IBM @ TREC Clinical Trials Track 2021

Laura Biester

Venkata Joopudi

Bharath Dandala

Topics

Free Text

75

Trials

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

375K

Top 1K trials per topic

Topics

Free Text

75

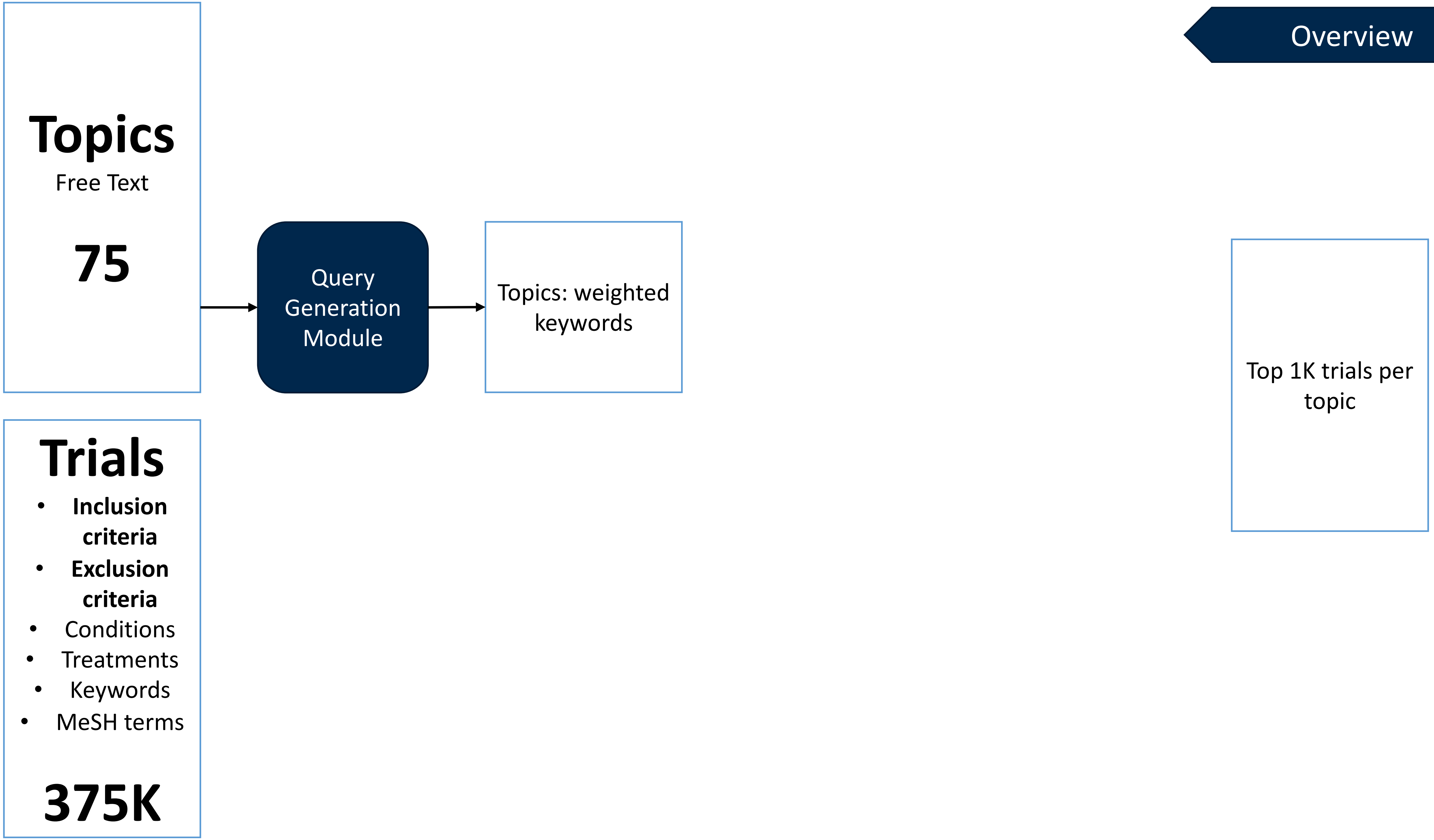
Trials

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

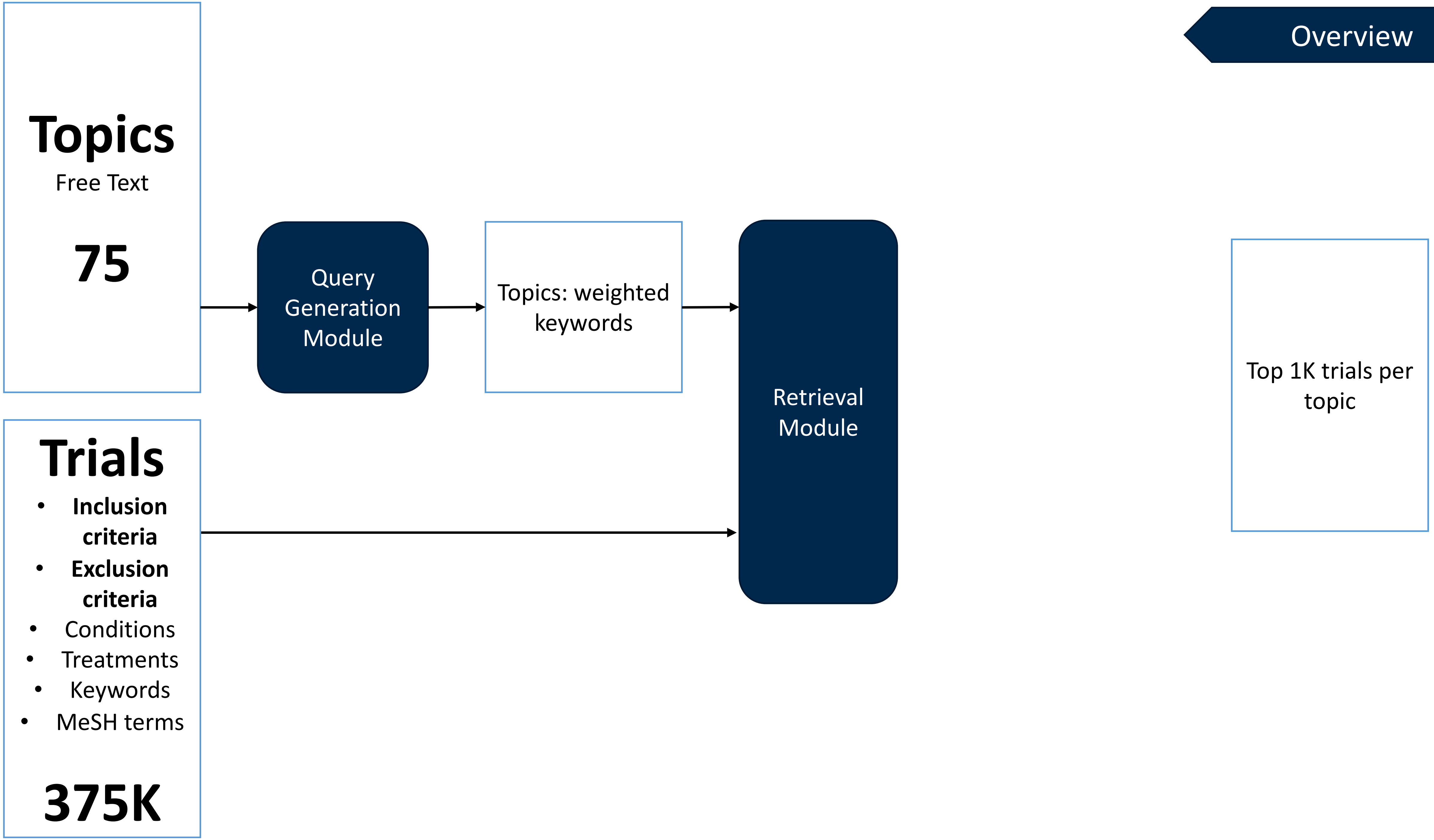
375K

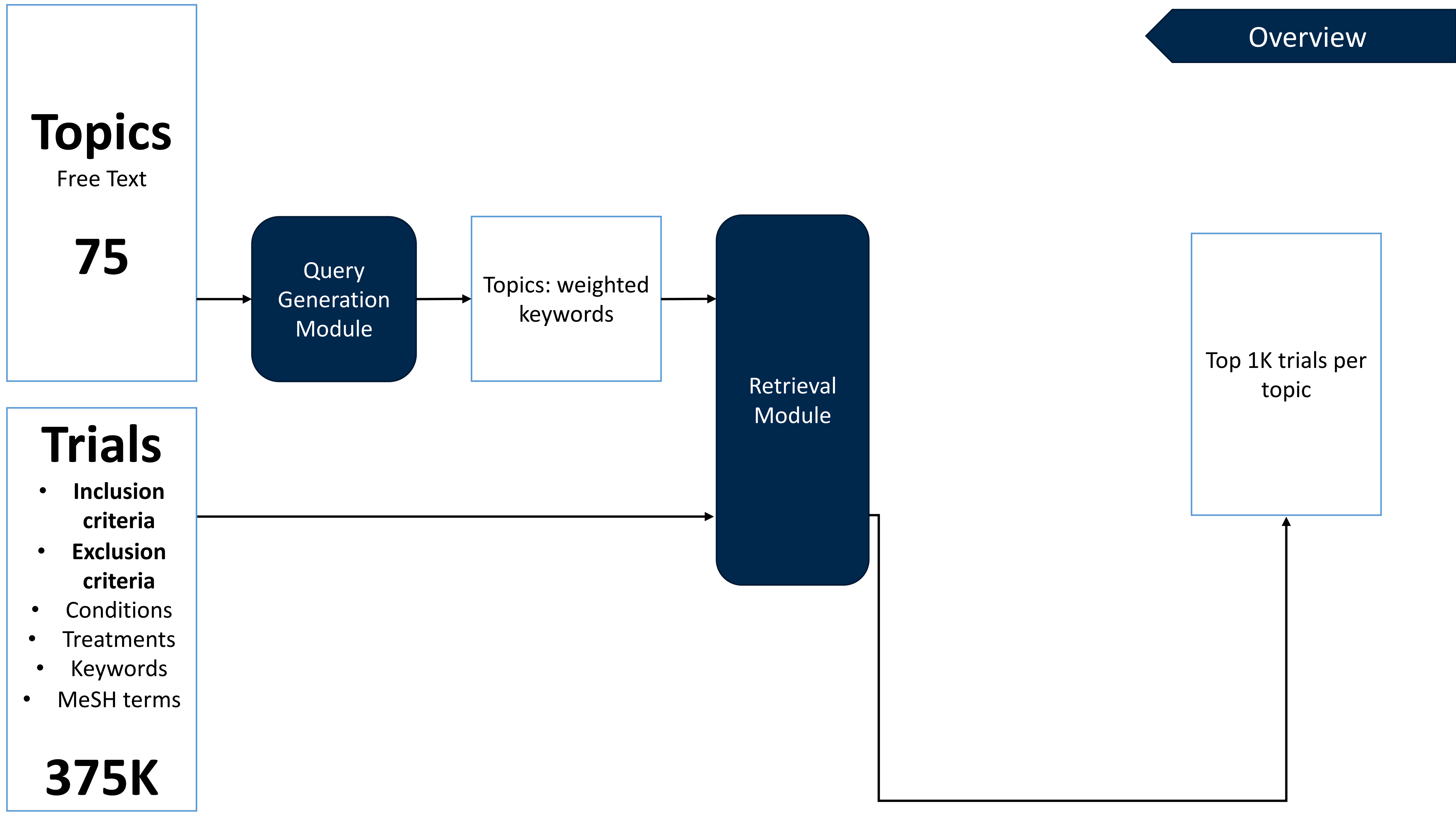
Top 1K trials per topic

Overview


Topics
Free Text
75
Trials
• Inclusion criteria
• Exclusion criteria
• Conditions
• Treatments
• Keywords
• MeSH terms
375K
Query Generation Module
Topics: weighted keywords
Top 1K trials per topic

