods:

- Classic: BM25 [Robertson 2009]
- **Neural: ColBERT** [Khattab et al., 2021]

Select top-k blocks from collection (e.g. Wikipedia)

Retrieve top-k passages

Q: How many of Warsaw's inhabitants spoke Polish in 1933?

Each evidence block b

Demographics
Demographically, Warsaw was the most diverse city in Poland,

Evidence block 1: $s_{retr}(b_1,q)$

34%).[115] Prior to the Second World War, Warsaw hosted the world's second largest Jewish population after New York – approximately 30 percent of the city's total population in the late

Evidence block 2: $s_{retr}(b_2,q)$

years.[53] Most of the modern-day population growth is based on internal migration and urbanisation.

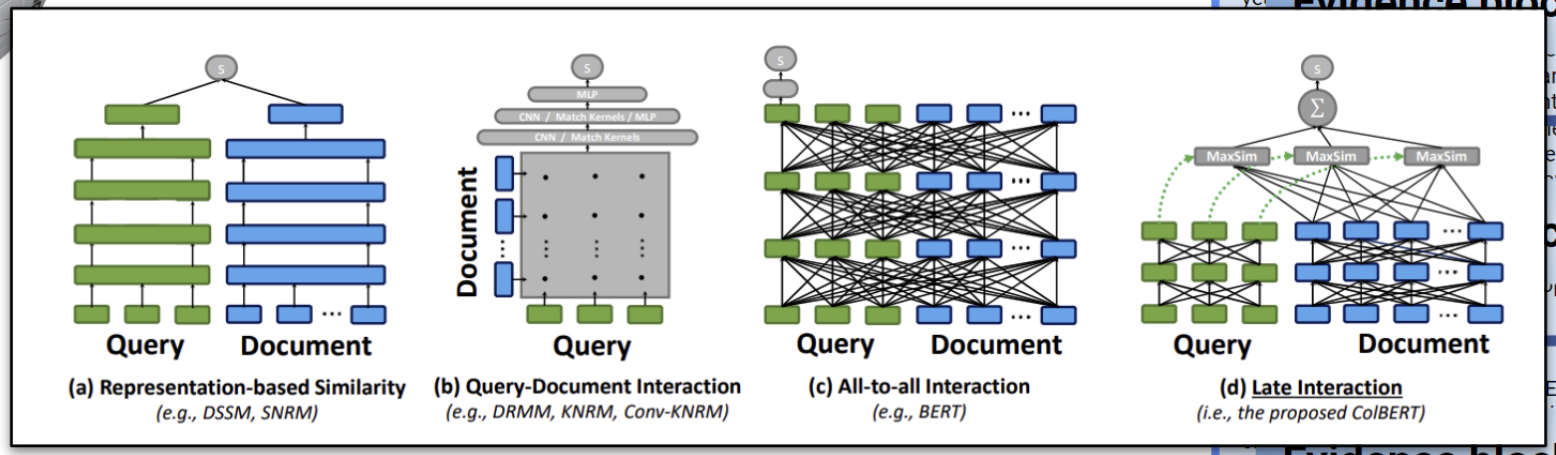
Other countries

Evidence block 3: $s_{retr}(b_3,q)$

Evidence block 4: $s_{retr}(b_4,q)$

Evidence block 5: $s_{retr}(b_5,q)$

Vietnamese, Belarusians, Russians and Indians were the most prominent groups.[124]



Other choices for Neural Retrievers

Question Answering (QA) vs. Information Retrieval (IR)

- QA and IR are related, but satisfy different info needs
- In QA, questions are in natural language sentences; in IR, queries tend to be short keyword phrases
- In QA, the answers are often short and to-the-point; in IR, the system returns lists of documents.
- In QA, the answer might be synthesized from multiple sources; In IR, a document is the atomic unit.

Slide stolen (with permission) from Kevin Duh

Types of QA

- Factoid Questions
  - Where is JHU located?
- List Questions
  - What centers are in Hopkins?
- Definitional Questions
  - Who is Ron Daniels?
- Relationship Questions
  - How is APL related to JHU?
- Opinion Questions
  - Why do people like Multilingual NLP?


