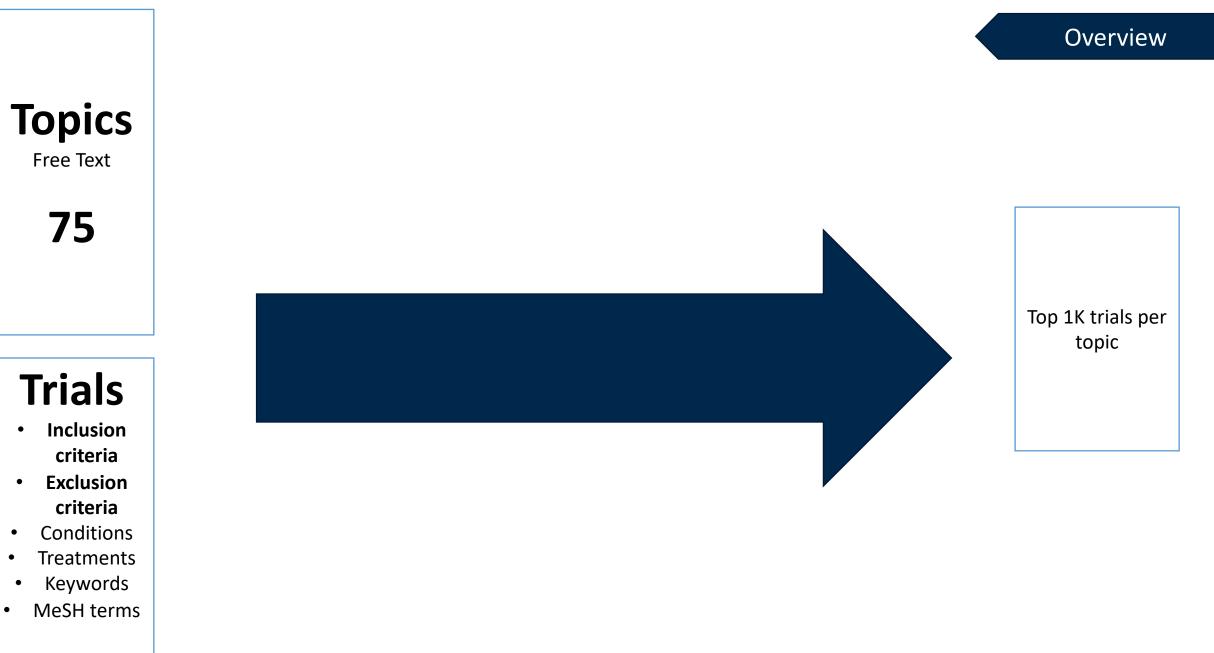
IBM @ TREC Clinical Trials Track 2021

Laura Biester

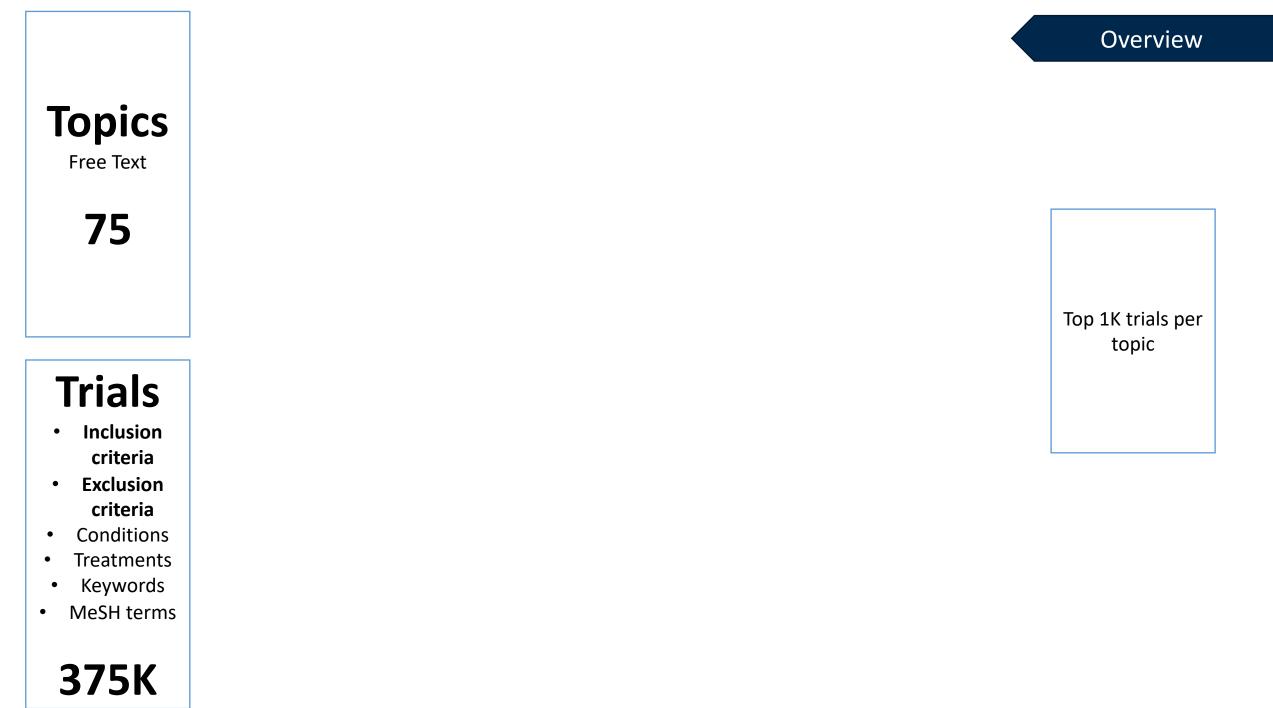
Venkata Joopudi

Bharath Dandala

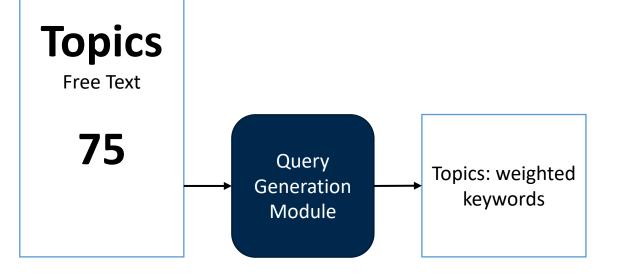




375K







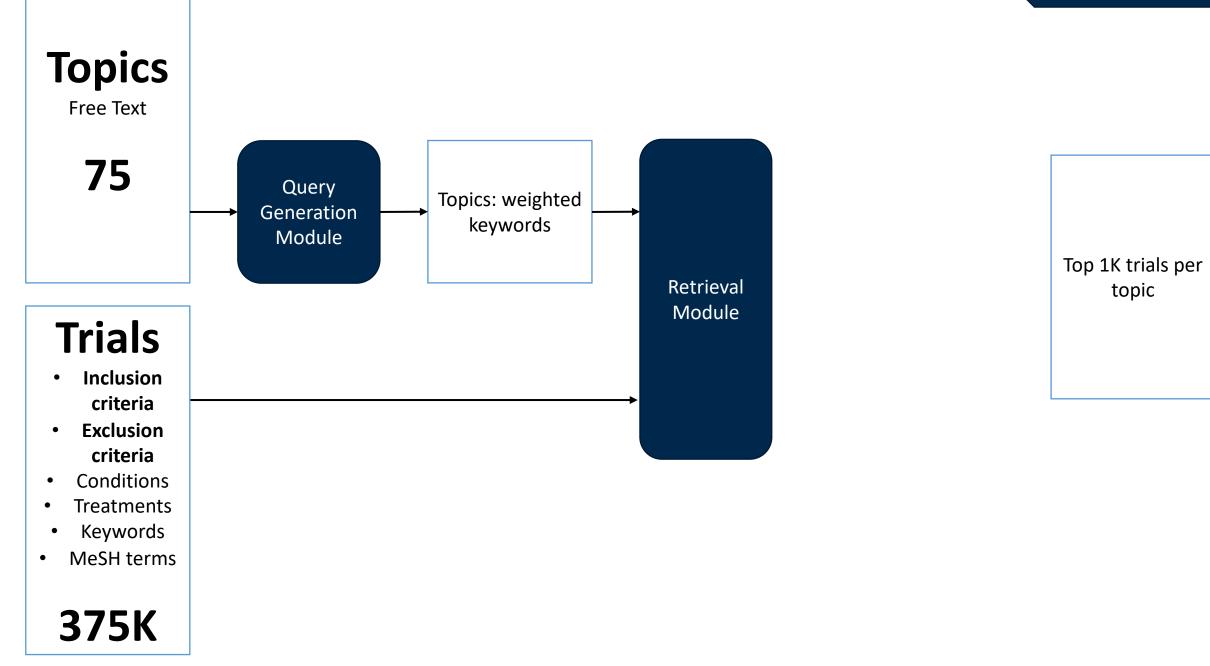
Trials

- Inclusion criteria
- Exclusion criteria
- Conditions
- Treatments
- Keywords
- MeSH terms