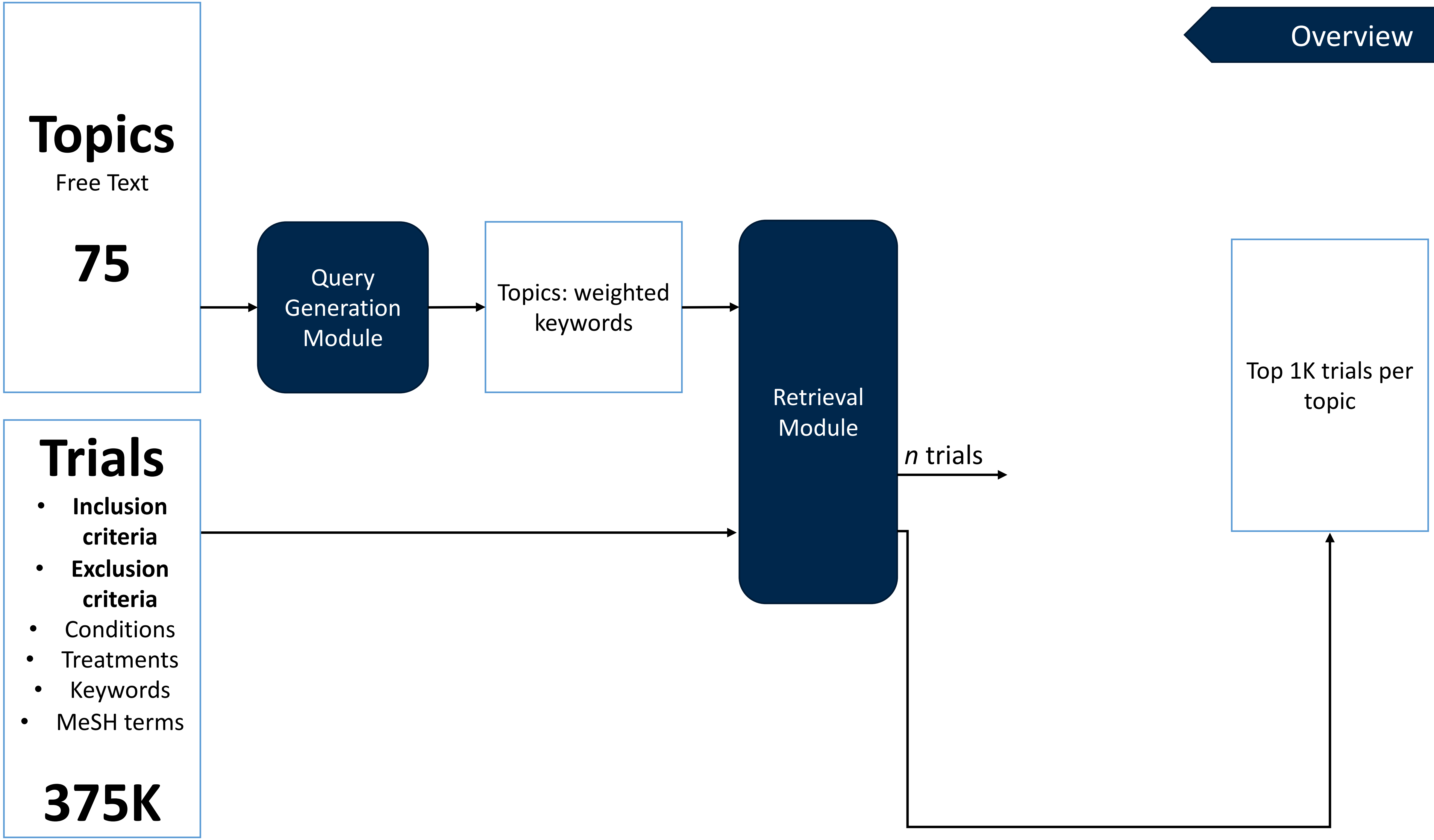
Overview

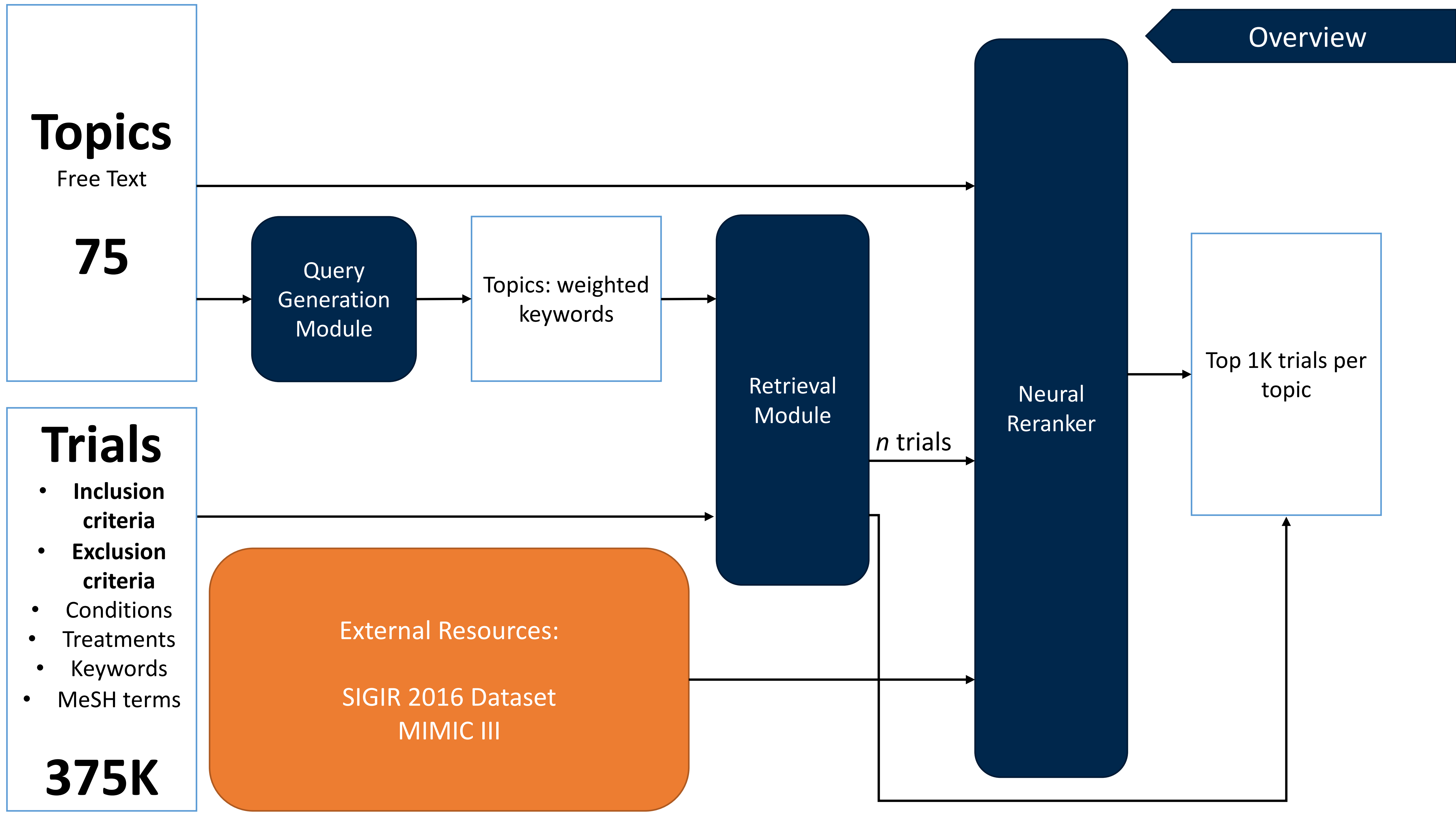

Topics
Free Text
75
Query Generation Module
Topics: weighted keywords
Retrieval Module
Top 1K trials per topic
Trials
• Inclusion criteria
• Exclusion criteria
• Conditions
• Treatments
• Keywords
• MeSH terms
375K



Overview
Topics
Free Text
75
Query Generation Module
Topics: weighted keywords
Retrieval Module
Top 1K trials per topic
Trials
• Inclusion criteria
• Exclusion criteria
• Conditions
• Treatments
• Keywords
• MeSH terms
375K

Overview


Topics
Free Text
75
Query Generation Module
Topics: weighted keywords
Retrieval Module
n trials
Top 1K trials per topic
Trials
• Inclusion criteria
• Exclusion criteria
• Conditions
• Treatments
• Keywords
• MeSH terms
375K


Overview
Topics
Free Text
75
Query Generation Module
Topics: weighted keywords
Retrieval Module
n trials
Neural Reranker
Top 1K trials per topic
Trials
• Inclusion criteria
• Exclusion criteria
• Conditions
• Treatments
• Keywords
• MeSH terms
375K
External Resources:
SIGIR 2016 Dataset
MIMIC III

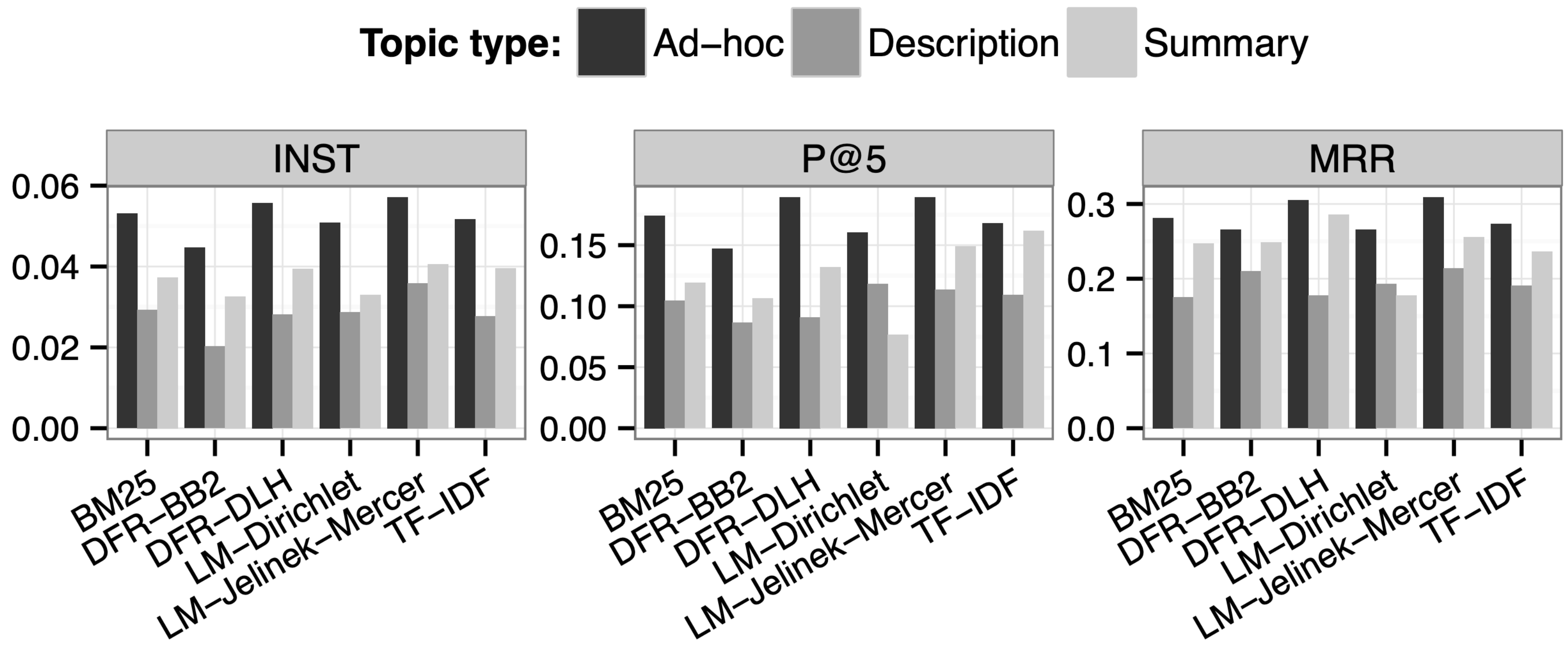
Ad-hoc Query Inspiration

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

Koopman, B., & Zuccon, G. (2016, July). A test collection for matching patients to clinical trials. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval* (pp. 669-672).

Ad-hoc Query Inspiration

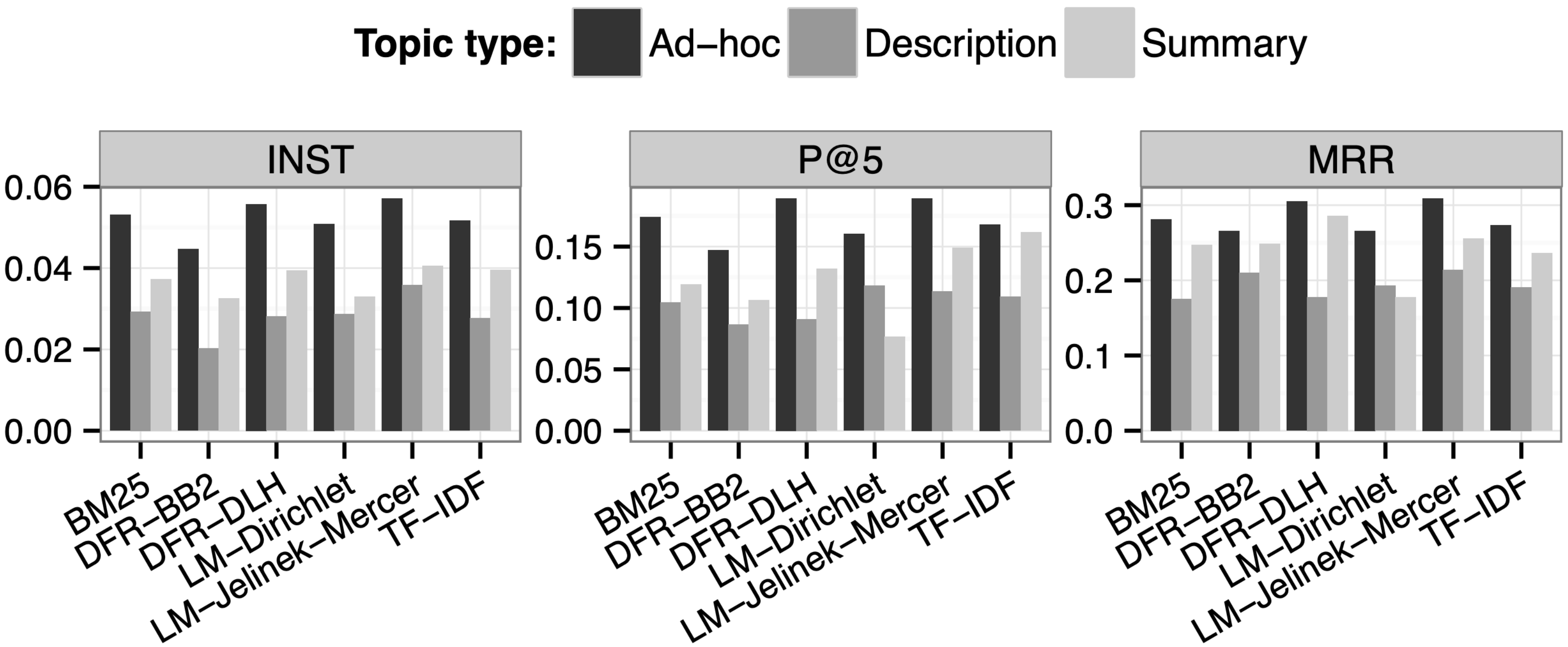
Koopman and Zuccon (2016) use **ad-hoc queries** for retrieval



Koopman, B., & Zuccon, G. (2016, July). A test collection for matching patients to clinical trials. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval* (pp. 669-672).

Ad-hoc Query Inspiration

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



Can we mimic this process with automated methods?

Koopman, B., & Zuccon, G. (2016, July). A test collection for matching patients to clinical trials. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval* (pp. 669-672).