THINK
$24,000
Who is Stoker?
((FOR ONE WELCOME OUR NEW COMPUTER OVERLORDS)
$ 1,000
$77,147
Who is Bram Stoker?
$ 17,973
$21,600
WHO IS BRAM STOKER?
$5600

Watson

- https://www.youtube.com/watch?v=P18EdAKuC1U

IBM Watson Wins Jeopardy!

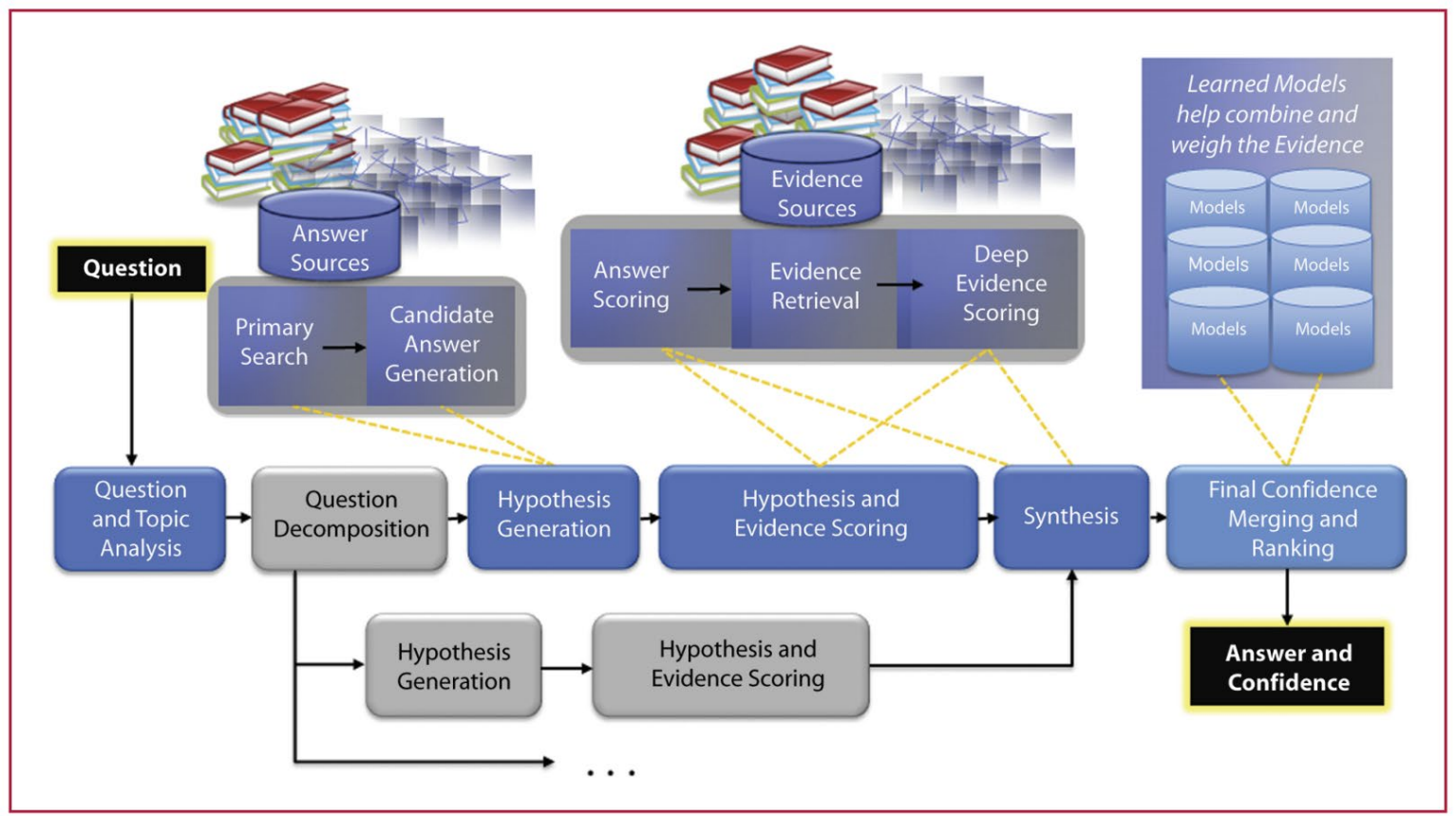


Figure 1

DeepQA architecture.