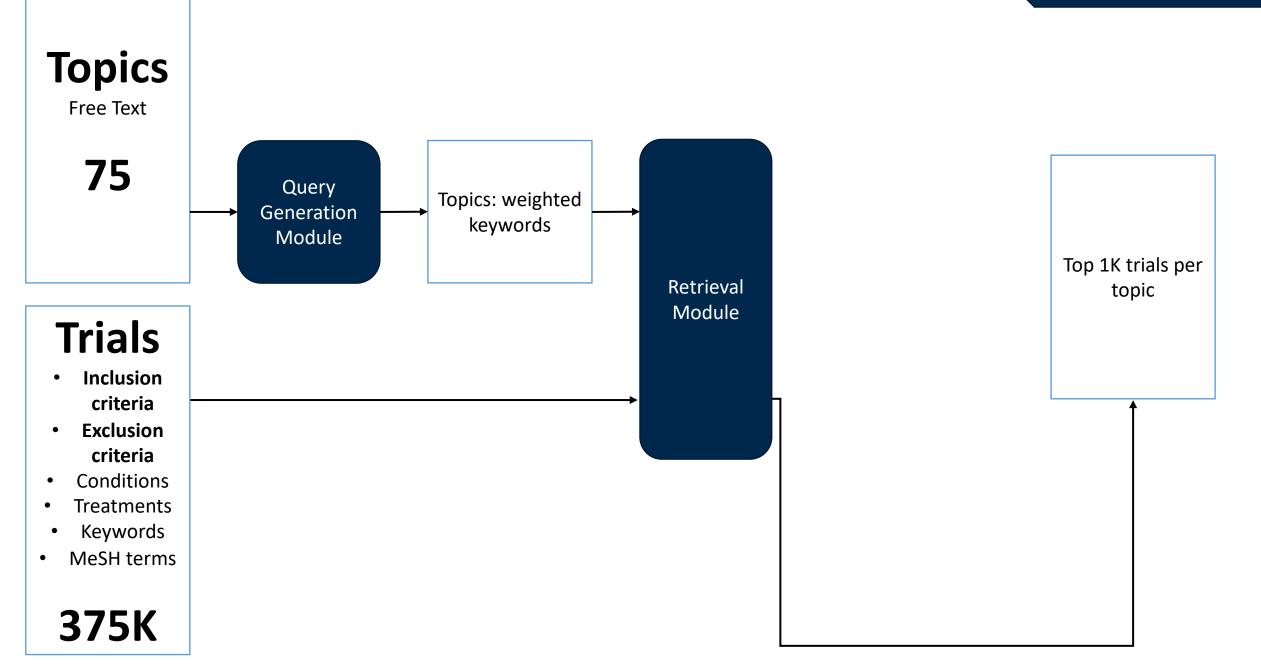
375K

Top 1K trials per topic

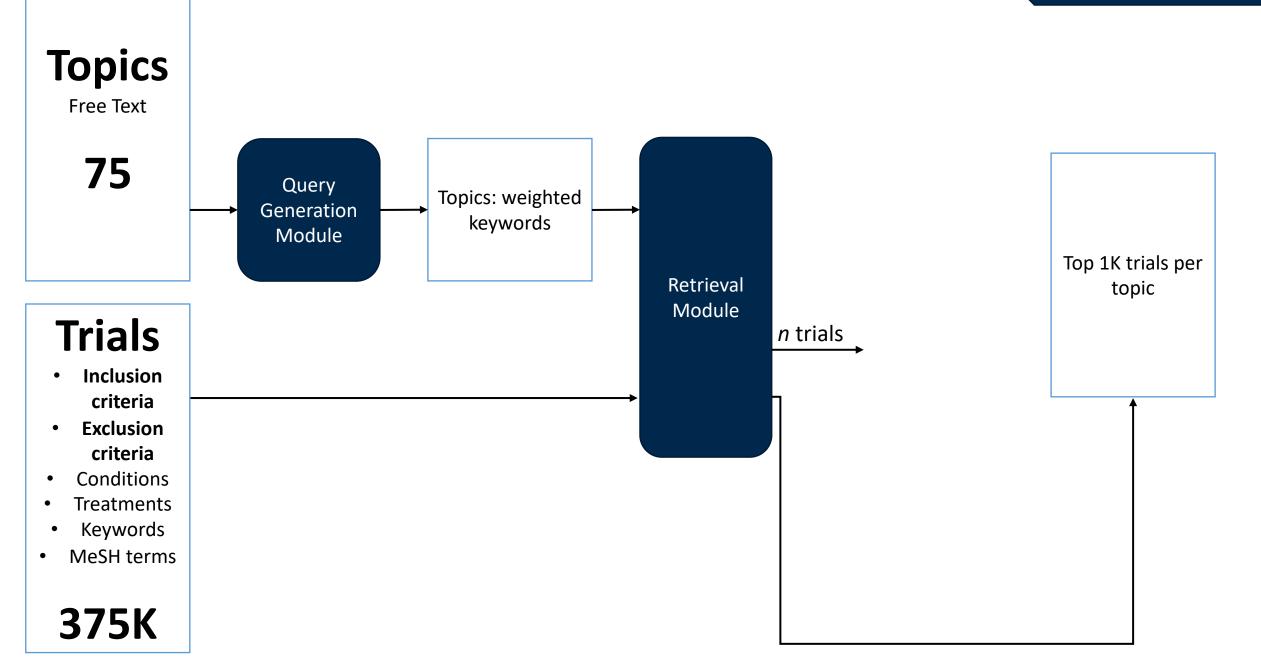


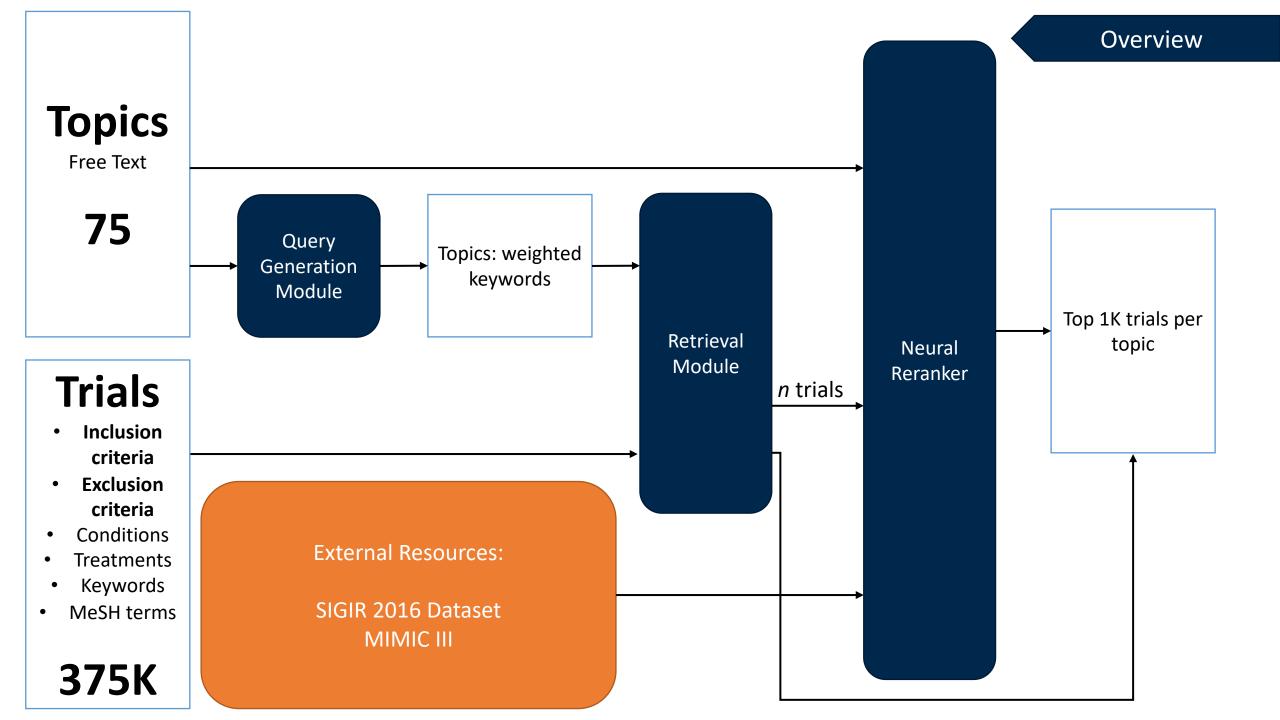












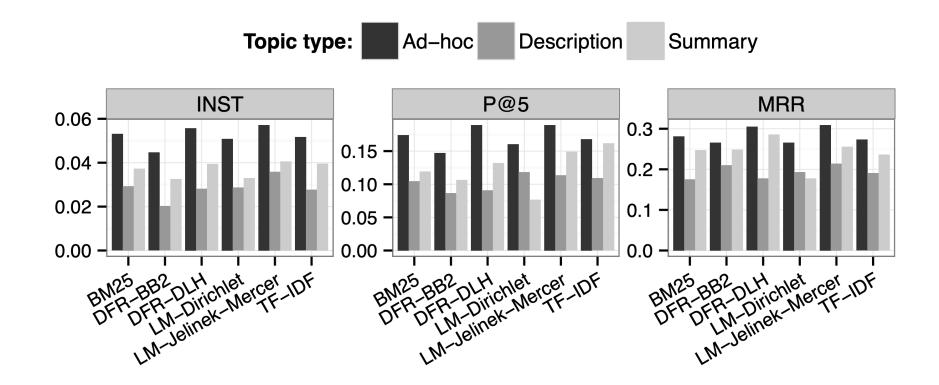
Query Generation

Ad-hoc Query Inspiration

Koopman and Zuccon (2016) use ad-hoc queries for retrieval

Ad-hoc Query Inspiration

Koopman and Zuccon (2016) use ad-hoc queries for retrieval

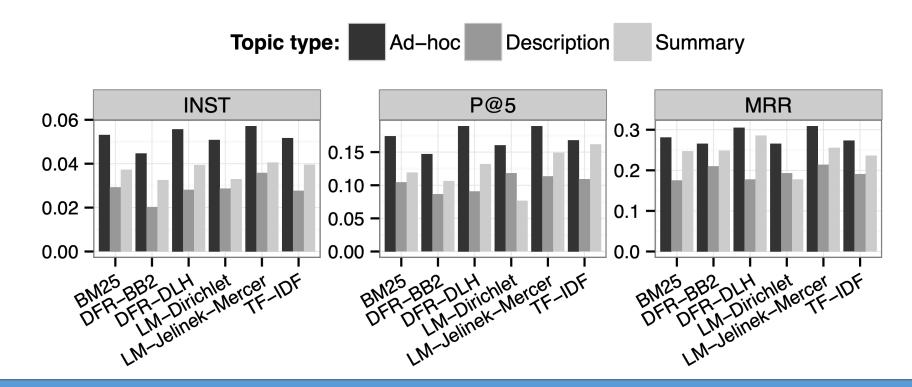


Methodology

Query Generation

Ad-hoc Query Inspiration

Koopman and Zuccon (2016) use ad-hoc queries for retrieval



Methodology

Query Generation

Can we mimic this process with automated methods?