Overview

Step 1

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

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

Overview

Step 1

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

Step 2

- Using *external resources*, identify unique diseases

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

Overview

Step 1
- Extract pertinent information using IBM Watson's Annotator for Clinical Data (ACD)

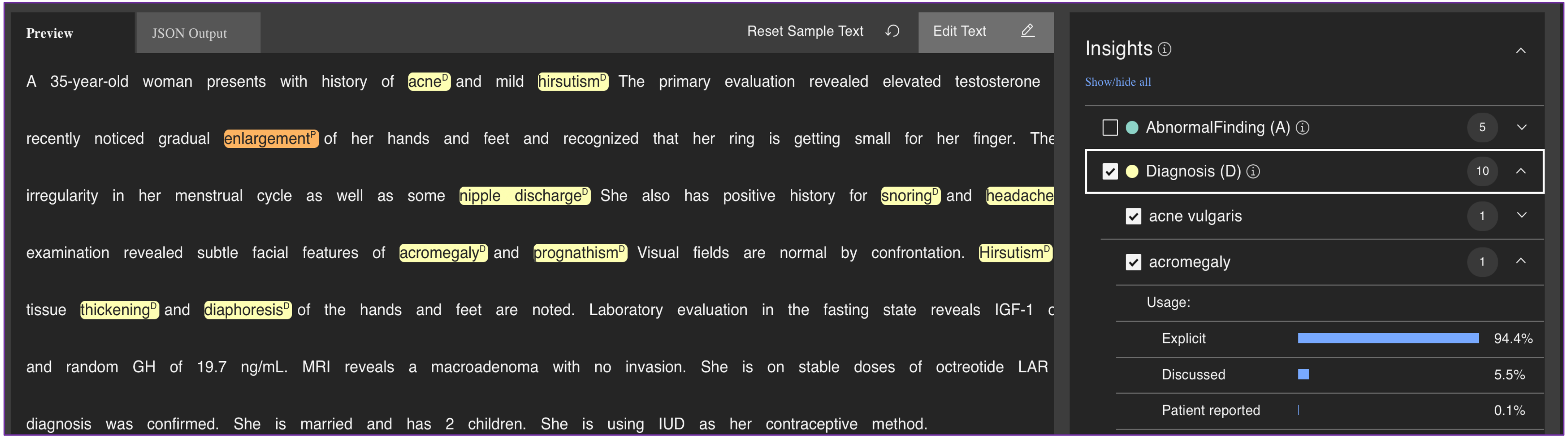
Step 2
- Using *external resources*, identify unique diseases

Step 3
- Normalize and standardize features, using the resulting sum as query weight

Methodology

Query Generation

Query Example

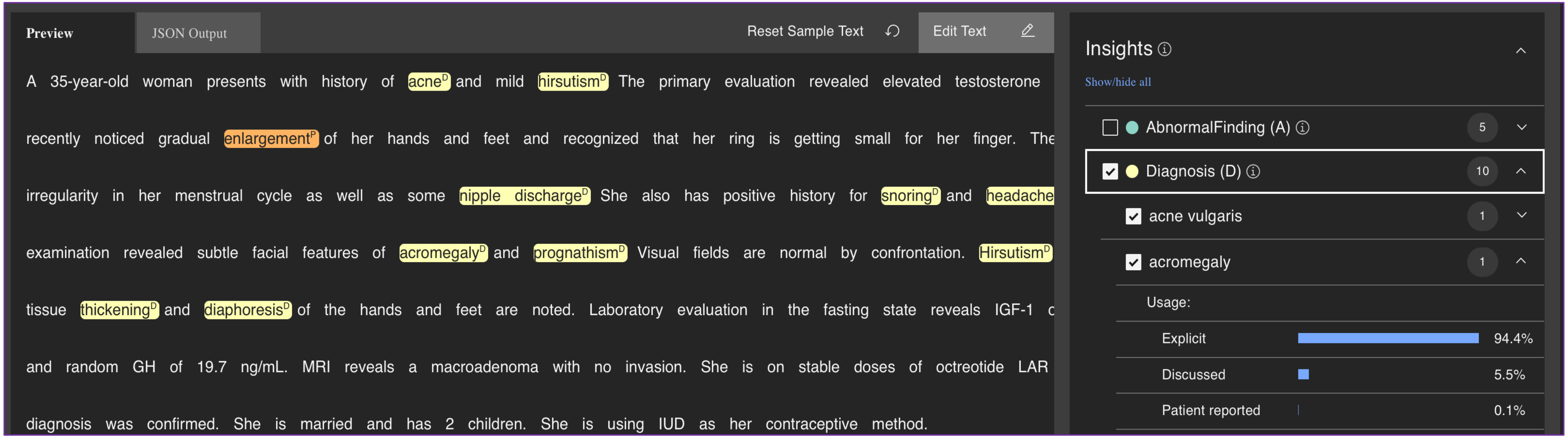


https://acd-try-it-out.mybluemix.net/preview

Methodology

Query Generation

Query Example



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

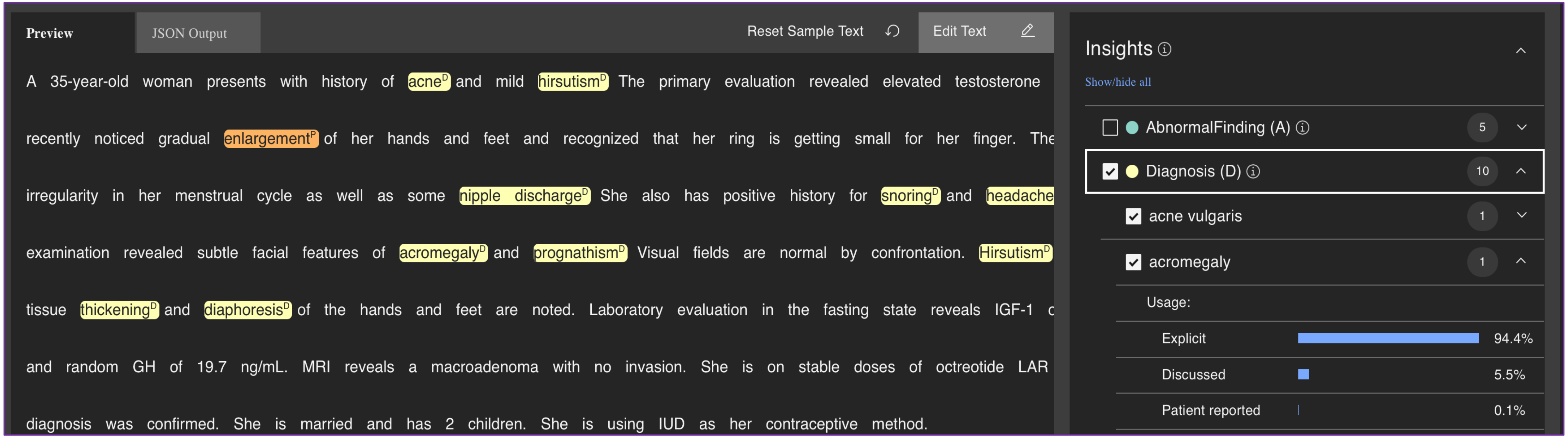
extracted entities

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

Methodology

Query Generation

Query Example



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

extracted entities

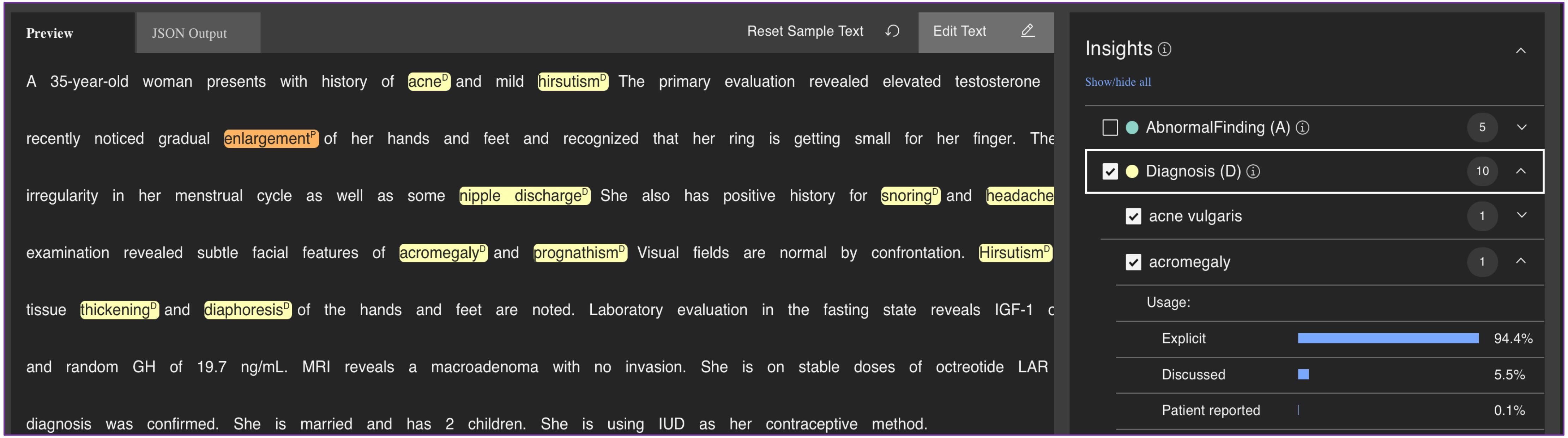
- acromegaly
- congenital prognathism
- nipple discharge
- hirsutism
- acne vulgaris
- snoring
- excessive sweating
- increased thickness
- headache

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

Methodology

Query Generation

Query Example



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

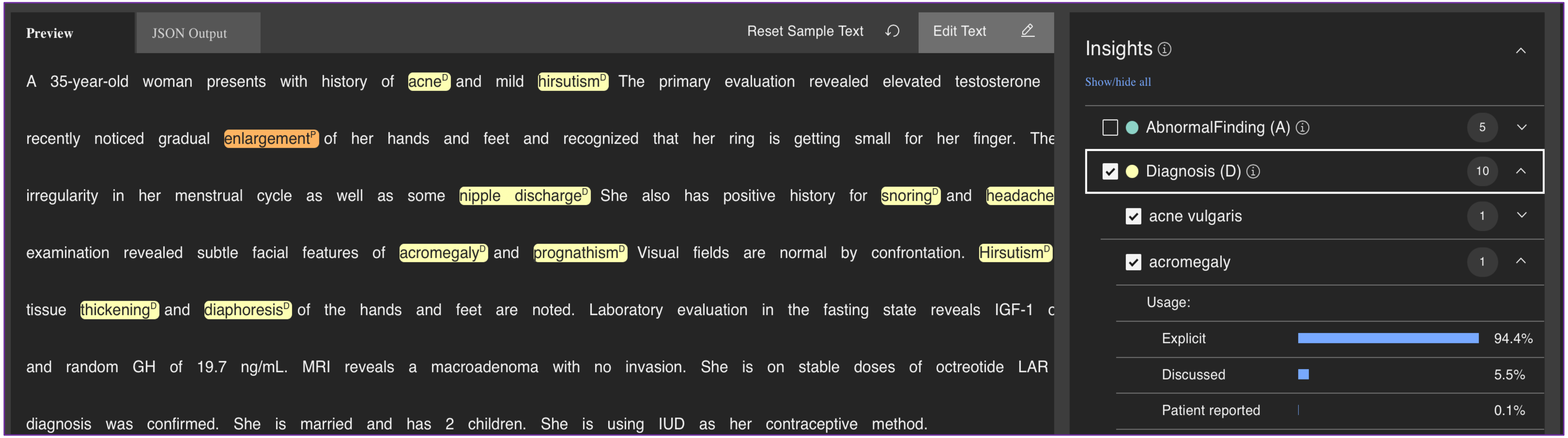
extracted entities

- 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

Standardization and normalization

Query Example



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

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

Standardization and normalization

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

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

Robertson, S. E., Walker, S., Jones, S., Hancock-Beaulieu, M. M., & Gatford, M. (1995). Okapi at TREC-3. *Nist Special Publication Sp*, *109*, 109.

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

Robertson, S. E., Walker, S., Jones, S., Hancock-Beaulieu, M. M., & Gatford, M. (1995). Okapi at TREC-3. *Nist Special Publication Sp*, *109*, 109.

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

Robertson, S. E., Walker, S., Jones, S., Hancock-Beaulieu, M. M., & Gatford, M. (1995). Okapi at TREC-3. *Nist Special Publication Sp*, *109*, 109.

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

Koopman, B., & Zuccon, G. (2016, July). A test collection for matching patients to clinical trials. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval* (pp. 669-672).

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

0 = would not refer patient to trial
1 = would consider referring with further investigation
2 = highly likely to refer patient to trial

Koopman, B., & Zuccon, G. (2016, July). A test collection for matching patients to clinical trials. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval* (pp. 669-672).

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

0 = would not refer patient to trial
1 = would consider referring with further investigation
2 = highly likely to refer patient to trial

Koopman, B., & Zuccon, G. (2016, July). A test collection for matching patients to clinical trials. In *Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval* (pp. 669-672).

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	Saved	Status	Study Title	Conditions	Interventions	Locations
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 studies per page Show/Hide Columns

Row	Saved	Status	Study Title	Conditions	Interventions	Locations
1		Completed	**Meningiomas** and Treatment With CYPROTERONE ACETATE or Progestin	• **Meningioma**		• Département de Nutrition - CHRU de Brest Brest, France

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 : 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	Saved	Status	Study Title	Conditions	Interventions	Locations
1		Not yet recruiting	The Use of Eye Patches and Earplugs in Intensive Care in Cases of **Craniotomy**.	• **Craniotomy**	• Device: eye patch and earplug	

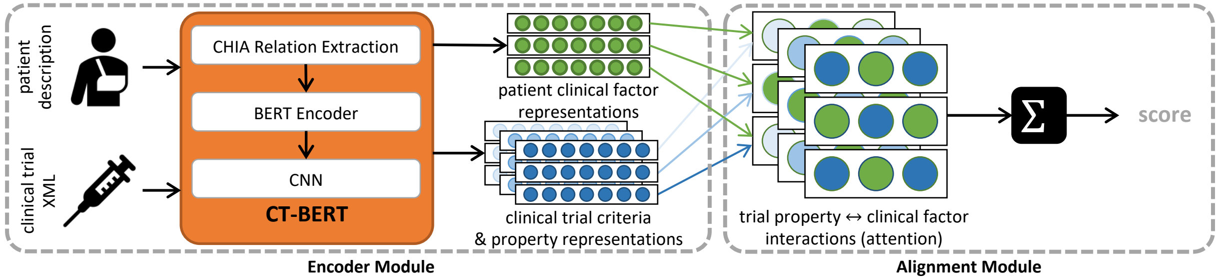
Showing: 1-100 of **214** studies 100 studies per page Show/Hide Columns

Row	Saved	Status	Study Title	Conditions	Interventions	Locations
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.