Ferucci 2012

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

Pranav Rajpurkar and **Jian Zhang** and **Konstantin Lopyrev** and **Percy Liang**
{pranavsr,zjian,klopyrev,pliang}@cs.stanford.edu
Computer Science Department
Stanford University

- Crowd-Sourced
- Spans, not answer list
- 107,785 question-answer pairs on 536 articles
- V2.0 contains 50k *unanswerable*

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".

What causes precipitation to fall?
gravity

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

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 SQuAD dataset. Each of the answers is a segment of text from the passage.

XQuAD

- 240 paragraphs and 1190 question-answer pairs from SQuAD v1.1
- Translated into ten languages by professional translators.
- "XQuAD translators see English questions and passages at the same time, priming them to use similar words"

On the Cross-lingual Transferability of Monolingual Representations

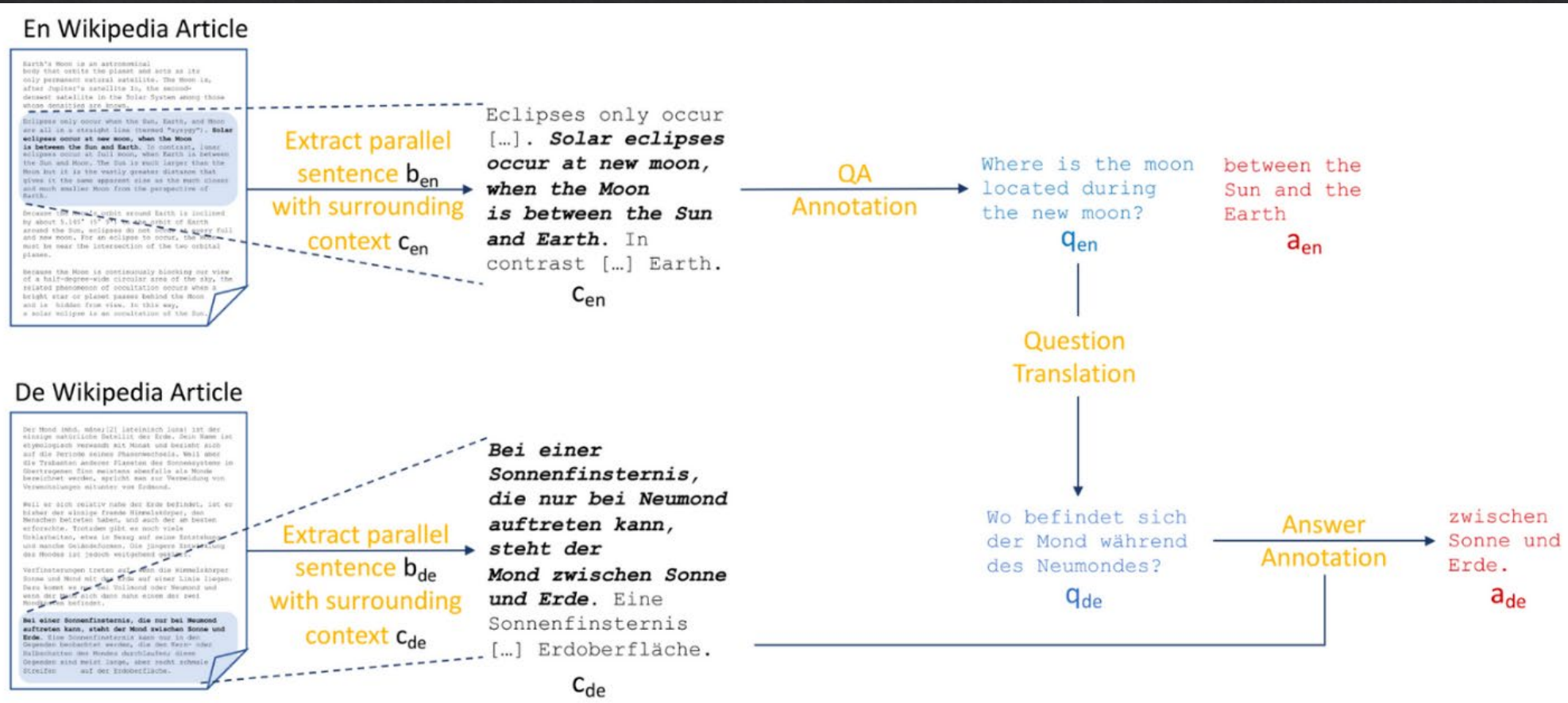
Mikel Artetxe[†*], **Sebastian Ruder**[‡], **Dani Yogatama**[‡]
[†]HiTZ Center, University of the Basque Country (UPV/EHU)
[‡]DeepMind
mikel.artetxe@ehu.eus
{ruder,dyogatama}@google.com