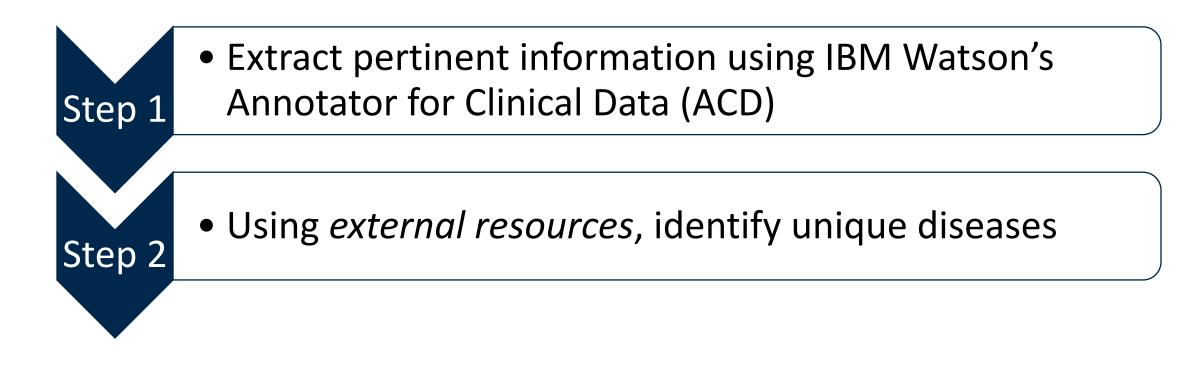
Query Generation

Overview

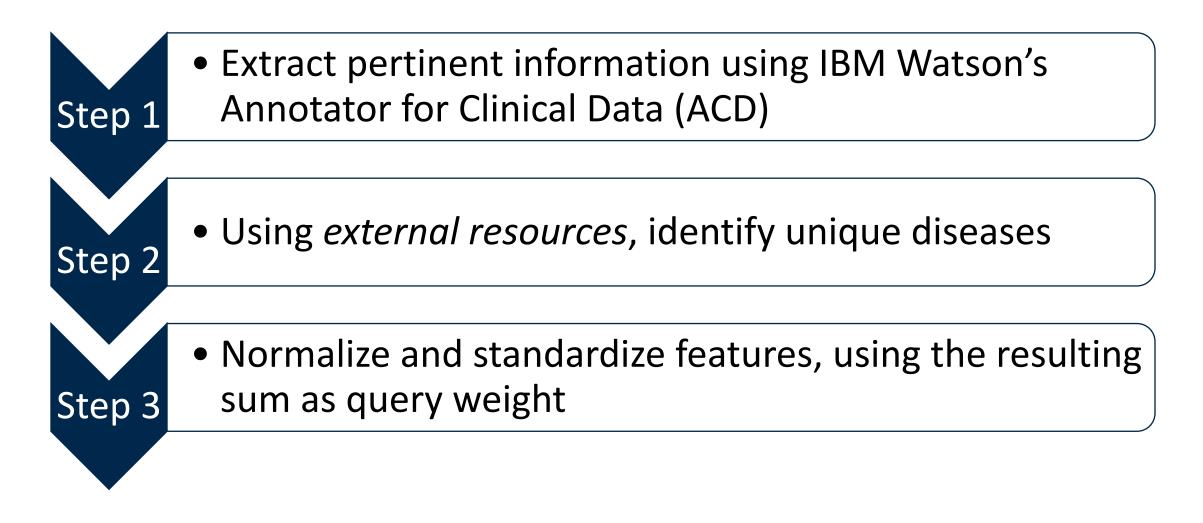


Extract pertinent information using IBM Watson's Annotator for Clinical Data (ACD)

Overview



Overview



Query Generation

Query Example

Preview	JSON Output		1	Reset Sample Text \$\circlerify Edit \$\circlerif	Text 🖉	Insights (1)		~
A 35-year-old	woman presents	with history of <mark>acne^p and mild (</mark>	<mark>nirsutism^D The primary eva</mark>	luation revealed elevated te	estosterone	Show/hide all		
recently notice	ed gradual <mark>enlarge</mark>	<mark>ment^e of her hands and feet an</mark>	d recognized that her ring	g is getting small for her	finger. The	☐ ● AbnormalFinding (A) (Ì	5 ~
irregularity in	her menstrual over	le as well as some <mark>nipple disc</mark>	harde ^D She also has nos	itive history for sporing ^D an	beadache	🖌 🕒 Diagnosis (D) 🛈		10 ^
inegularity in	nei nenstruar cyc	ie as well as some (lipple disc	narge she also has pos	and shoring an		✓ acne vulgaris		1 ~
examination re	evealed subtle faci	al features of <mark>acromegaly^D and p</mark>	<mark>rognathism^D</mark> Visual fields a	re normal by confrontation.	. <mark>Hirsutism^D</mark>	✓ acromegaly		1 ^
tissue <mark>thickeni</mark>	ng ^D and diaphoresi	⁹ of the hands and feet are n	oted. Laboratory evaluation	in the fasting state revea	als IGF-1 c	Usage:		
			with an investige Obs. it			Explicit I		94.4%
and random	GH OF 19.7 ng/m	MRI reveals a macroadenoma	with no invasion. She is	s on stable doses of octro	eolide LAR	Discussed		5.5%
diagnosis was	confirmed. She	s married and has 2 children.	She is using IUD as he	r contraceptive method.		Patient reported		0.1%

Query Generation

Query Example

Preview	JSON Output	Reset Sample Text の Edit Text ⊿ Insights ①	~
A 35-year-old	woman presents	with history of acne ^D and mild hirsutism ^D The primary evaluation revealed elevated testosterone Show/hide all	
recently notice	ed gradual <mark>enlarg</mark>	ement ^e of her hands and feet and recognized that her ring is getting small for her finger. The	Ĵ 5 ~
irregularity in	her menstrual c	ycle as well as some <mark>nipple discharge^p She also has positive history for <mark>snoring^p and headache</mark></mark>	10 ^
		✓ acne vulgaris	1 ~
examination re	evealed subtle fa	cial features of acromegaly ^D and prognathism ^D Visual fields are normal by confrontation. Hirsutism ^D	1 ^
tissue <mark>thickeni</mark>	<mark>ng^D and diaphore</mark>	sis ^D of the hands and feet are noted. Laboratory evaluation in the fasting state reveals IGF-1 c	
and random	GH of 19.7 ng/i	mL. MRI reveals a macroadenoma with no invasion. She is on stable doses of octreotide LAR	● 94.4% ● 5.5%
		Patient reported	0.1%
diagnosis was	confirmed. She	is married and has 2 children. She is using IUD as her contraceptive method.	

extracted entities

acromegaly congenital prognathism nipple discharge hirsutism acne vulgaris snoring excessive sweating increased thickness headache

Query Generation

Query Example

Preview	JSON Output				Reset Sample Text 🏼 🎝	Edit Text 🖉	Insigh	nts (i)			~
A 35-year-old	woman present	s with history of <mark>ac</mark>	<mark>ne^p and mild <mark>hirsutism</mark></mark>	P The primary e	evaluation revealed elev	vated testosterone	Show/hide				
recently notice	d gradual <mark>enlar</mark>	<mark>gement^p)</mark> of her hand	s and feet and recog	nized that her ı	ring is getting small f	or her finger. The		AbnormalFinding (A) (1)	5	~
irregularity in	her menstrual of	cvcle as well as so	me <mark>nipple discharge^p :</mark>	She also has p	ositive history for <mark>snor</mark>	ring ^D and headache		Diagnosis (D) 🛈		10	^
								acne vulgaris			~
examination re	evealed subtle ta	acial features of <mark>acro</mark>	omegaly ^D and prognathis	sm ^b Visual fields	are normal by confro	ontation. Hirsutism ^D		 acromegaly 			^
tissue <mark>thickenii</mark>	ng ^D and diaphore	<mark>esis^D</mark> of the hands a	nd feet are noted. La	aboratory evaluati	on in the fasting stat	e reveals IGF-1 c		Usage: Explicit		0/	4.4%
and random	GH of 19.7 ng,	/mL. MRI reveals a	macroadenoma with n	io invasion. She	is on stable doses	of octreotide LAR		Discussed	•		.5%
diagnosis was	confirmed. She	e is married and ha	s 2 children. She is	using IUD as I	her contraceptive metho	od.		Patient reported		0.1	.1%

extracted entities

acromegaly
congenital prognathism
nipple discharge
hirsutism
acne vulgaris
snoring
excessive sweating
increased thickness
headache

feature	sign	snoring	acromegaly
IsPotentialDiagnosis	+	0	1
IsDiagnosis	+	0	1
IsPatientReportedCondition	+	0	0
IsSymptom	-	1	0
IsRareDisease	+	0	1
IsPMH	-	1	0
IDF_MIMIC	-	5	000001
IDF_PUBMED	-	5	000002

sight w/hide			^
	AbnormalFinding (A) (j)	5 ~
~ •	Diagnosis (D) 🛈		10 ^
~	acne vulgaris		1 ~
~	acromegaly		1 ^
	Usage:		
	Explicit		94.4%
	Discussed		5.5%
	Patient reported		0.1%

Query Generation

https://acd-try-it-out.mybluemix.net/preview

Query Example

Preview	JSON Output		Reset Sample Text 🏼 🅥	Edit Text 🖉	Insights ①	^
A 35-year-old	woman presents	; with history of <mark>acne^p and mild <mark>hirsutism^p T</mark></mark>	he primary evaluation revealed eleva	ted testosterone	Show/hide all	
recently notic	ed gradual <mark>enlar</mark> g	<mark>jement^e)</mark> of her hands and feet and recognize	d that her ring is getting small for	r her finger. The	☐ ● AbnormalFinding (A) ③	5 ~
irregularity in	her menstrual o	ycle as well as some <mark>nipple discharge^p Sh</mark> e	also has positive history for snorin	ng ^D and headache	✓ ● Diagnosis (D) ①	10 ^
					✓ acne vulgaris	1 ~
examination r	evealed subtle fa	cial features of <mark>acromegaly^D and prognathism^D</mark>	Visual fields are normal by confron	ntation. <mark>Hirsutism^D</mark>	acromegaly	1 ^
tissue <mark>thicken</mark>	<mark>ing[⊅] and diaphore</mark>	<mark>sis^D</mark> of the hands and feet are noted. Labora	atory evaluation in the fasting state	reveals IGF-1 c	Usage:	
and random	GH of 19.7 ng/	mL. MRI reveals a macroadenoma with no ii	nvasion. She is on stable doses of	f octreotide LAR	Explicit Explicit	94.4% 5.5%
diagnosis was	s confirmed. She	is married and has 2 children. She is usir	ng IUD as her contraceptive method	J.	Patient reported	0.1%

extracted	entities
extracted	entities

acromegaly congenital prognathism nipple discharge hirsutism acne vulgaris snoring excessive sweating increased thickness headache

feature	sign	snoring	acromegaly
IsPotentialDiagnosis	+	0	1
IsDiagnosis	+	0	1
IsPatientReportedCondition	+	0	0
IsSymptom	-	1	0
IsRareDisease	+	0	1
IsPMH	-	1	0
IDF_MIMIC	-	5	000001
IDF_PUBMED	-	5	000002

Query Generation

Query Example

Preview	JSON Output			Reset Sample Text	Edit Text 🖉	i li	nsights ①		_
A 35-year-old	woman presents	s with history of <mark>acne[⊅] and mild</mark>	<mark>hirsutism^D The primary ev</mark>	aluation revealed eleva	ated testosteron		how/hide all		
recently notice	ed gradual <mark>enlarg</mark>	<mark>jement^e)</mark> of her hands and feet a	nd recognized that her ri	ng is getting small fo	r her finger. ⁻	The _	□ ● AbnormalFinding (A)	Ì	5 ~
							<mark> -</mark> Diagnosis (D) 🗓		10 ^
irregularity in	ner menstrual c	ycle as well as some <mark>nipple dis</mark>	scharge ^e She also has po	isitive history for <mark>shoril</mark>	ng ^e and headad	cne	✓ acne vulgaris		1 ~
examination re	evealed subtle fa	cial features of <mark>acromegaly^D and</mark>	<mark>prognathism[⊅] Visual fields</mark>	are normal by confro	ntation. <mark>Hirsutis</mark> i	m ^D	✓ acromegaly		1 ^
tissue <mark>thicken</mark>	ing ^D and diaphore	<mark>sis^p of the hands and feet are</mark>	noted. Laboratory evaluatio	n in the fasting state	reveals IGF-1	с	Usage:		
							Explicit		94.4%
and random	GH of 19.7 ng/i	mL. MRI reveals a macroadenom	a with no invasion. She	is on stable doses o	f octreotide LA	٨R	Discussed	-	5.5%
diagnosis was	s confirmed She	is married and has 2 children	She is using ILID as h	er contracentive metho	d		Patient reported		0.1%

diagnosis was confirmed. She is married and has 2 children. She is using IUD as her contraceptive method.