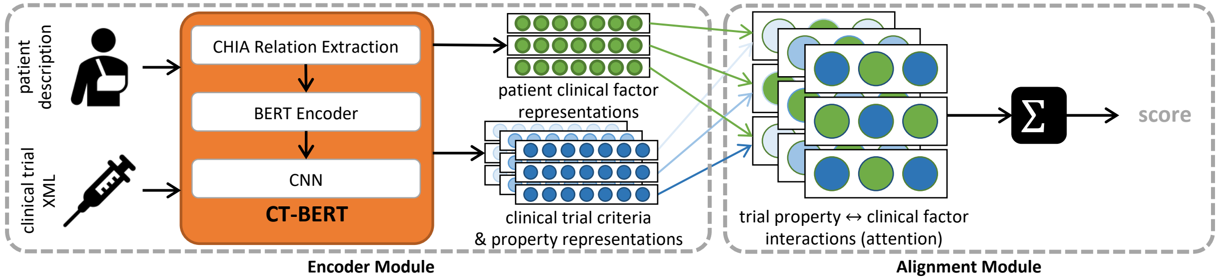
Overview and Inspiration



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

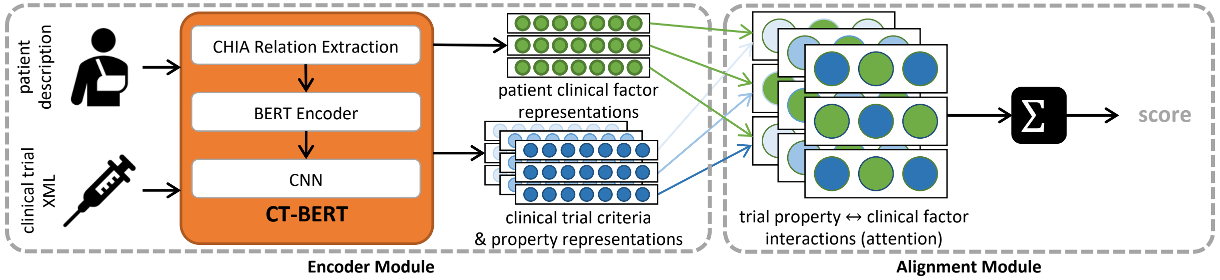
1. The model identifies + encodes meaningful topic/trial information



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

1. The model identifies + encodes meaningful topic/trial information

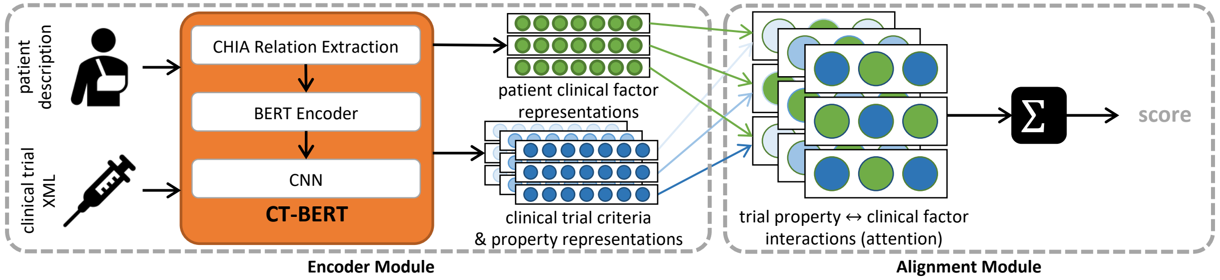


2. The model computes alignment scores between trials and topics to find matches

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

1. The model identifies + encodes meaningful topic/trial information



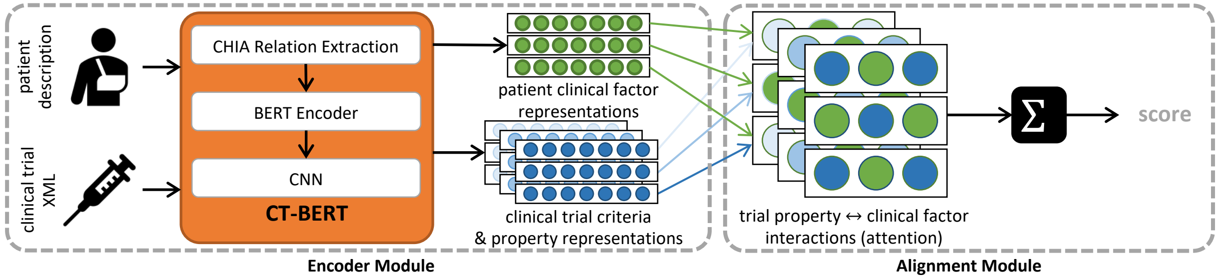
2. The model computes alignment scores between trials and topics to find matches

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

1. 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!

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).

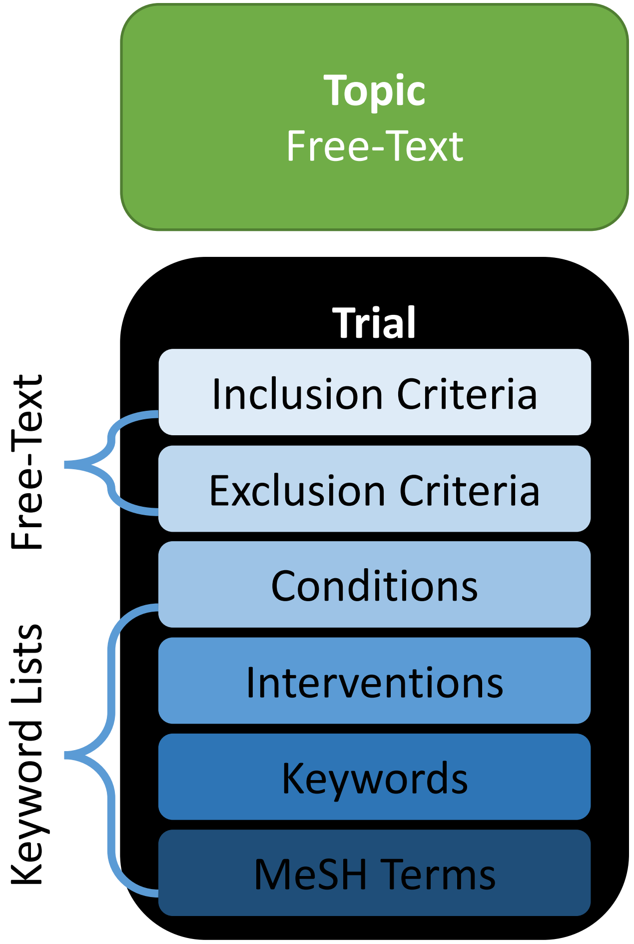
Encoding

Topic
Free-Text

Trial
- Inclusion Criteria
- Exclusion Criteria
- Conditions
- Interventions
- Keywords
- MeSH Terms

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



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*.

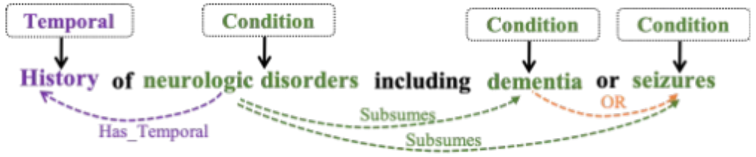
Encoding

Topic
Free-Text

Trial

Free-Text
- Inclusion Criteria
- Exclusion Criteria

Keyword Lists
- Conditions
- Interventions
- Keywords
- MeSH Terms



CHIA relation extraction

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

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*.

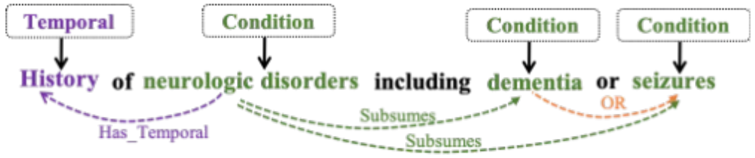
Encoding

Topic
Free-Text

Trial
- Inclusion Criteria
- Exclusion Criteria
- Conditions
- Interventions
- Keywords
- MeSH Terms

Free-Text

Keyword Lists



CHIA relation extraction
Extract entities with joint entity-relation 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

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

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*.

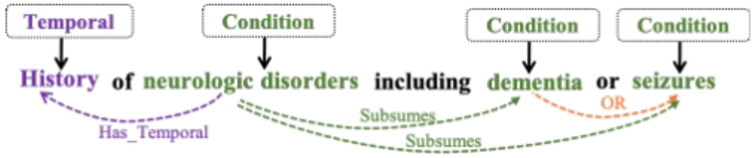
Encoding

Topic
Free-Text

Trial

Free-Text:
- Inclusion Criteria
- Exclusion Criteria

Keyword Lists:
- Conditions
- Interventions
- Keywords
- MeSH Terms



CHIA relation extraction
Extract entities with joint entity-relation 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

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

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*.

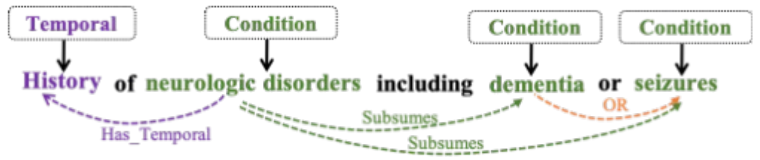
Encoding

Topic
Free-Text

Trial

Free-Text
- Inclusion Criteria
- Exclusion Criteria

Keyword Lists
- Conditions
- Interventions
- Keywords
- MeSH Terms



CHIA relation extraction
Extract entities with joint entity-relation 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

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

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*.

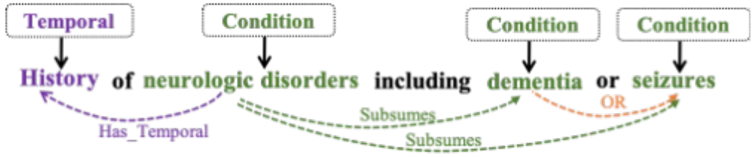
Encoding

Topic
Free-Text

Trial

Free-Text
- Inclusion Criteria
- Exclusion Criteria

Keyword Lists
- Conditions
- Interventions
- Keywords
- MeSH Terms

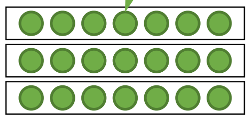


CHIA relation extraction
Extract entities with joint entity-relation extraction model (Wang and Lu, 2020) trained on CHIA annotations (Kury et al., 2020)

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

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

CNN
We use a CNN to handle spans with have multiple tokens



patient representation

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*.

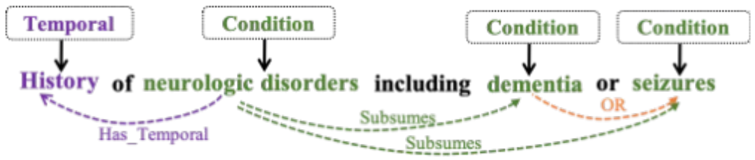
Encoding

Topic
Free-Text

Trial

Free-Text
- Inclusion Criteria
- Exclusion Criteria

Keyword Lists
- Conditions
- Interventions
- Keywords
- MeSH Terms

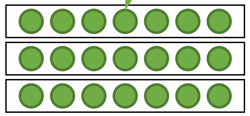


CHIA relation extraction
Extract entities with joint entity-relation extraction model (Wang and Lu, 2020) trained on CHIA annotations (Kury et al., 2020)

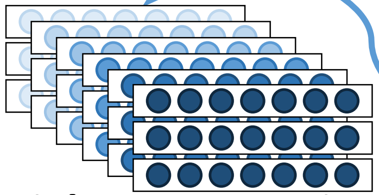
BERT-Based Encoder
Full text encoded using BERT model (Clinical BERT, BlueBERT)
Embeddings extracted for spans from CHIA

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

CNN
We use a CNN to handle spans with have multiple tokens



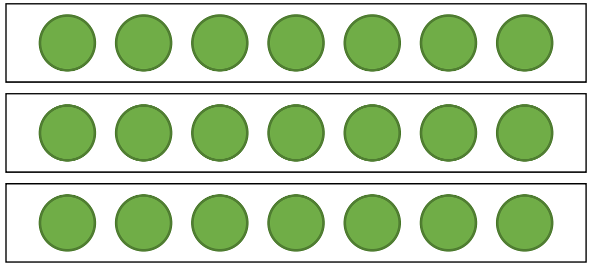
patient representation



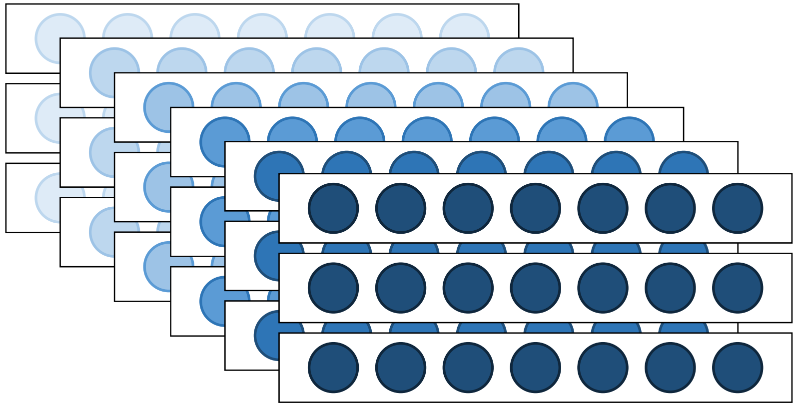
trial representations

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*.