MLQA: Evaluating Cross-lingual Extractive Question Answering

Patrick Lewis[*†] **Barlas Oğuz**[*] **Ruty Rinott**[*] **Sebastian Riedel**[*†] **Holger Schwenk**[*]
[*]Facebook AI Research [†]University College London
{plewis,barlaso,ruty,sriedel,schwenk}@fb.com

- 7 Languages (English, Arabic, German, Spanish, Hindi, Vietnamese, Simpl. Chinese)
- Extractive QA
- SQuAD-like
- Average parallel of 4 languages
- Automatically align sentences across languages → Get Paragraph → Crowd-Source Questions on English → Question translated by professional translators

MLQA



The MLQA alignment and annotation process (Lewis et al., 2020)

TyDiQA

- Typologically Diverse Question Answering
- 11 Language Pairs
- 200k QA pairs

TyDi QA: A Benchmark for Information-Seeking Question Answering in *Ty*pologically *Di*verse Languages

Jonathan H. Clark **Eunsol Choi** **Michael Collins** **Dan Garrette**
Tom Kwiatkowski **Vitaly Nikolaev** **Jennimaria Palomaki**

Google Research

tydiqa@google.com

Language	Latin script[a]	White space tokens	Sentence boundaries	Word formation[b]	Gender[c]	Prodrop
English	+	+	+	+	+[d]	—
Arabic	—	+	+	++	+	+
Bengali	—	+	+	+	+	+
Finnish	+	+	+	+++	—	—
Indonesian	+	+	+	+	—	+
Japanese	—	—	+	+	—	+
Kiswahili	+	+	+	+++	—[e]	+
Korean	—	+[f]	+	+++	—	+
Russian	+	+	+	++	+	+
Telugu	—	+	+	+++	+	+
Thai	—	—	—	+	+	+

[a] '—' indicates **Latin script** is not the conventional writing system. Intermixing of Latin script should still be expected.

[b] We include inflectional and derivation phenomena in our notion of **word formation**.

[c] We limit the **gender** feature to sex-based gender systems associated with coreferential gendered personal pronouns.

[d] English has grammatical gender only in third person personal and possessive pronouns.

[e] Kiswahili has morphological noun classes (Corbett, 1991), but here we note sex-based gender systems.

[f] In Korean, tokens are often separated by whitespace, but prescriptive spacing conventions are commonly flouted.

Table 1: Typological features of the 11 languages in TyDi QA. We use + to indicate that this phenomena occurs, ++ to indicate that it occurs frequently, and +++ to indicate very frequently.

Question Word	TyDi QA	SQuAD
What	30%	51%
How	19%	12%
When	14%	8%
Where	14%	5%
(Yes/no)	10%	<1%
Who	9%	11%
Which	3%	5%
Why	1%	2%

Table 2: Distribution of question words in the English portion of the development data.

Spelling variation in Arabic transliteration

Q: من هو موزارت ؟
? mwzArt hw mn
Who is Mozart ?

A: فولفغانغ أماديوس موتسارت (27 يناير 1756 - 5 ديسمبر 1791) ولد في 27 يناير 1756 في سالزبورغ بالنمسا
fwlfgAng A#mAdyws mwtsArt (27 ynAyr 1756 - 5 dysmbr 1791) wld fy 27 ynAyr 1756 fy sAlzbwrg bAlnmsA
Wolfgang Amadeus Mozart (January 27, 1756 - December 5, 1791) was born on January 27, 1756 in Salzburg, Austria

This Arabic example demonstrates variation in the spelling of non-native names. Both spellings of Mozart are correct and refer to the same entity across the QA pair.