	teature
extracted entities	IsPotentialDiagnosis
acromegaly	IsDiagnosis
congenital prognathism nipple discharge	IsPatientReportedCondition
hirsutism	IsSymptom
acne vulgaris	IsRareDisease
snoring	IsPMH
excessive sweating	IDF MIMIC
increased thickness	
headache	IDF_PUBMED

feature	sign	snoring	acromegaly
IsPotentialDiagnosis	+	0	1
IsDiagnosis	+	0	1
IsPatientReportedCondition	+	0	0
IsSymptom	-	1	0
IsRareDisease	+	0	1
IsPMH	-	1	0
IDF_MIMIC	-	5	000001
IDF_PUBMED	-	5	000002

weighted ad-hoc query list

acromegaly **0.8165** congenital prognathism **0.54** nipple discharge 0.24 hirsutism 0.24 acne vulgaris 0.09 snoring 0.00 excessive sweating 0.00 increased thickness 0.00 headache 0.00

Standardization

and normalization

Retrieval

Retrieval Modules

Lucene

- Using our ad-hoc queries, we search a Lucene Index with BM25
- Index includes trial conditions and treatment fields
- Query-level boosting used to boost terms according to our query weights

Retrieval Modules

Lucene

- Using our ad-hoc queries, we search a Lucene Index with BM25
- Index includes trial *conditions* and *treatment* fields
- Query-level boosting used to boost terms according to our query weights

Semantic Textual Similarity (STS)

- Transformer-based model used to encode text
- Measure weight-normalized pairwise cosine similarity between topic queries and trial condition/intervention fields

Retrieval Modules

Lucene

- Using our ad-hoc queries, we search a Lucene Index with BM25
- Index includes trial *conditions* and *treatment* fields
- Query-level boosting used to boost terms according to our query weights

Semantic Textual Similarity (STS)

- Transformer-based model used to encode text
- Measure weight-normalized pairwise cosine similarity between topic queries and trial condition/intervention fields

As input to neural rerankers, use a weighted average of Lucene and STS ranks to compute top 2k trials

Data

Two obvious data sources: TREC and Koopman and Zuccon's dataset from SIGIR

Dataset	Patient Descriptions	Clinical Trials	Labeled Pairs
TREC	75	375K	0
SIGIR	60	204K	3870

Data

Two obvious data sources: TREC and Koopman and Zuccon's dataset from SIGIR

Dataset	Patient Descriptions	Clinical Trials	Labeled Pairs
TREC	75	375K	0
SIGIR	60	204K	3870

Data

Two obvious data sources: TREC and Koopman and Zuccon's dataset from SIGIR

Dataset	Patient Descriptions	Clinical Trials	Labeled Pairs	Ē
TREC	75	375K	0	7
SIGIR	60	204K	3870	
0 = would not refer patient to trial 1 = would consider referring with further investigation 2 = highly likely to refer patient to trial				

Neural Reranker : Data Generation

Patient is a 50-year-old left-handed man who is here for a follow up of his left sphenoid meningioma. He presented electively for left sided craniotomy for mass resection. His neurological problem began last year when he became confused and disoriented in a hotel bathroom. At that time, he was visiting his daughter for a wedding. His wife found him slumped over in the bath tube. According to her, his eyes looked funny. He could not stand up. His verbal output did not make sense. He was brought to Hospital in Placentia, CA. He woke up 7 to 8 hours later in the emergency room. He felt very tired after the event. He was hospitalized from [**2142-6-22**] to [**2142-6-25**]. He had a cardiac pacemaker placement due to irregular heart rate and bradycardia. He also had a head MRI that showed a less than 1 cm diameter sphenoid meningioma.

DIAGNOSIS CODES

388887,42130,114236,1,"2252", meningioma 388888,42130,114236,2,"4019", hypertension 388889,42130,114236,3,"42731",atrial fibrillation

PROCEDURE CODES 238000,42130,114236,1,"0151",craniotomy

Showing: 1-100 of 180 studies 100 Studies per page						Show/Hide Columns
Row	Row Saved Status Conditions Interventions Lo					
1			The Use of Eye Patches and Earplugs in Intensive Care in Cases of Craniotomy .	Craniotomy	Device: eye patch and earplug	
		recruiting				

Showing: 1-100 of 214 studies 100 Show						Show/Hide Columns
Row	Row Saved Status Conditions Interventions					
1		Completed	Meningiomas and Treatment With CYPROTERONE ACETATE or Progestin	Meningioma		Département de Nutrition - CHRU de Brest Brest, France

Neural Reranker : Data Generation

Patient is a 50-year-old left-handed man who is here for a follow up of his left sphenoid meningioma. He presented electively for left sided craniotomy for mass resection. His neurological problem began last year when he became confused and disoriented in a hotel bathroom. At that time, he was visiting his daughter for a wedding. His wife found him slumped over in the bath tube. According to her, his eyes looked funny. He could not stand up. His verbal output did not make sense. He was brought to Hospital in Placentia, CA. He woke up 7 to 8 hours later in the emergency room. He felt very tired after the event. He was hospitalized from [**2142-6-22**] to [**2142-6-25**]. He had a cardiac pacemaker placement due to irregular heart rate and bradycardia. He also had a head MRI that showed a less than 1 cm diameter sphenoid meningioma.

DIAGNOSIS CODES

388887,42130,114236,1,"2252", meningioma 388888,42130,114236,2,"4019", hypertension 388889,42130,114236,3,"42731",atrial fibrillation

PROCEDURE CODES 238000,42130,114236,1,"0151",craniotomy

Showing: 1-100 of 180 studies 100 Studies per page						Show/Hide Columns
Row	Row Saved Status Conditions Interventions Local					
1		Not yet recruiting	The Use of Eye Patches and Earplugs in Intensive Care in Cases of Craniotomy .	Craniotomy	Device: eye patch and earplug	

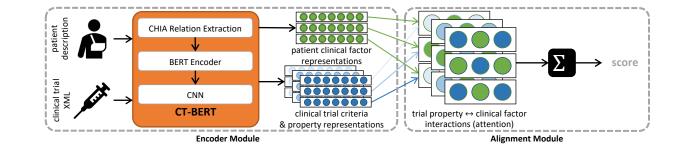
Showing: 1-100 of 214 studies 100 Show/Hide						Show/Hide Columns
Row	Row Saved Status Conditions Interventions					
1		Completed	Meningiomas and Treatment With CYPROTERONE ACETATE or Progestin	Meningioma		Département de Nutrition - CHRU de Brest
						Brest, France

Use MIMIC III to generate **more than 700K** silver-standard training pairs

Johnson, A. E., Pollard, T. J., Shen, L., Li-Wei, H. L., Feng, M., Ghassemi, M., ... & Mark, R. G. (2016). MIMIC-III, a freely accessible critical care database. Scientific data, 3(1), 1-9.

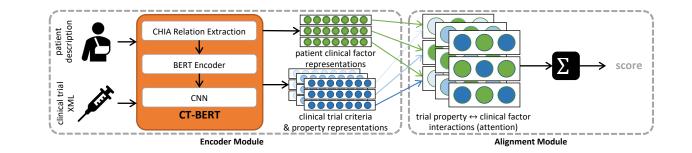
Neural Reranker

Overview and Inspiration



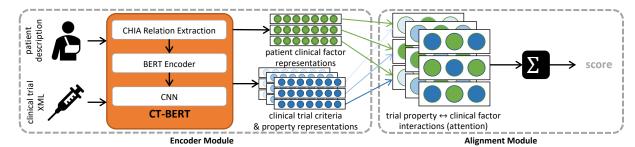
Overview and Inspiration

 The model identifies + encodes meaningful topic/trial information



Overview and Inspiration

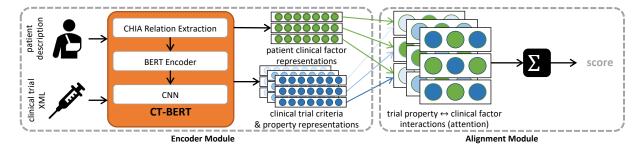
 The model identifies + encodes meaningful topic/trial information



2. The model computes alignment scores between trials and topics to find matches

Overview and Inspiration

 The model identifies + encodes meaningful topic/trial information



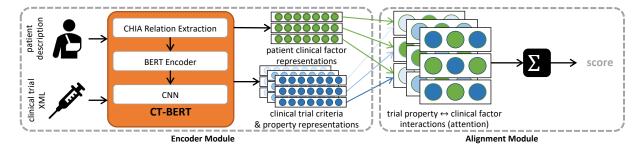
2. The model computes alignment scores between trials and topics to find matches

Inspired by DeepEnroll; uses natural language inference model to determine if trials match patients.

Zhang, X., Xiao, C., Glass, L. M., & Sun, J. (2020, April). Deepenroll: Patient-trial matching with deep embedding and entailment prediction. In Proceedings of The Web Conference 2020 (pp. 1029-1037).

Overview and Inspiration

 The model identifies + encodes meaningful topic/trial information

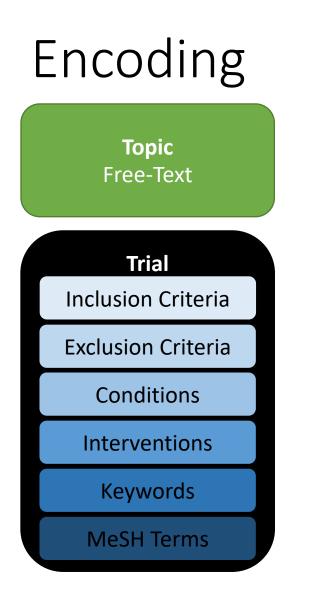


2. The model computes alignment scores between trials and topics to find matches

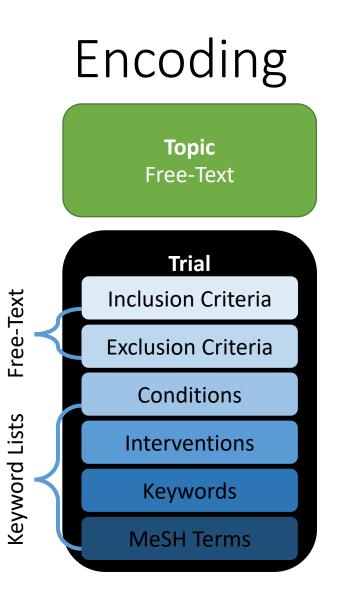
Inspired by DeepEnroll; uses natural language inference model to determine if trials match patients.