Alignment



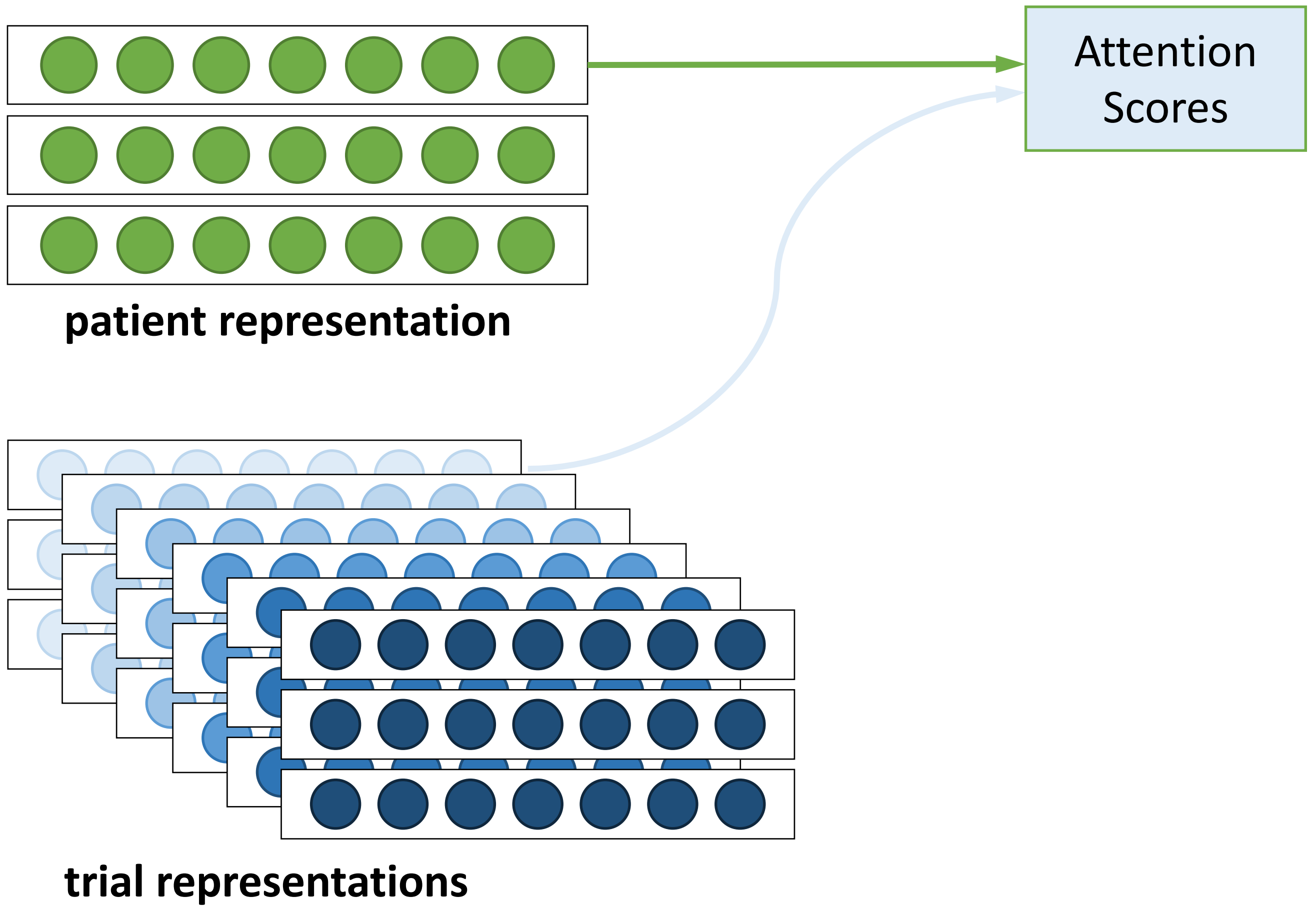
patient representation



trial representations

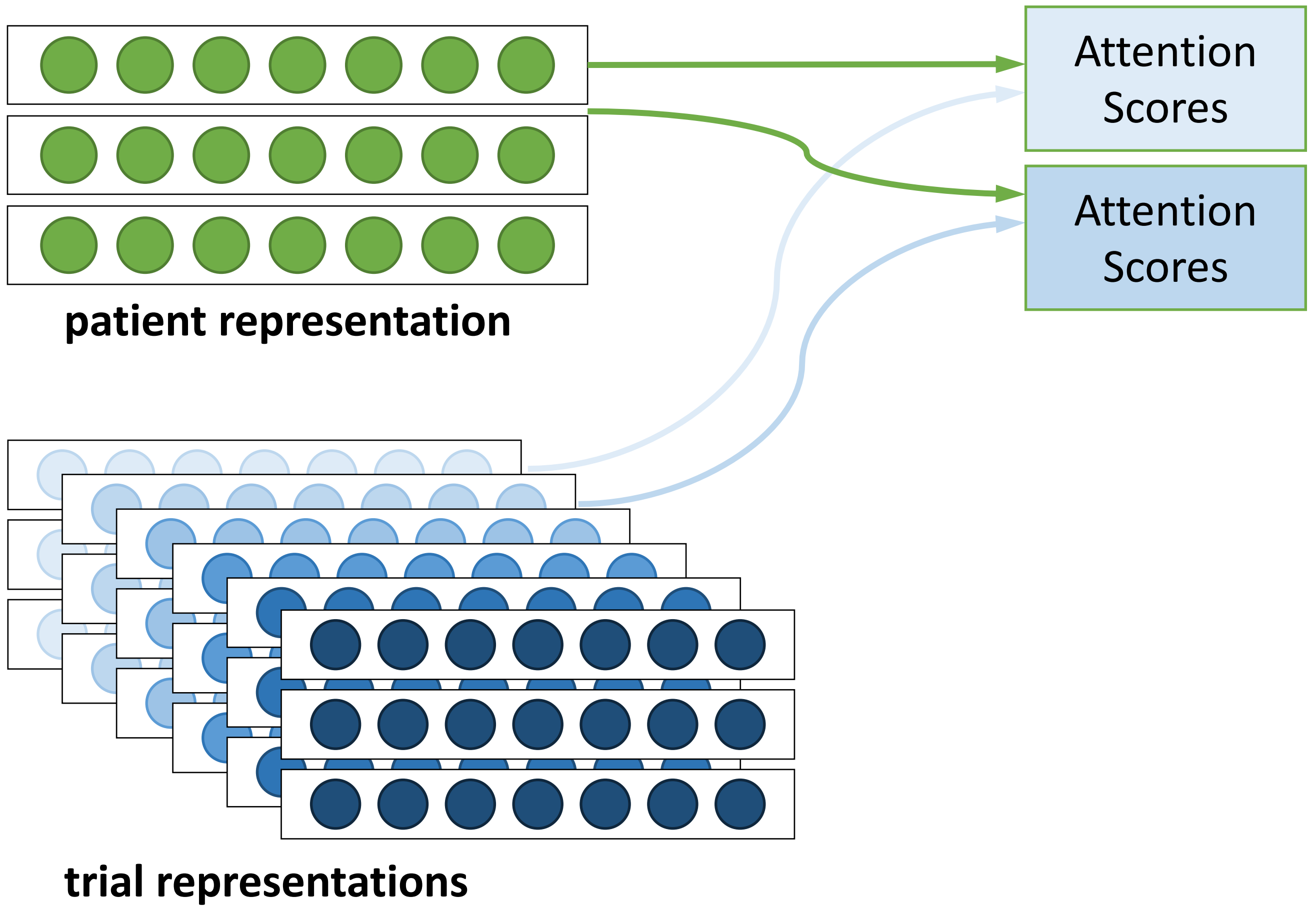
Weston, J., Bengio, S., & Usunier, N. (2011, June). Wsabie: Scaling up to large vocabulary image annotation. In *Twenty-Second International Joint Conference on Artificial Intelligence*.

Alignment



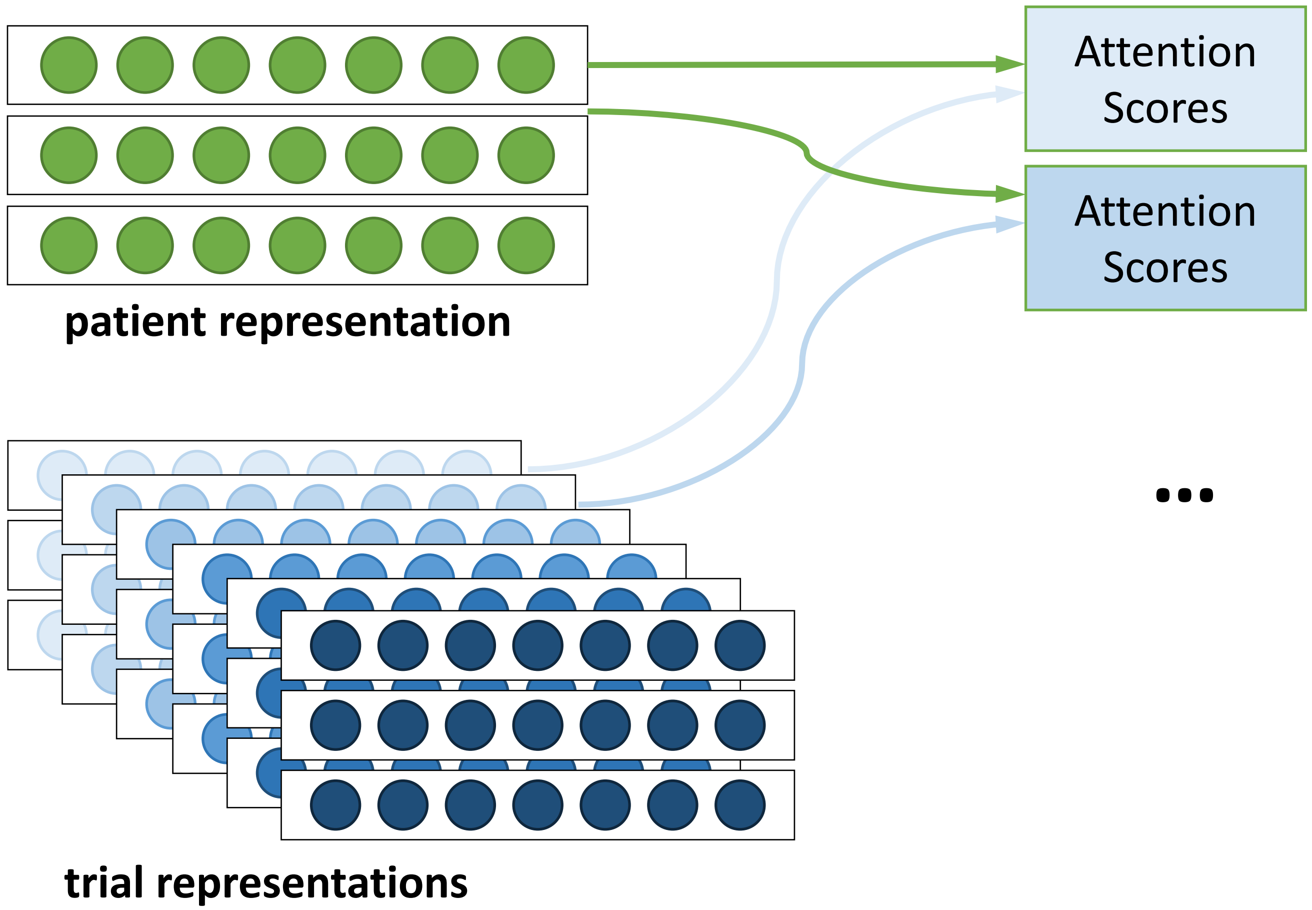
Weston, J., Bengio, S., & Usunier, N. (2011, June). Wsabie: Scaling up to large vocabulary image annotation. In *Twenty-Second International Joint Conference on Artificial Intelligence*.

Alignment



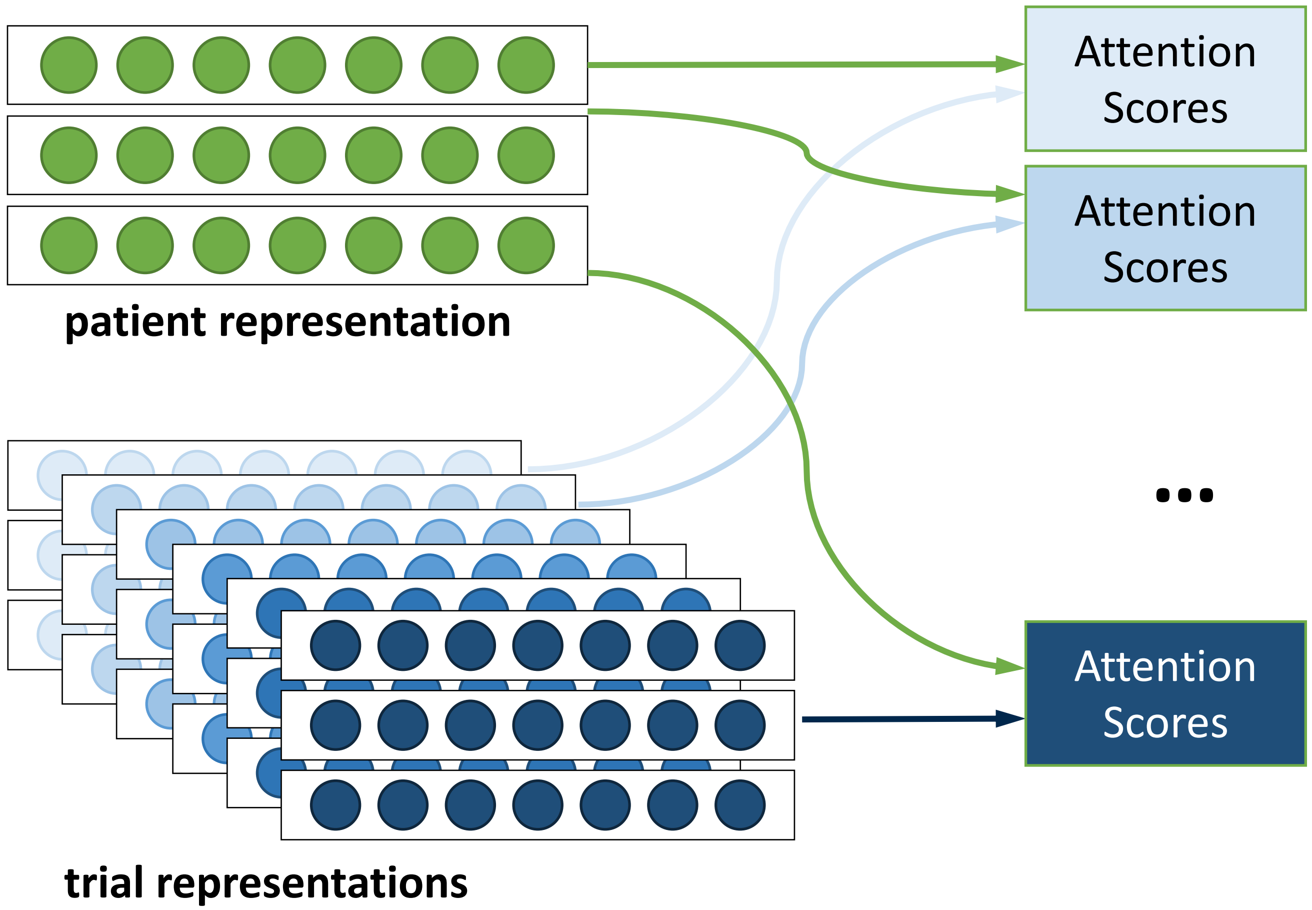
Weston, J., Bengio, S., & Usunier, N. (2011, June). Wsabie: Scaling up to large vocabulary image annotation. In *Twenty-Second International Joint Conference on Artificial Intelligence*.