Script switching in Russian

Q:	Кто	изобрел	телефон	?
	Kto	izobrel	telefon	?
	who	invented	telephone	?

Who invented the telephone ?

A:	Сам	Рейс	назвал	сконструированное	им	устройство	Telephone .
	Sam	Reis	nazyval	skonstruirovannoe	im	ustroistvo	Telephone .
	self	Reis	called	constructed	him	device	Telephone .

Reis himself called the device he created the Telephone .

Vowel diacritization in Arabic

Q: ما هي الوان العلم العُماني ؟

? mA hy AlwAn AlElm AlEumAny ?

What are the colors of the Omani flag?

A: العلم الوطني لسلطنة عمان انشئ بقرار سلطاني ورفع لأول مرة في 18 شوال 1391 هـ الموافق 17 ديسمبر 1970 ، 1970 .

AlElm AlwTny lslTnp EmAn An$y# bqrAr slTAny wrfE lA#wl mrp fy 18 $wAl 1391 h_ AlmwAfq 17 dysmbr 1970 .

The national flag of the Sultanate of Oman was established by a royal decree and was raised for the first time on Shawwal 18, 1391 AH corresponding to December 17, 1970.

Vowel diacritization in Russian

Q:	Что	такое	атом	?			
	Chto	takoe	atom	?			
	What	such	atom	?			
	What is an		*atom*	*?*			

A:	Á том	—	частица	вещества	микроскопических	размеров	...
	Á tom	—	chastitsa	veschestva	mikroskopicheskih	razmerov	...
	Atom	PRED	particle	matter	microscopic	sizes	...
	An	*atom*	*is a microscopic particle of matter...*				

Word boundary variation in Arabic

Q: متى ولد عبدالسلام بن محمد ؟

? mHmd bn EbdAlslAm wld mtY

When was AbdulSalam bin Muhammad born?

A: عبد السلام بن محمد بن أحمد العَلَمي (1246 - هـ 1322 / 1830 - م 1904) .

Ebd AlslAm bn mHmd bn A#Hmd AlEalamy (h_ 1246 - h_ 1322 / 1830 - m 1904) .

Abdul Salam bin Muhammad bin Ahmed Al-Alami (1246 AH - 1322 AH / 1830 - 1904 AD).

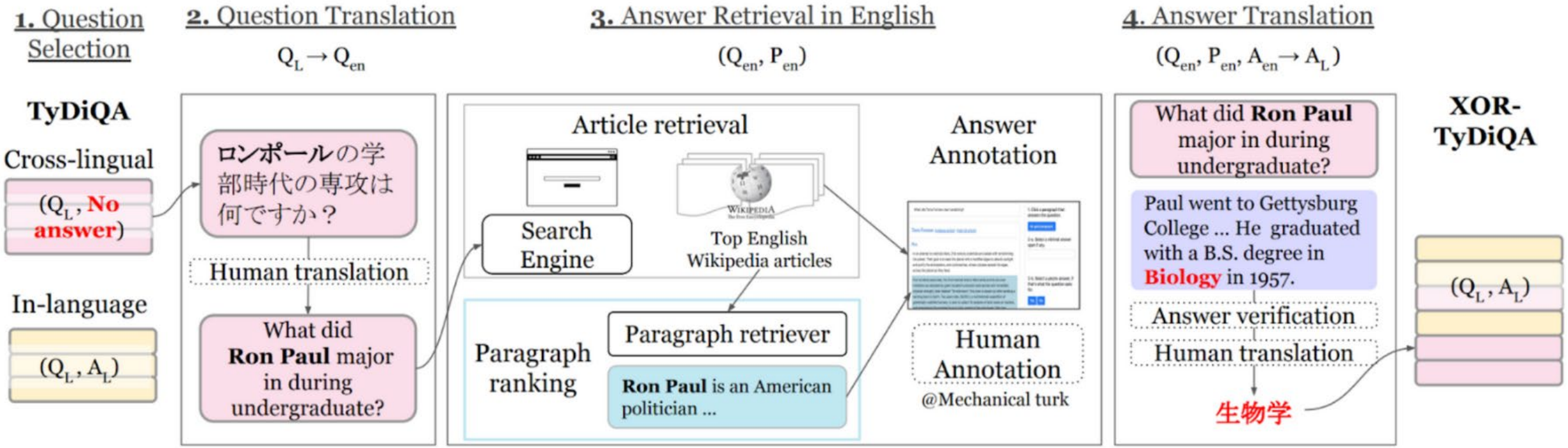
Why not translate?