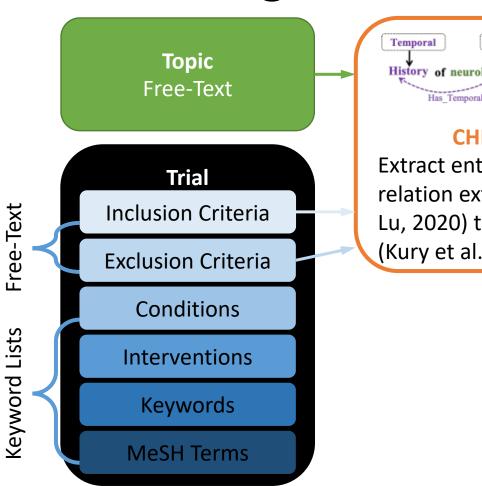
Differs from us in that they use structured EHR!







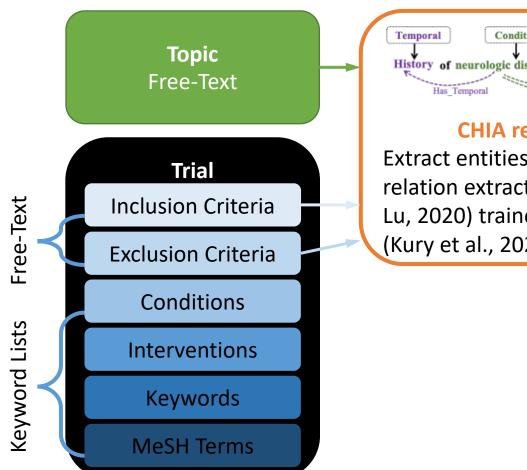




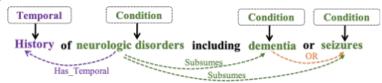
Encoding

Temporal Condition Condition Condition History of neurologic disorders including dementia or seizures OR CHIA relation extraction Extract entities with joint entityrelation extraction model (Wang and Lu, 2020) trained on CHIA annotations (Kury et al., 2020)

Neural Reranker



Encoding



CHIA relation extraction

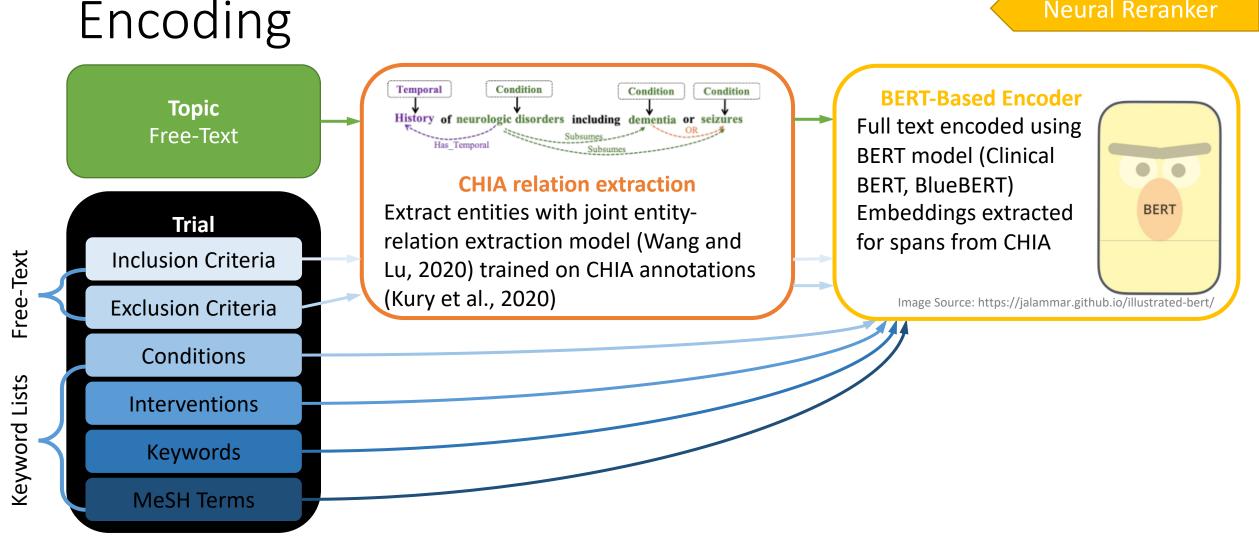
Extract entities with joint entityrelation extraction model (Wang and Lu, 2020) trained on CHIA annotations (Kury et al., 2020)

BERT-Based Encoder

Full text encoded using BERT model (Clinical BERT, BlueBERT) Embeddings extracted for spans from CHIA BERT

Image Source: https://jalammar.github.io/illustrated-bert/

Neural Reranker



Kury, F., Butler, A., Yuan, C., Fu, L. H., Sun, Y., Liu, H., ... & Weng, C. (2020). Chia, a large annotated corpus of clinical trial eligibility criteria. Scientific data, 7(1), 1-11. Wang, J., & Lu, W. (2020). Two are better than one: Joint entity and relation extraction with table-sequence encoders. arXiv preprint arXiv:2010.03851.

Methodology **Neural Reranker** Encoding Condition Temporal Condition Condition **BERT-Based Encoder** Topic History of neurologic disorders including dementia or seizures Full text encoded using Free-Text Has Temporal **BERT model (Clinical CHIA** relation extraction **BERT, BlueBERT) Embeddings** extracted Extract entities with joint entity-BERT Trial relation extraction model (Wang and for spans from CHIA **Inclusion Criteria** Lu, 2020) trained on CHIA annotations (Kury et al., 2020) Image Source: https://jalammar.github.io/illustrated-bert/ **Exclusion Criteria** Conditions Interventions Keywords

CNN

We use a CNN to handle spans with

have multiple tokens

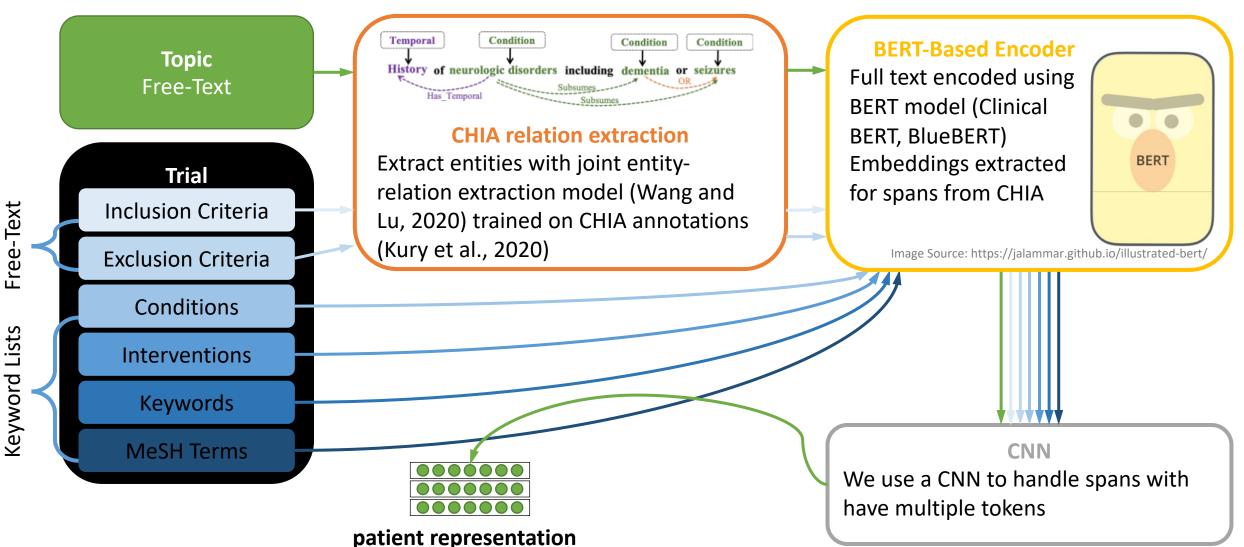
Kury, F., Butler, A., Yuan, C., Fu, L. H., Sun, Y., Liu, H., ... & Weng, C. (2020). Chia, a large annotated corpus of clinical trial eligibility criteria. Scientific data, 7(1), 1-11. Wang, J., & Lu, W. (2020). Two are better than one: Joint entity and relation extraction with table-sequence encoders. arXiv preprint arXiv:2010.03851.

Free-Text

Keyword Lists

MeSH Terms

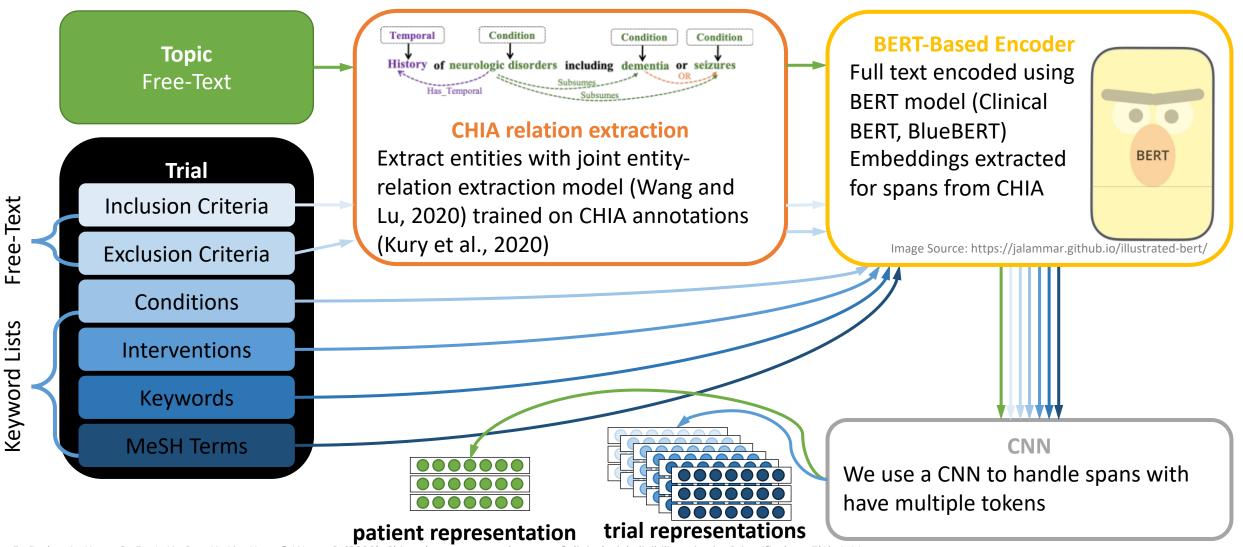
Methodology Neural Reranker



Kury, F., Butler, A., Yuan, C., Fu, L. H., Sun, Y., Liu, H., ... & Weng, C. (2020). Chia, a large annotated corpus of clinical trial eligibility criteria. *Scientific data*, 7(1), 1-11. Wang, J., & Lu, W. (2020). Two are better than one: Joint entity and relation extraction with table-sequence encoders. *arXiv preprint arXiv:2010.03851*.

Encoding

Methodology Neural Reranker

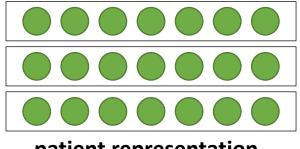


Kury, F., Butler, A., Yuan, C., Fu, L. H., Sun, Y., Liu, H., ... & Weng, C. (2020). Chia, a large annotated corpus of clinical trial eligibility criteria. *Scientific data*, 7(1), 1-11. Wang, J., & Lu, W. (2020). Two are better than one: Joint entity and relation extraction with table-sequence encoders. *arXiv preprint arXiv:2010.03851*.