Alignment



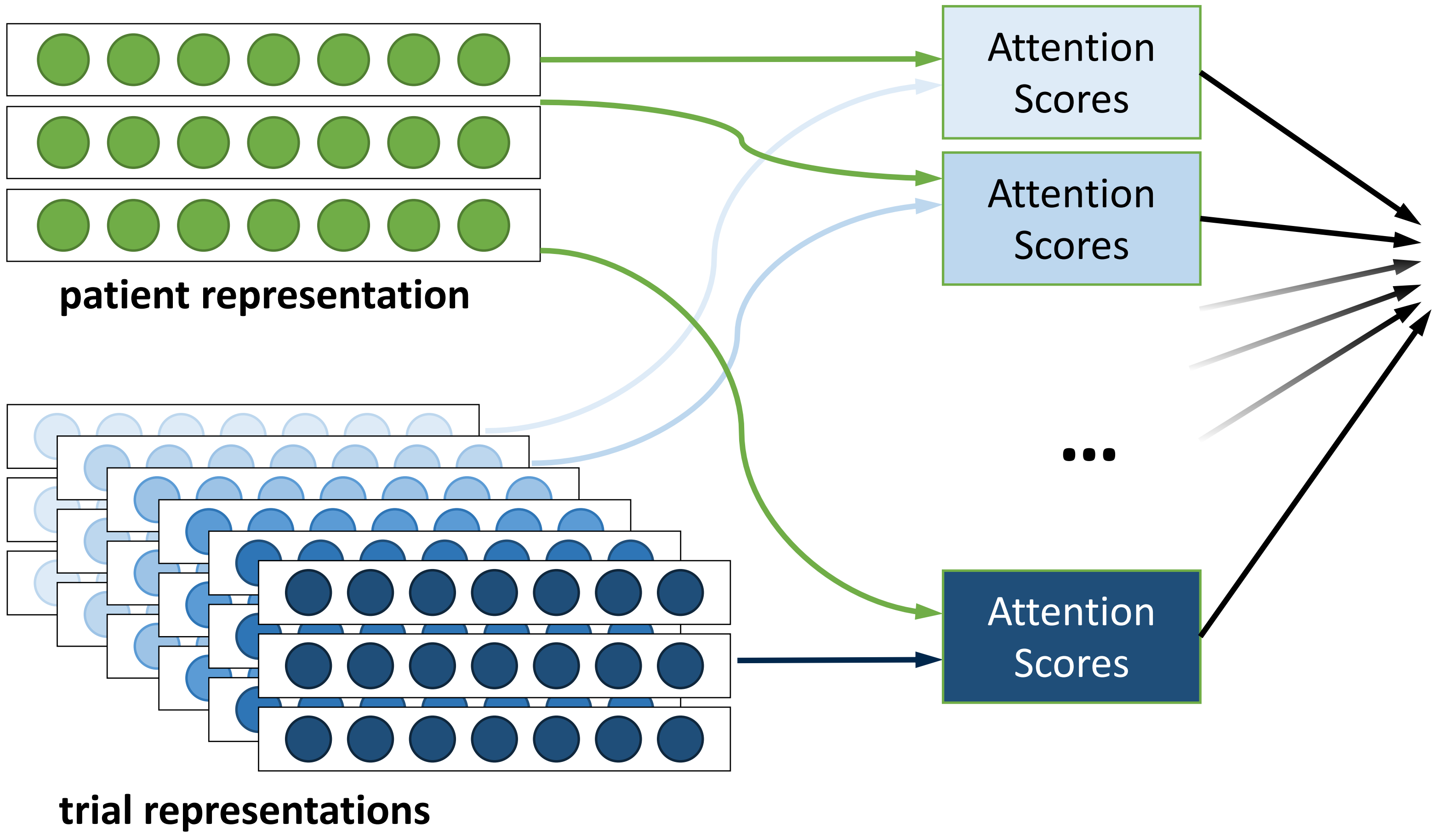
Weston, J., Bengio, S., & Usunier, N. (2011, June). Wsabie: Scaling up to large vocabulary image annotation. In *Twenty-Second International Joint Conference on Artificial Intelligence*.

Alignment



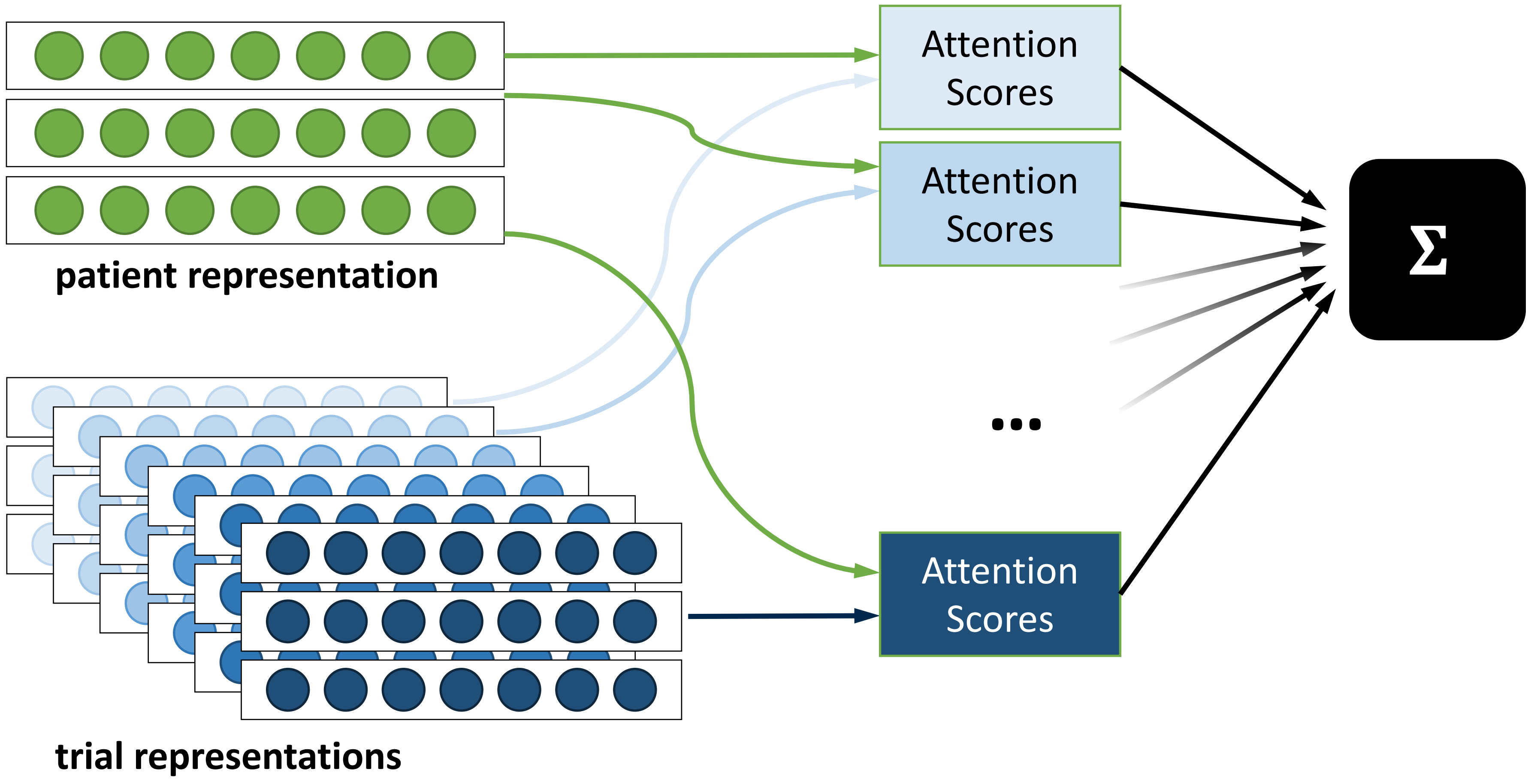
Weston, J., Bengio, S., & Usunier, N. (2011, June). Wsabie: Scaling up to large vocabulary image annotation. In *Twenty-Second International Joint Conference on Artificial Intelligence*.

Alignment



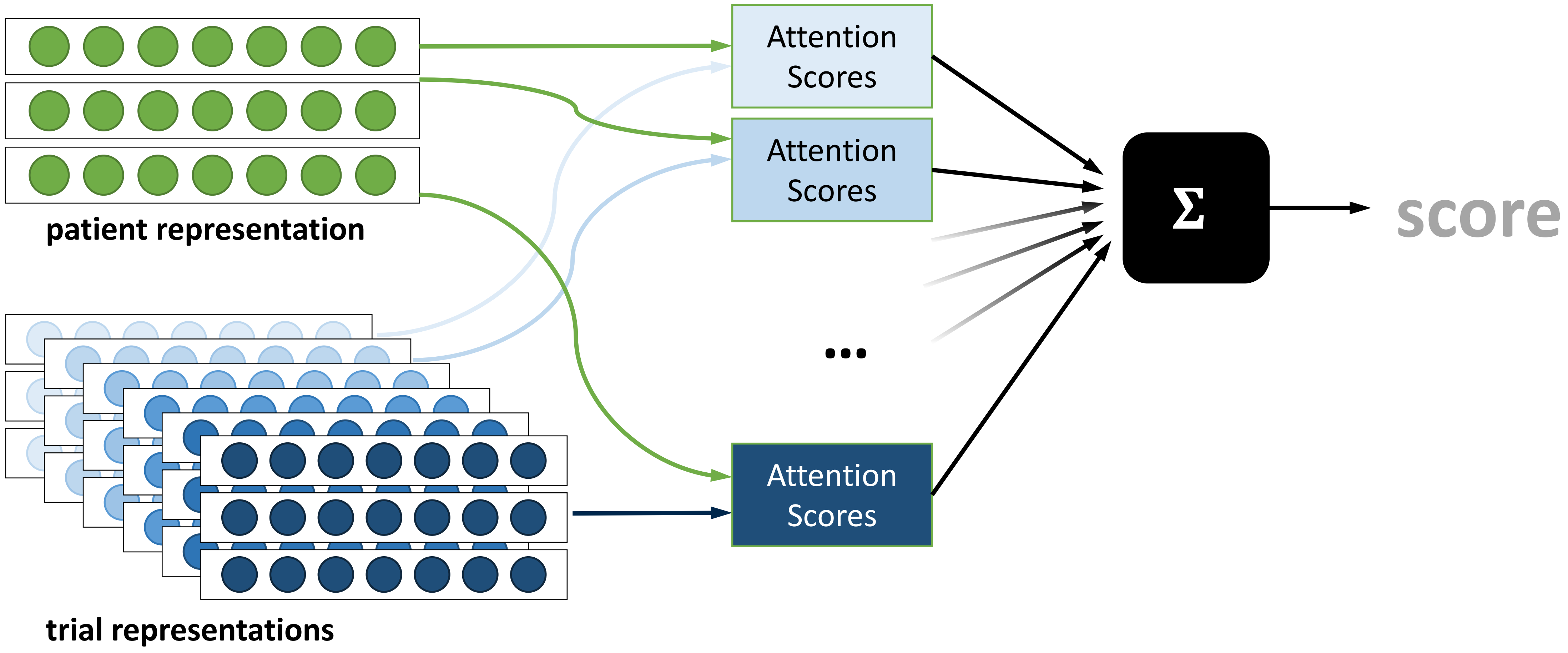
Weston, J., Bengio, S., & Usunier, N. (2011, June). Wsabie: Scaling up to large vocabulary image annotation. In *Twenty-Second International Joint Conference on Artificial Intelligence*.