- Translationese
- “For example, in TYDI QA, one Bengali question asks What does sapodilla taste like?”

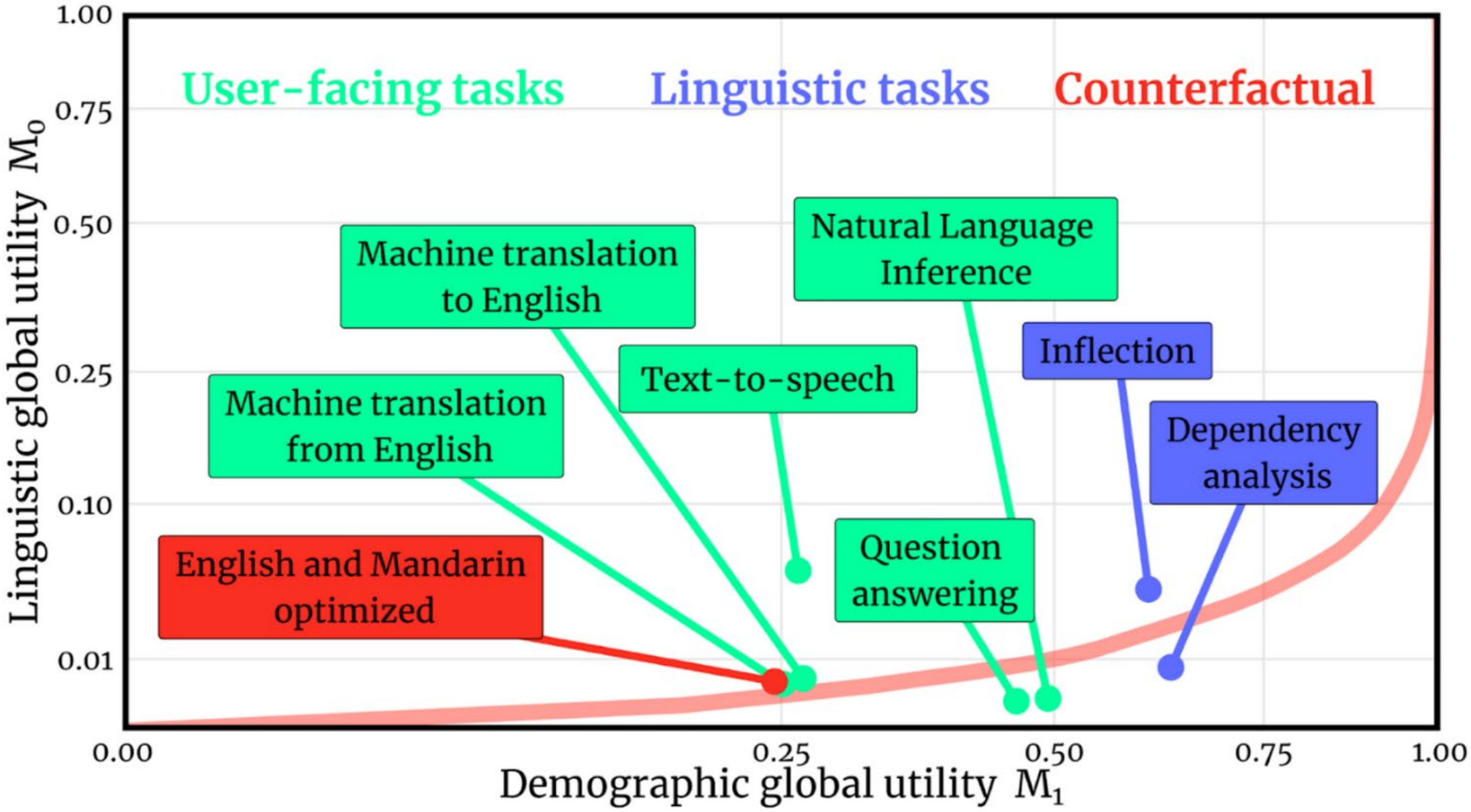
	TyDiQA-GoldP	SQuAD Zero Shot	Human
(English)	(76.8)	(73.4)	(84.2)
Arabic	81.7	60.3	85.8
Bengali	75.4	57.3	94.8
Finnish	79.4	56.2	87.0
Indonesian	84.8	60.8	92.0
Kiswahili	81.9	52.9	92.0
Korean	69.2	50.0	82.0
Russian	76.2	64.4	96.3
Telugu	83.3	49.3	97.1
Overall	79.0	56.4	90.9