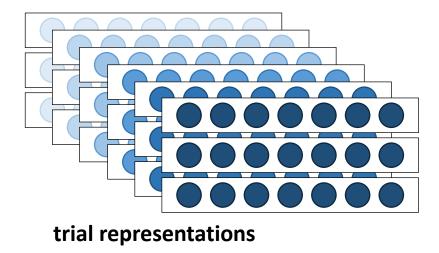
Encoding

Neural Reranker

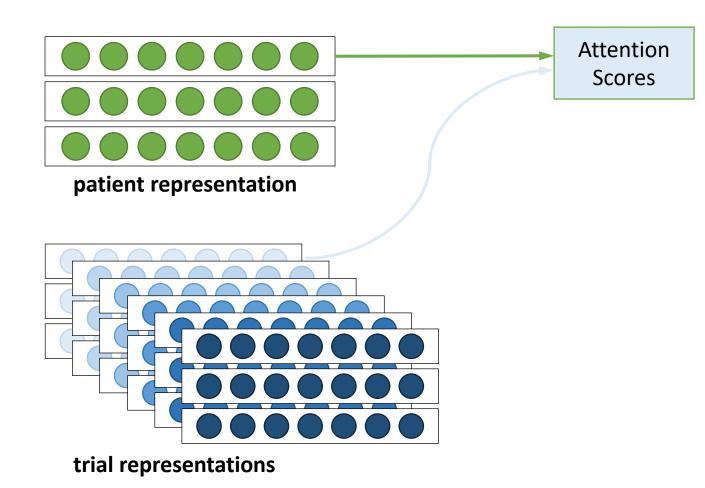
Alignment



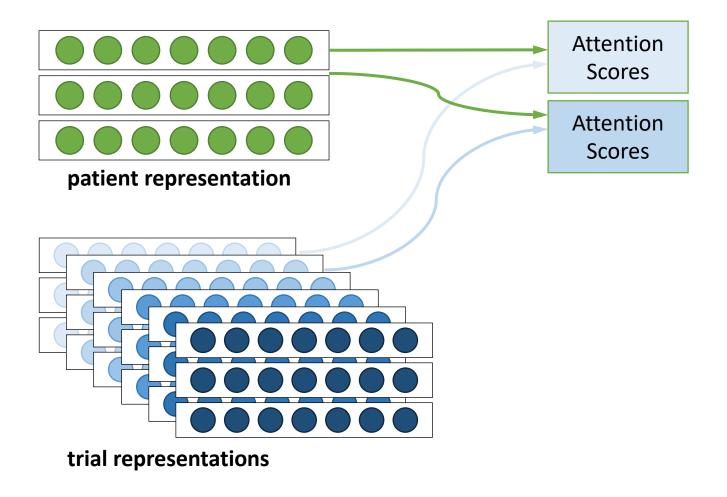
patient representation



Neural Reranker

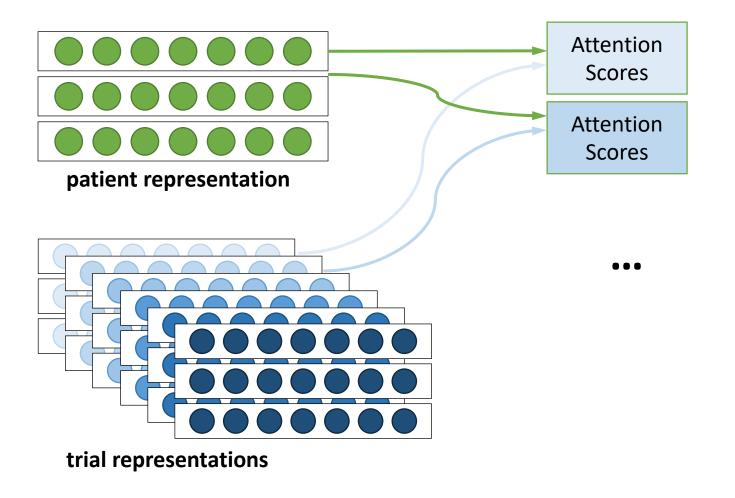


Neural Reranker



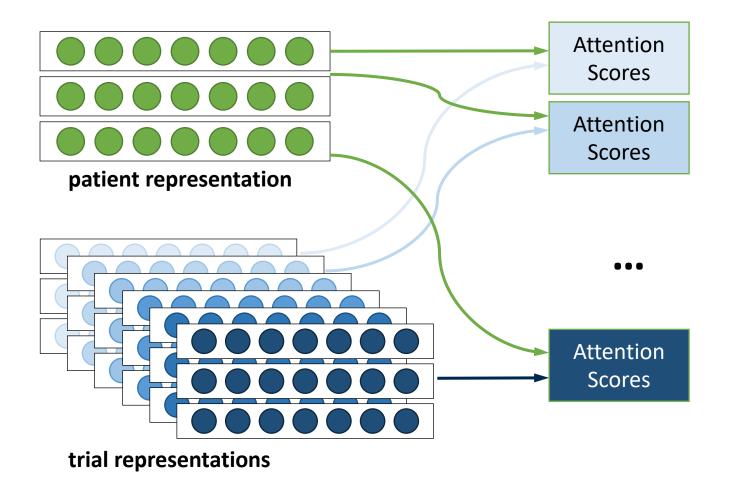
Neural Reranker

Alignment

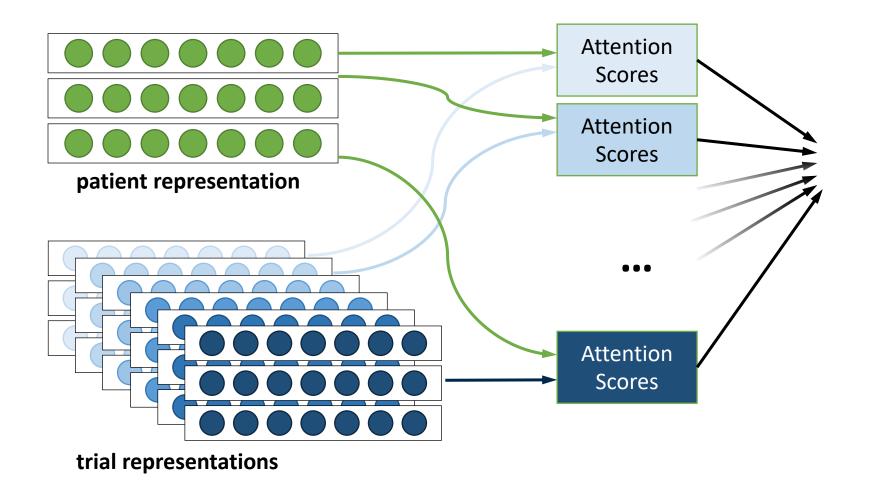


Weston, J., Bengio, S., & Usunier, N. (2011, June). Wsabie: Scaling up to large vocabulary image annotation. In Twenty-Second International Joint Conference on Artificial Intelligence.