Alignment



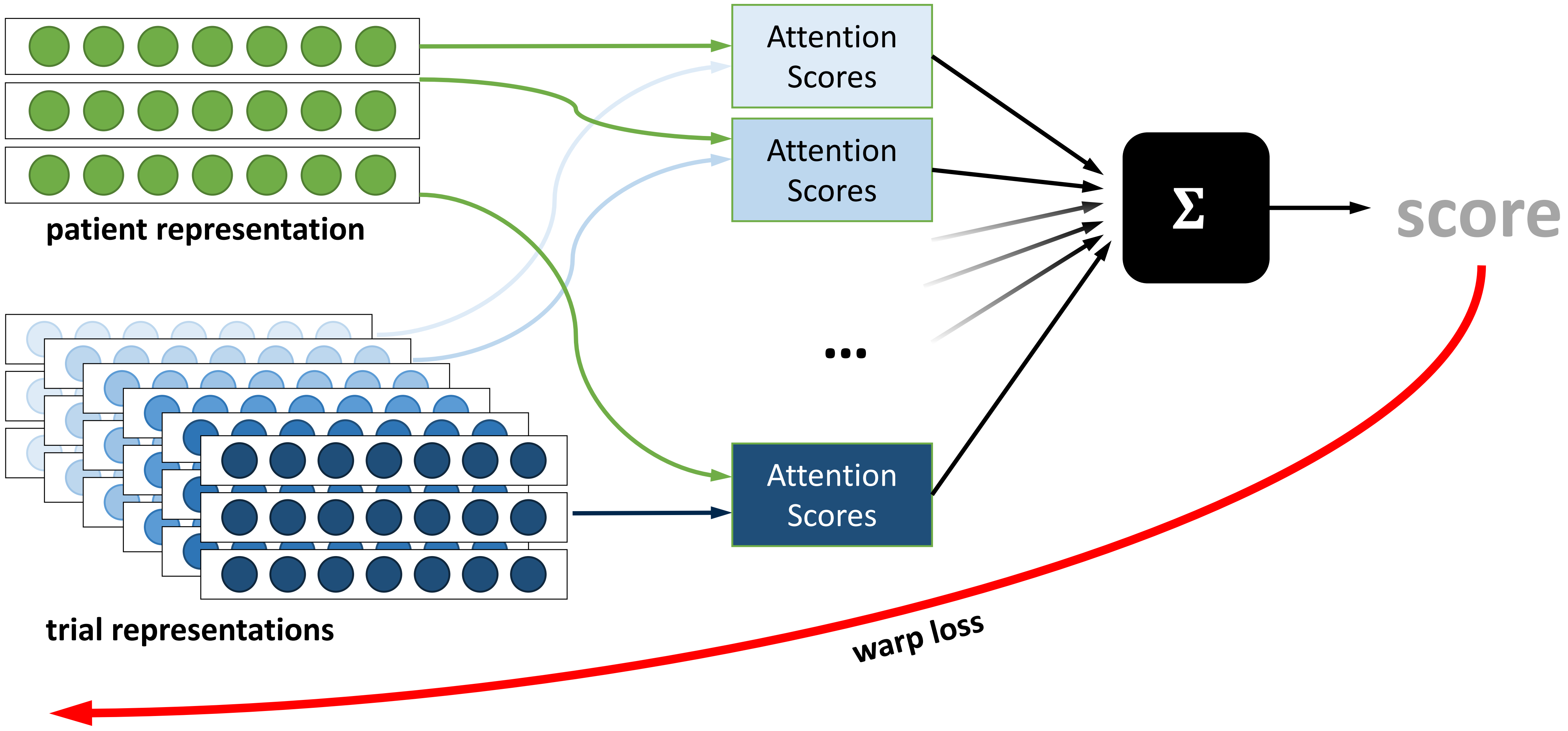
Weston, J., Bengio, S., & Usunier, N. (2011, June). Wsabie: Scaling up to large vocabulary image annotation. In *Twenty-Second International Joint Conference on Artificial Intelligence*.

Alignment

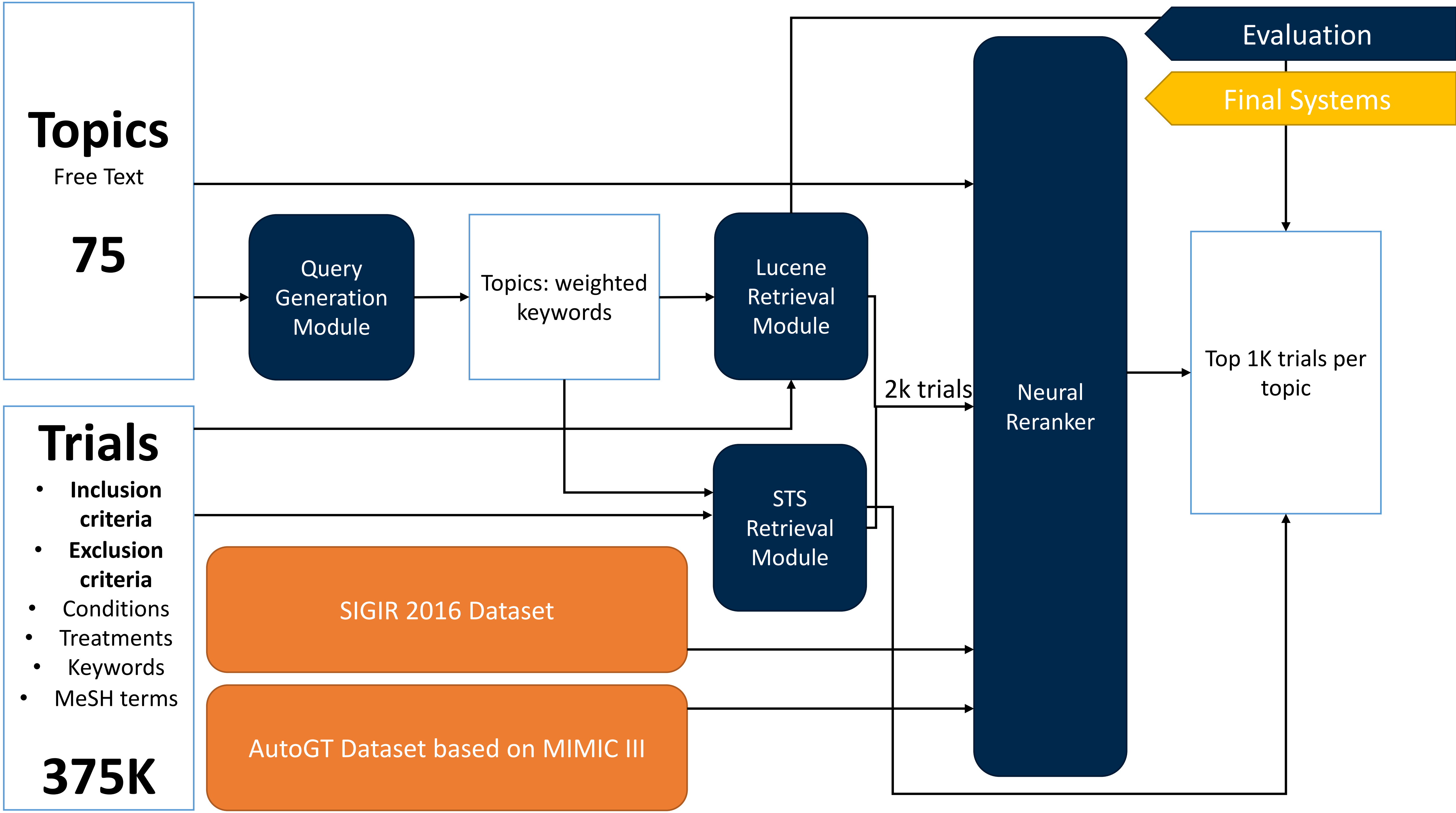


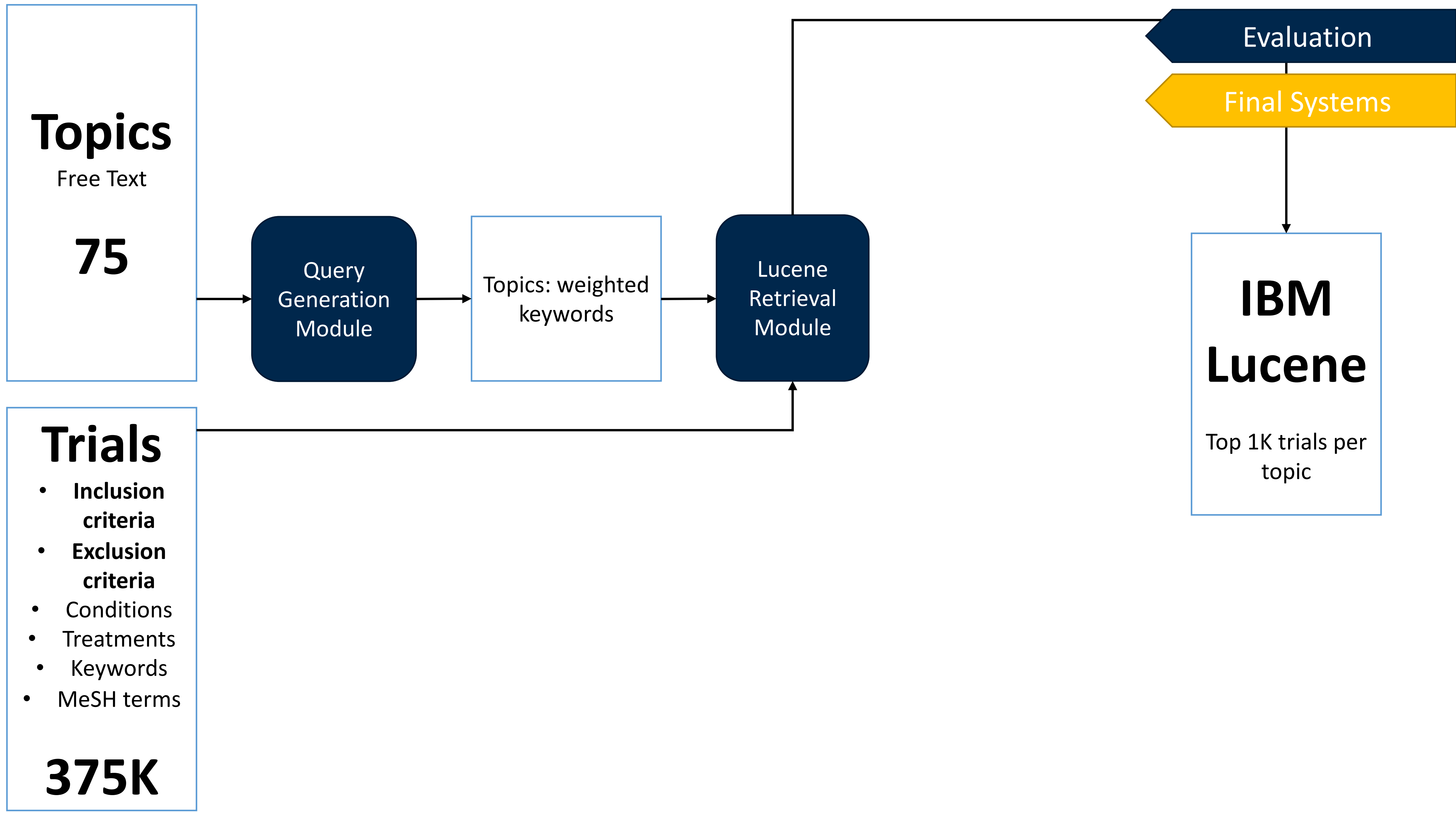
Weston, J., Bengio, S., & Usunier, N. (2011, June). Wsabie: Scaling up to large vocabulary image annotation. In *Twenty-Second International Joint Conference on Artificial Intelligence*.