Table 7: F1 scores for the simplified TyDiQA-GoldP task v1.1. *Left:* Fine tuned and evaluated on the TyDiQA-GoldP set. *Middle:* Fine tuned on SQuAD v1.1 and evaluated on the TyDiQA-GoldP dev set, following the XQuAD zero-shot setting. *Right:* Estimate of human performance on TyDiQA-GoldP. Models are averaged over 5 fine tunings.

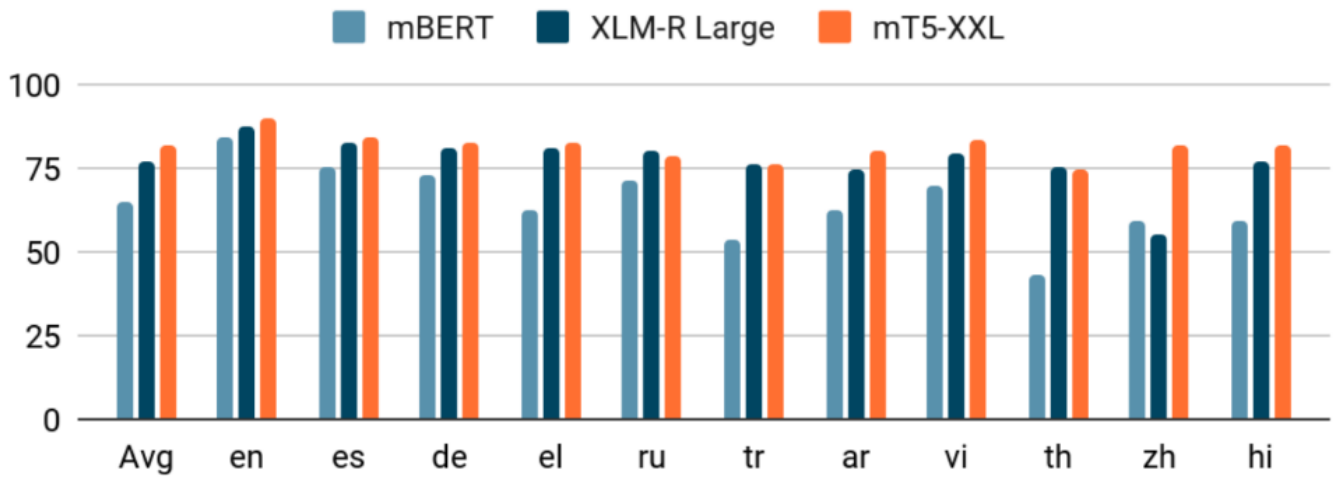
XOR-TyDi QA

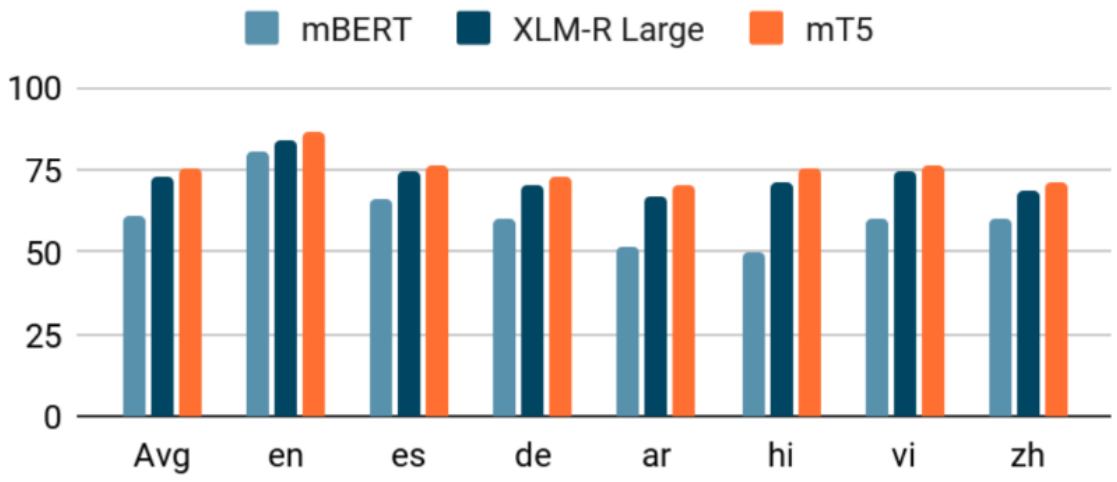


The XOR-TyDi QA annotation process (Asai et al., 2021)



Linguistic and demographic utility of different NLP applications (Blasi et al., 2021)





Zero-shot cross-lingual transfer performance (F1) of representative models on XQuAD (left) and MLQA (right) (Ruder et al., 2021)

$(q^{\mathrm{En}}, a^{\mathrm{En}}, p^{\mathrm{En}}_{\mathrm{gold}})$

"who sings nothing's gonna stop us now?"

Starship (band)

From Wikipedia, the free encyclopedia

"**Nothing's Gonna Stop Us Now**" is a song co-written by Diane Warren and Albert Hammond,[3] recorded by American rock band Starship for their second studio album, *No Protection* (1987). It is a power ballad[4] duet featuring Starship vocalists Grace Slick and Mickey Thomas. It is the theme to the romantic comedy film *Mannequin*.[5][6]

Nothing's Gonna Stop Us Now