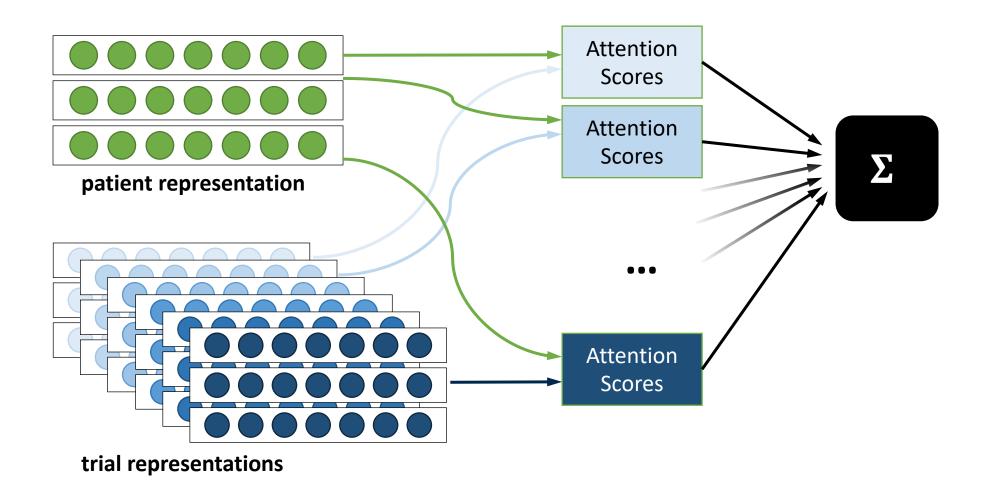
Neural Reranker



Neural Reranker

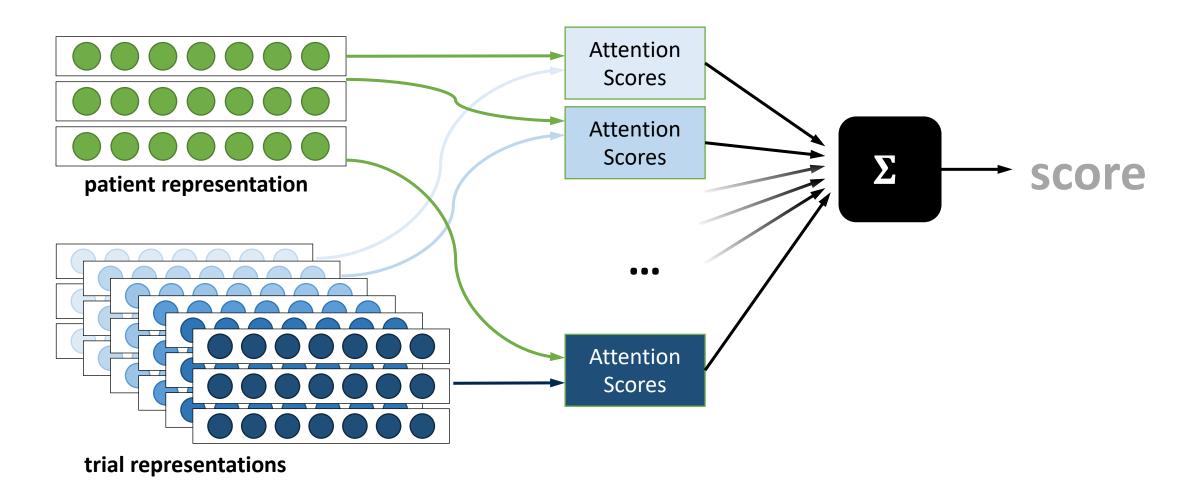


Neural Reranker



Neural Reranker

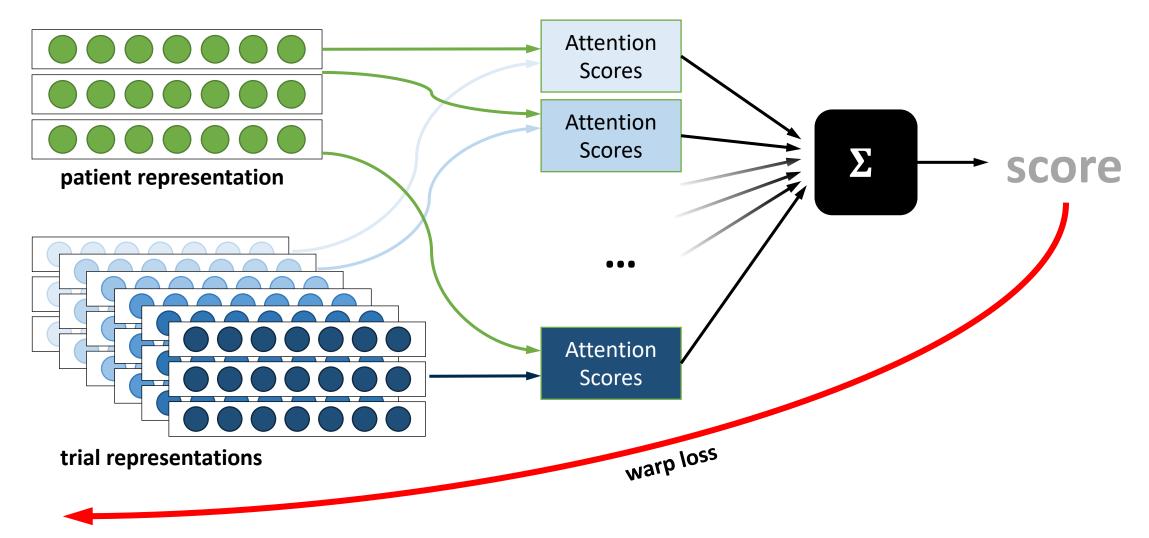
Alignment



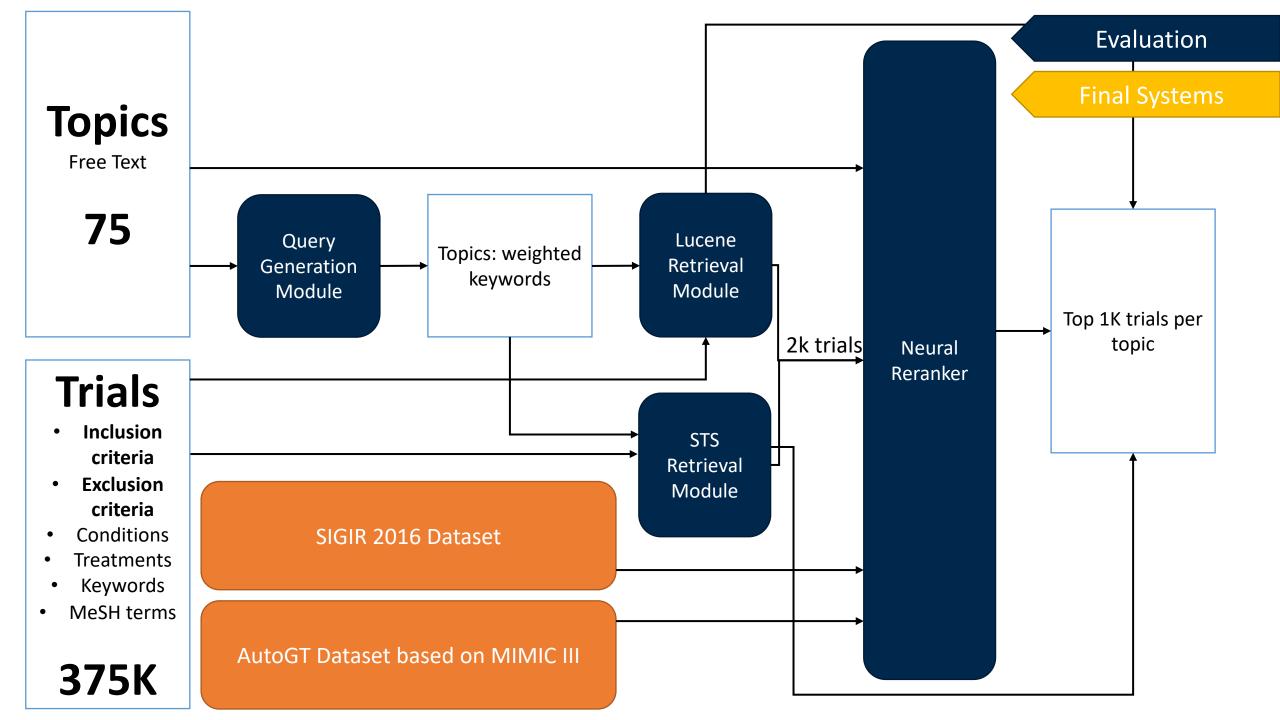
Weston, J., Bengio, S., & Usunier, N. (2011, June). Wsabie: Scaling up to large vocabulary image annotation. In Twenty-Second International Joint Conference on Artificial Intelligence.