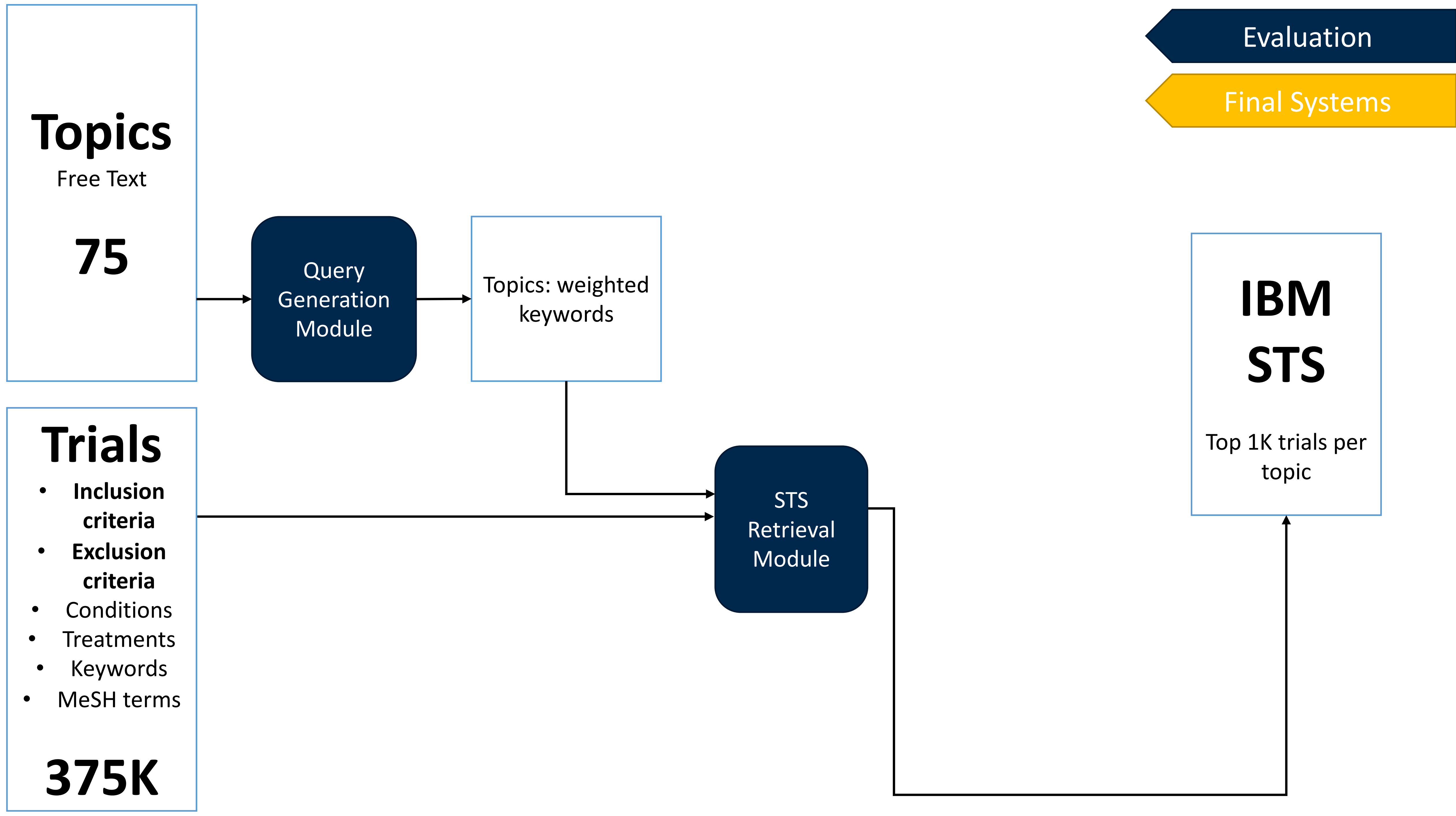
Alignment

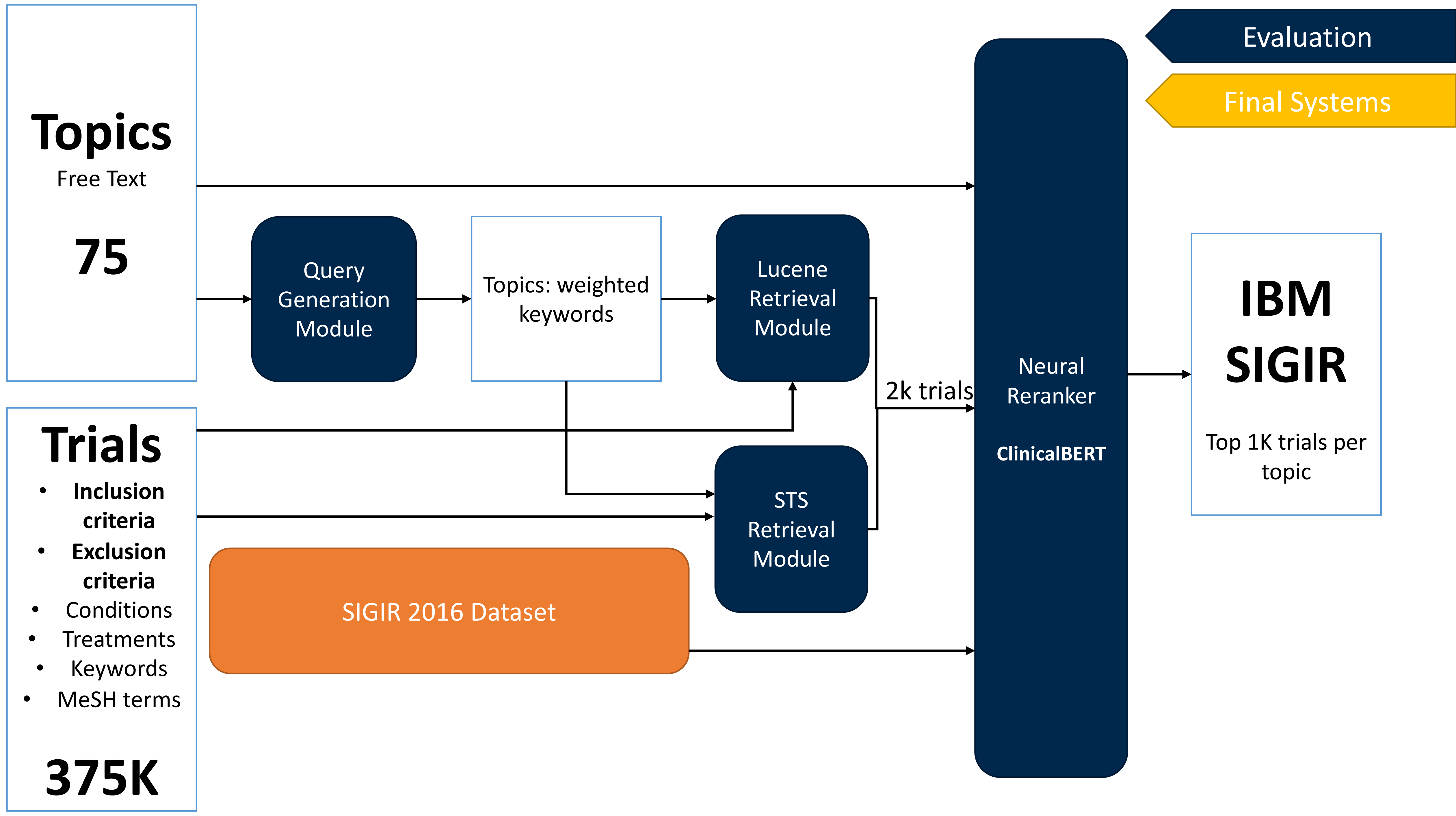


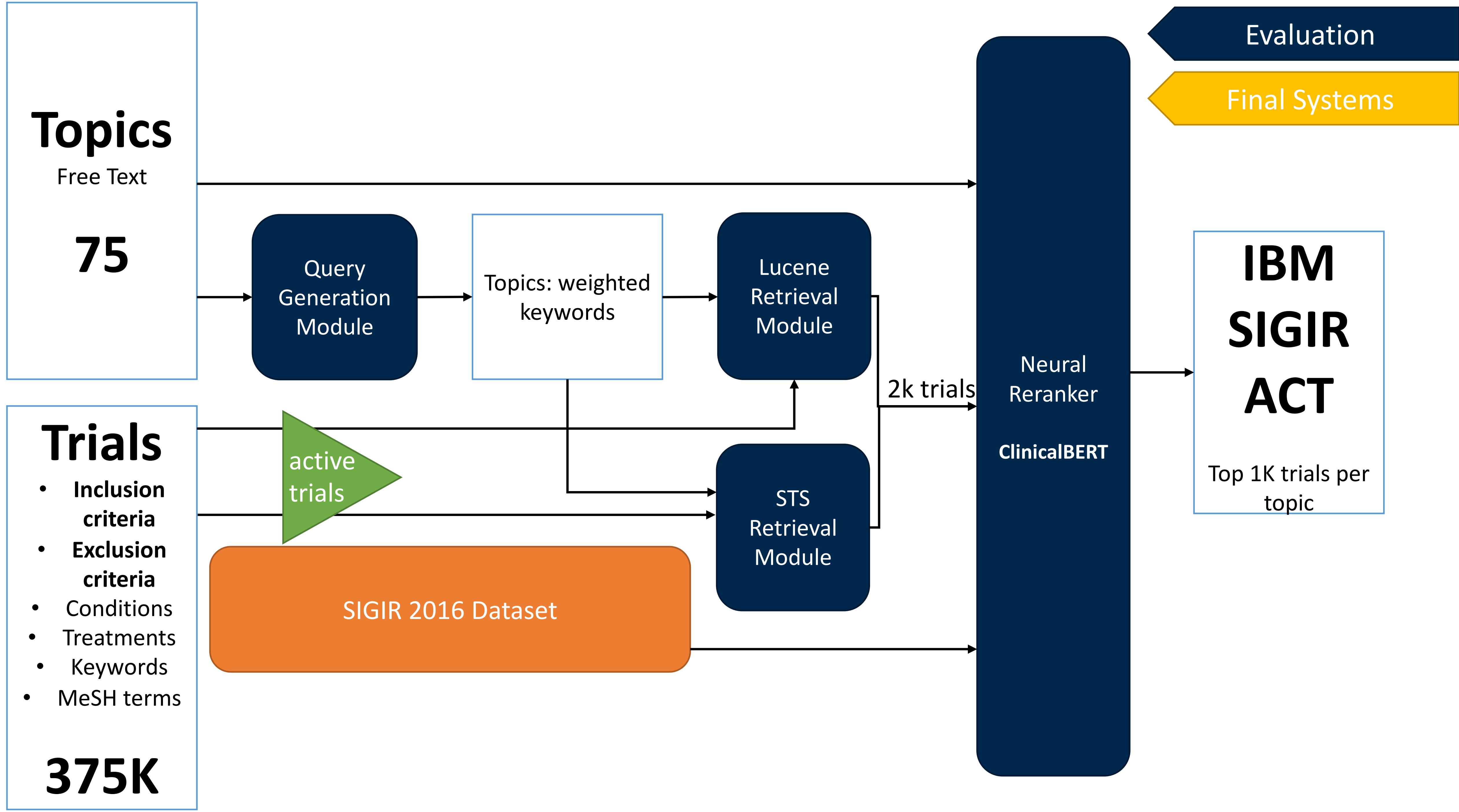
Weston, J., Bengio, S., & Usunier, N. (2011, June). Wsabie: Scaling up to large vocabulary image annotation. In *Twenty-Second International Joint Conference on Artificial Intelligence*.

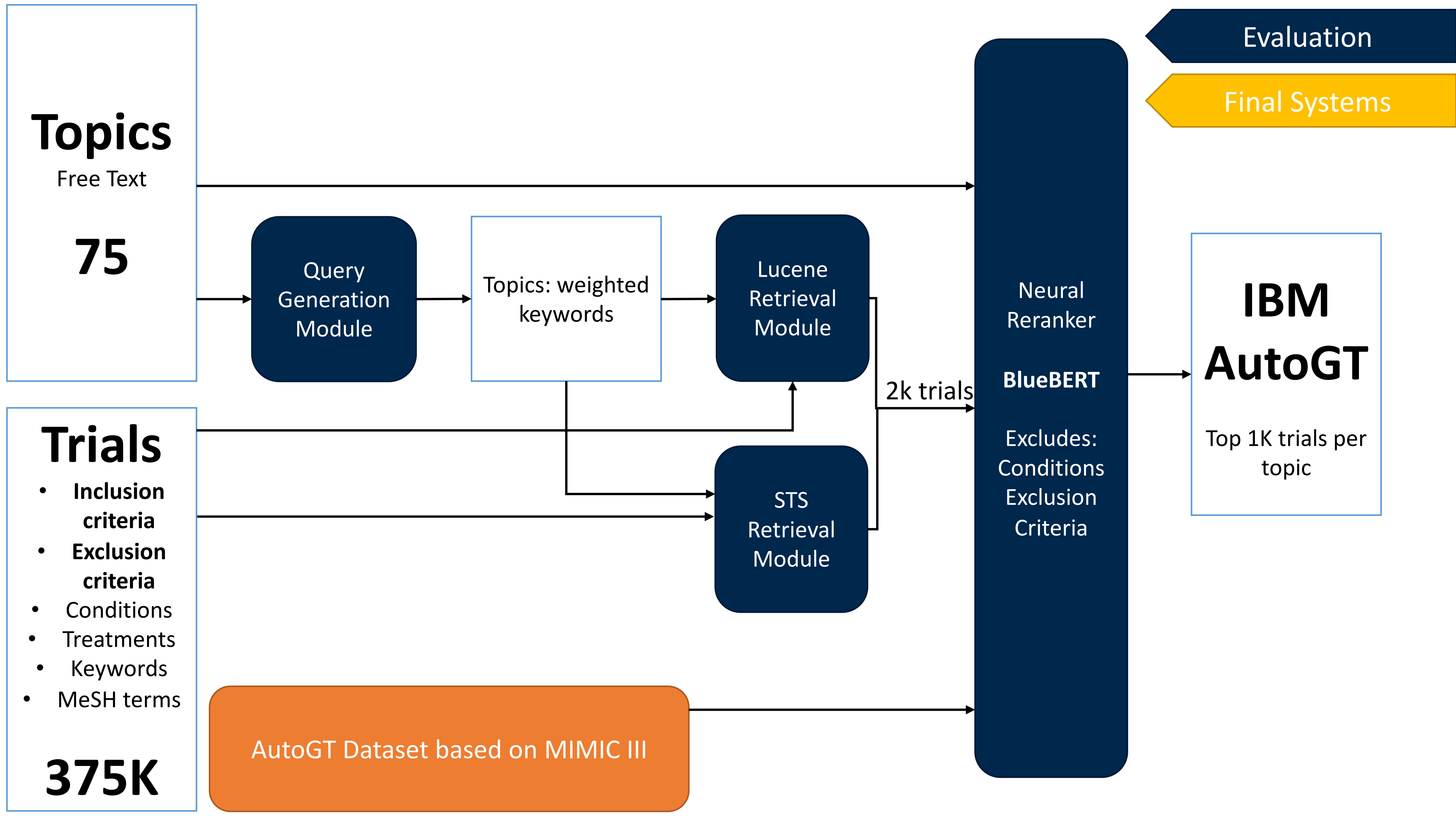

Topics
Free Text
75
Query Generation Module
Topics: weighted keywords
Lucene Retrieval Module
2k trials
Neural Reranker
Top 1K trials per topic
Evaluation
Final Systems
Trials
Inclusion criteria
Exclusion criteria
Conditions
Treatments
Keywords
MeSH terms
375K
STS Retrieval Module
SIGIR 2016 Dataset
AutoGT Dataset based on MIMIC III


Topics
Free Text
75
Query Generation Module
Topics: weighted keywords
Lucene Retrieval Module
Evaluation
Final Systems
IBM Lucene
Top 1K trials per topic
Trials
Inclusion criteria
Exclusion criteria
Conditions
Treatments
Keywords
MeSH terms
375K


Evaluation
Final Systems
Topics
Free Text
75
Query Generation Module
Topics: weighted keywords
Trials
Inclusion criteria
Exclusion criteria
Conditions
Treatments
Keywords
MeSH terms
375K
STS Retrieval Module
IBM STS
Top 1K trials per topic


Topics
Free Text
75
Query Generation Module
Topics: weighted keywords
Lucene Retrieval Module
2k trials
Neural Reranker
ClinicalBERT
Evaluation
Final Systems
IBM
SIGIR
Top 1K trials per topic
Trials
• Inclusion criteria
• Exclusion criteria
• Conditions
• Treatments
• Keywords
• MeSH terms
375K
STS Retrieval Module
SIGIR 2016 Dataset


Topics
Free Text
75
Query Generation Module
Topics: weighted keywords
Lucene Retrieval Module
2k trials
Neural Reranker
ClinicalBERT
Evaluation
Final Systems
IBM SIGIR ACT
Top 1K trials per topic
Trials
• Inclusion criteria
• Exclusion criteria
• Conditions
• Treatments
• Keywords
• MeSH terms
375K
active trials
STS Retrieval Module
SIGIR 2016 Dataset


Evaluation
Final Systems
Topics
Free Text
75
Query Generation Module
Topics: weighted keywords
Lucene Retrieval Module
STS Retrieval Module
2k trials
Neural Reranker
BlueBERT
Excludes: Conditions Exclusion Criteria
IBM AutoGT
Top 1K trials per topic
Trials
Inclusion criteria
Exclusion criteria
Conditions
Treatments
Keywords
MeSH terms
375K
AutoGT Dataset based on MIMIC III

Results

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

Results

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

Results

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

Results

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

Takeaways

Takeaways

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

Takeaways

- *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

Takeaways

- *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

Takeaways

- *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