From Wikipedia, the free encyclopedia

WIKIDATA

スターシップ (バンド)

出典: フリー百科事典『ウィキペディア（Wikipedia）』

愛はとまらない

出典: フリー百科事典『ウィキペディア（Wikipedia）』

「**愛はとまらない**（原題：*Nothing's Gonna Stop Us Now*）」（あいはとまらない）はアルバート・ハモンドとダイアン・ウォーレンの共作による楽曲[2]。アメリカ合衆国のロックバンド、スターシップにより録音された。スターシップの二人のヴォーカリスト、ミッキー・トーマスとグレイス・スリックのデュエットを前面に出した曲である。ロマンティック・コメディ映画である『マネキン』のテーマ曲になった[3][4]。1987年4月4日ビルボードホット100で第1位となる。また、同年夏5月、UKシングルチャートで4週間第1位の座にとどまった。1987年中にイギリス国内で最も売れたシングルレコードのチャートで第2位を獲得した。当時、グレイス・スリック（47歳）は、アメリカ国内で第1位となったシングルレコードを歌った女性歌手としての最年長記録保持者となった（この記録は、1999年にシェールが52歳で「ビリーヴ」を歌ってヒットさせたことで塗り替えられた）[5]。日本のオリコン洋楽シングルチャートでは1987年4月20日付から3週連続1位を獲得した[6]。

この曲は、第60回のアカデミー歌曲賞候補にも指名された。映画『マネキン』のサウンドトラックに含まれている他、1987年7月に発売されたスターシップのアルバム『No Protection』に含まれている。島田歌穂が1992年に『CHANCE〜夢はとまらない〜』として日本語カバーしている。

ラジオのインタヴューでアルバート・ハモンドは次のように語っている：「この曲のアイデアは、自分の前妻との離婚が成立した直後、7年にわたって同棲してきた彼女といよいよ結婚しようかという時に閃いた。」ハモンドがダイアン・ウォーレンに言ったところでは、「自分の彼女との結婚を7年間も、まるで結婚されているような感じだったんだけど、結局最終的には彼らは失敗したんだ。もう俺が結婚するのを止めることはできないんだ」[7]。

最近は、この曲はイギリスのテレビ番組『The X Factor』で、視聴者が演奏内容を確認する時にそのバックで流れる曲として使用されている。この曲は以前から「力強いシンセによるビートが、情感を心地良く盛り上げてくれる」曲として受け止められている[8]。

$(q^{\mathrm{En}}, a^{\mathrm{Ja}}, p^{\mathrm{Ja}})$

Cross-lingual data expansion via Wikidata language links; based on an example by Asai et al. (2021)