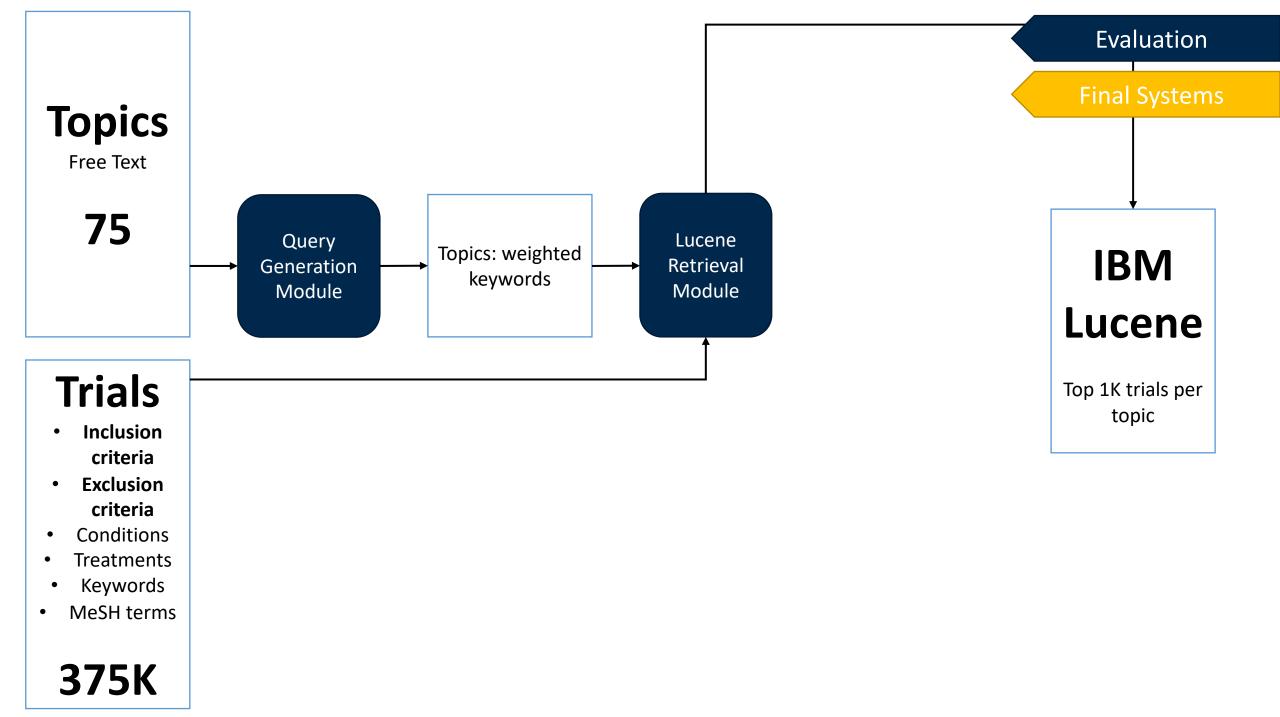
Neural Reranker

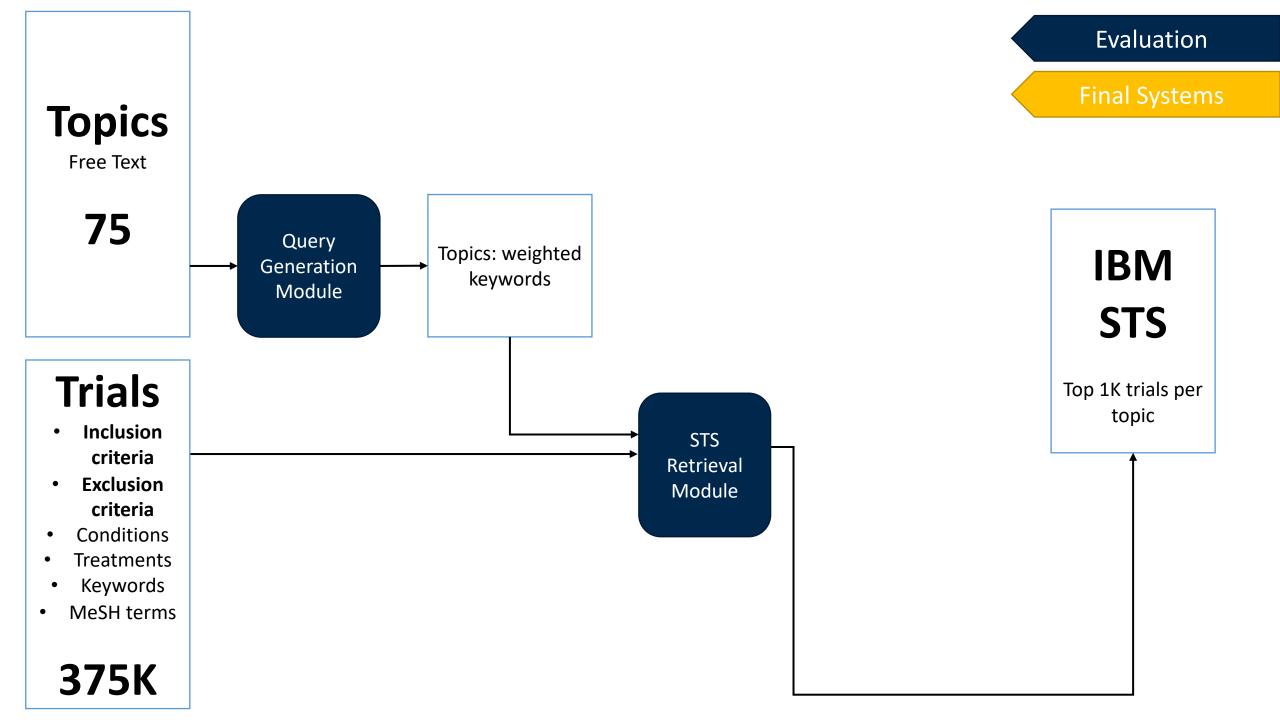
Alignment

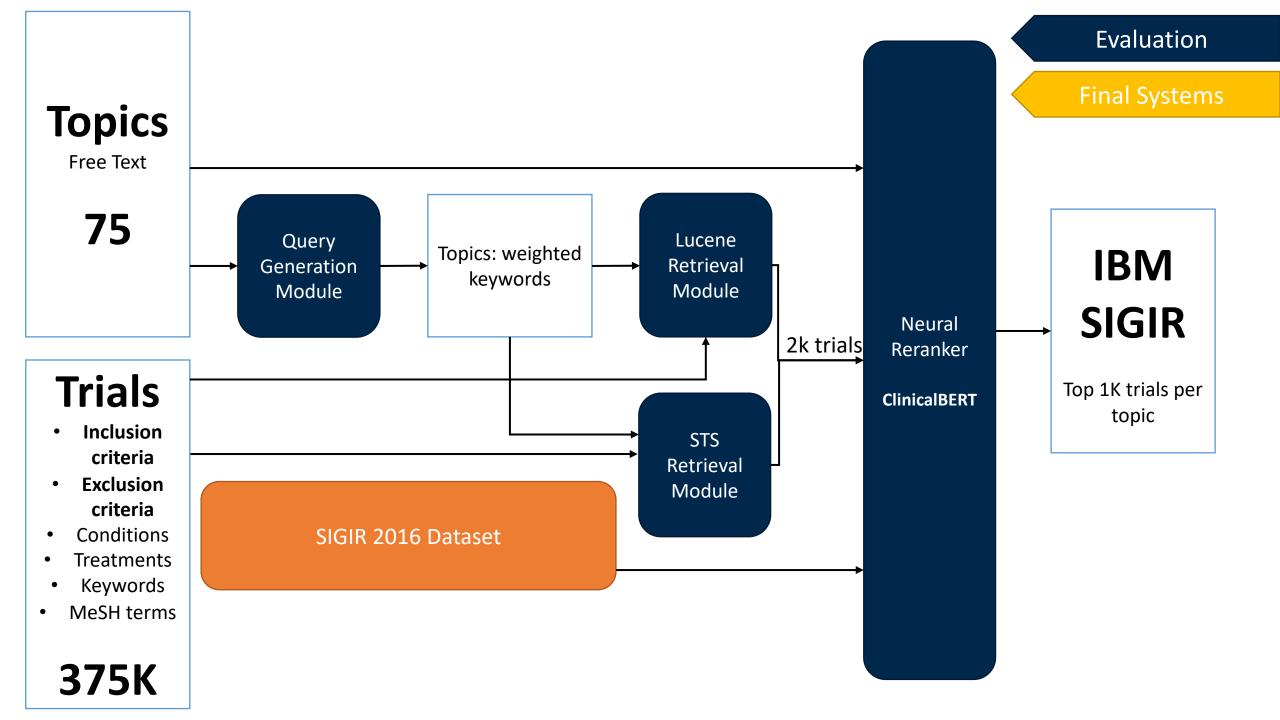


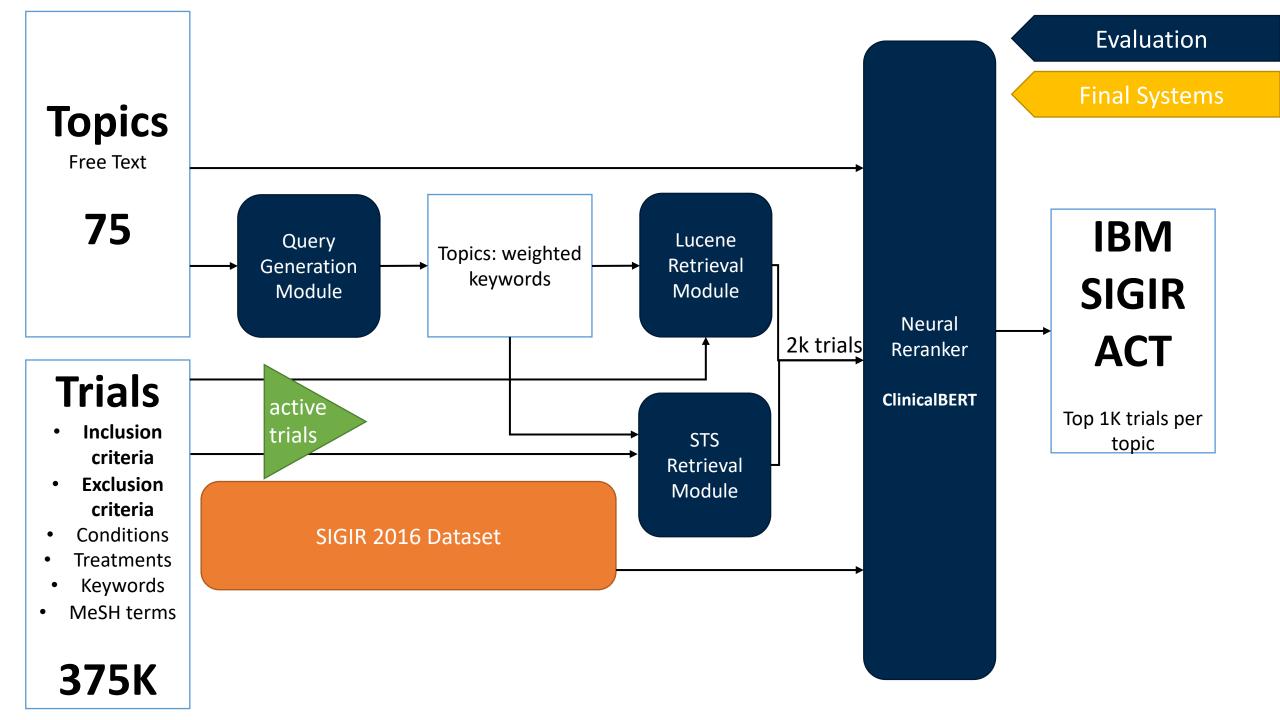
Weston, J., Bengio, S., & Usunier, N. (2011, June). Wsabie: Scaling up to large vocabulary image annotation. In Twenty-Second International Joint Conference on Artificial Intelligence.

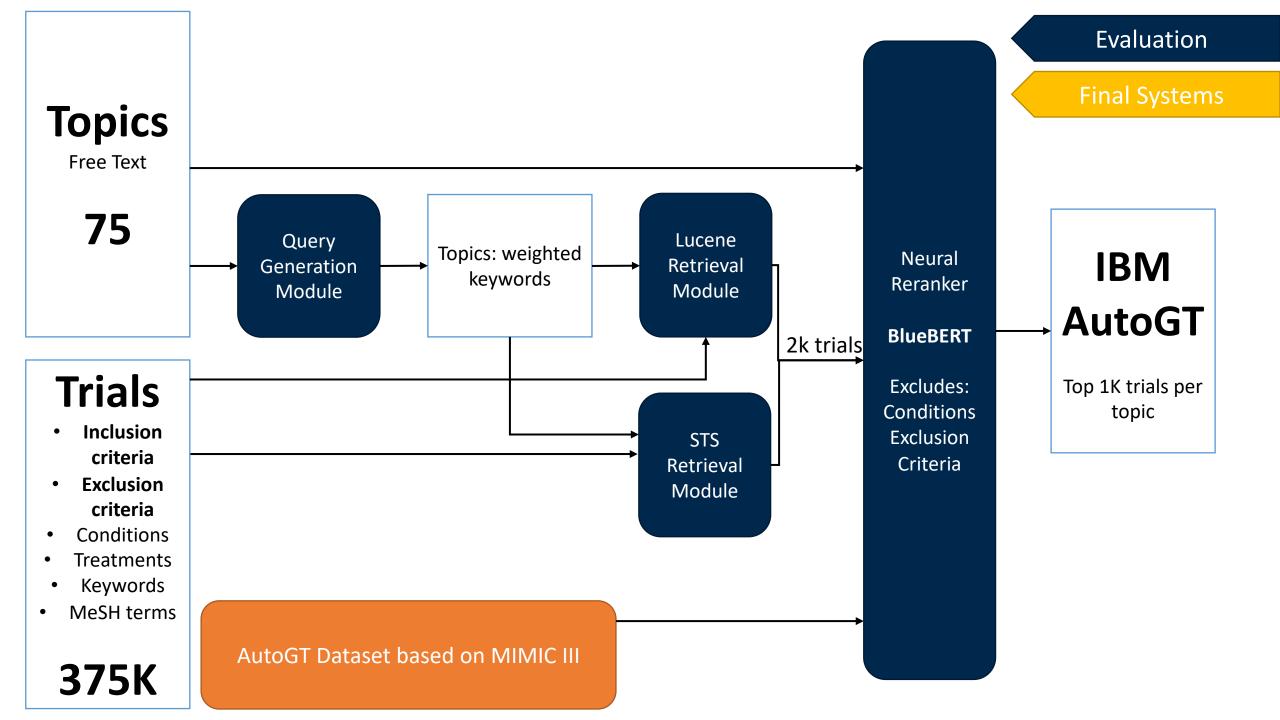












System	NDCG@10	PREC@10	Reciprocal Rank
IBMLucene	.32	.20	.39
IBMSTS	.22	.15	.27
IBM SIGIR	.14	.09	.19
IBMAUTOGT	.13	.09	.14
IBMSIGIRACT		.06	.13

System	NDCG@10	PREC@10	Reciprocal Rank
IBMLucene	.32	.20	.39
IBMSTS	.22	.15	.27
IBM SIGIR	.14	.09	.19
IBMAUTOGT	.13	.09	.14
IBMSIGIRACT		.06	.13

System	NDCG@10	PREC@10	Reciprocal Rank
IBMLucene	.32	.20	.39
IBMSTS	.22	.15	.27
IBM SIGIR	.14	.09	.19
IBMAUTOGT	.13	.09	.14
IBMSIGIRACT		.06	.13

System	NDCG@10	PREC@10	Reciprocal Rank
IBMLucene	.32	.20	.39
IBMSTS	.22	.15	.27
IBM SIGIR	.14	.09	.19
IBMAUTOGT	.13	.09	.14
IBMSIGIRACT		.06	.13

Conclusions

• In addition to the small size, the differing labeling scheme of the SIGIR doesn't capture the inclusion vs. exclusion labels

- In addition to the small size, the differing labeling scheme of the SIGIR doesn't capture the inclusion vs. exclusion labels
- Silver-standard training data does not account for nuances in criteria, only accounting for conditions
 - These are likely already handled well by our retrieval module

- In addition to the small size, the differing labeling scheme of the SIGIR doesn't capture the inclusion vs. exclusion labels
- Silver-standard training data does not account for nuances in criteria, only accounting for conditions
 - These are likely already handled well by our retrieval module
- More attention to factors such as patient demographics would likely help our approach

- In addition to the small size, the differing labeling scheme of the SIGIR doesn't capture the inclusion vs. exclusion labels
- Silver-standard training data does not account for nuances in criteria, only accounting for conditions
 - These are likely already handled well by our retrieval module
- More attention to factors such as patient demographics would likely help our approach
- Evaluation was challenging without gold standard data