Attentive Neural Architecture for Ad-hoc Structured Document Retrieval

Saeid Balaneshin [1] Alexander Kotov [1] Fedor Nikolaev [1,2]

[1]Textual Data Analytics Lab, Department of Computer Science, Wayne State University

[2]Kazan Federal University

Ad-hoc Structured (Multi-field) Document Retrieval

- IR research traditionally views documents as *holistic* and *homogeneous* units of text
- The task of retrieving structured (multi-field) documents arises in many information access scenarios:
  - Entity retrieval from knowledge graph(s)
  - Web document retrieval
  - Product search in e-Commerce

Entity Retrieval from Knowledge Graph(s)

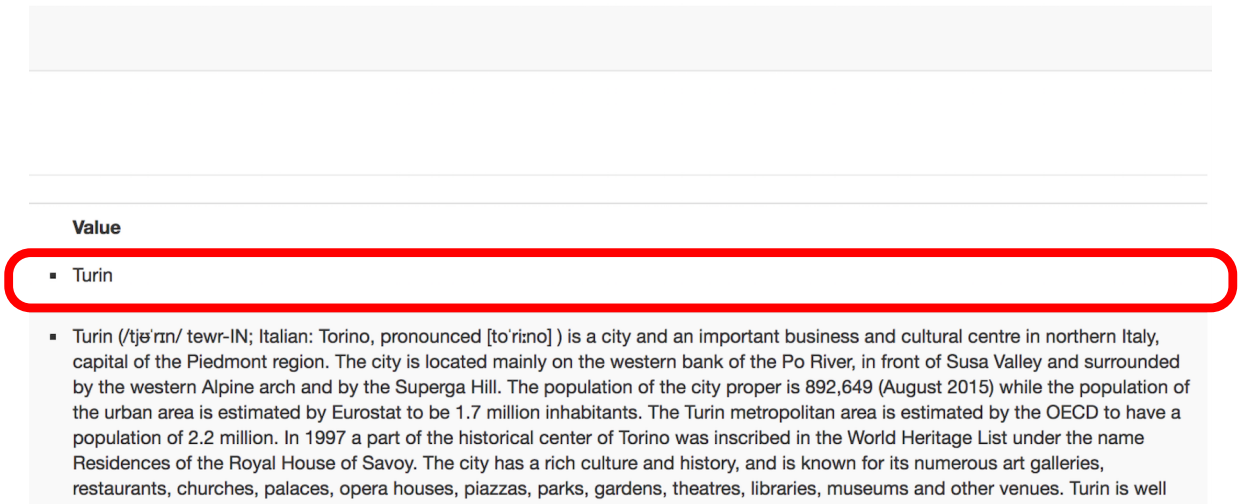
DBpedia

About: Turin

Property	Value
dbo:names	▪ Turin
dbo:attributes	▪ Turin (/tjʊˈrɪn/ tewr-IN; Italian: Torino, pronounced [toˈriːno]) is a city and an important business and cultural centre in northern Italy, capital of the Piedmont region. The city is located mainly on the western bank of the Po River, in front of Susa Valley and surrounded by the western Alpine arch and by the Superga Hill. The population of the city proper is 892,649 (August 2015) while the population of the urban area is estimated by Eurostat to be 1.7 million inhabitants. The Turin metropolitan area is estimated by the OECD to have a population of 2.2 million. In 1997 a part of the historical center of Torino was inscribed in the World Heritage List under the name Residences of the Royal House of Savoy. The city has a rich culture and history, and is known for its numerous art galleries, restaurants, churches, palaces, opera houses, piazzas, parks, gardens, theatres, libraries, museums and other venues. Turin is well known for its renaissance, baroque, rococo, neo-classical, and art nouveau architecture. Much of the city's public squares, castles, gardens and elegant palazzi such as Palazzo Madama, were built in the 16th to 18th century, after the capital of the Duchy of Savoy (later Kingdom of Sardinia) was moved to Turin from Chambery (now in France) as part of the urban expansion. The city used to be a major European political centre, being Italy's first capital city in 1861 and being home to the House of Savoy, Italy's royal family.It was the capital of the Duchy of Savoy from 1563, then of the Kingdom of Sardinia ruled by the Royal House of Savoy and finally the first capital of the unified Italy.Turin is sometimes called the cradle of Italian liberty, for having been the birthplace and home of notable politicians and people who contributed to the Risorgimento, such as Cavour. The city currently hosts some of Italy's best universities, colleges, academies, lycea and gymnasia, such as the six-century-old University of Turin and the Turin Polytechnic. Prestigious and important museums, such as the Museo Egizio and the Mole Antonelliana are also found in the city. Turin's several monuments and sights make it one of the world's top 250 tourist destinations, and the tenth most visited city in Italy in 2008. Even though much of its political significance and importance had been lost by World War II, it became a major European crossroad for industry, commerce and trade, and currently is one of Italy's main industrial centres, being part of the famous "industrial triangle", along with Milan and Genoa. Turin is ranked third in Italy, after Milan and Rome, for economic strength. With a GDP of $58 billion, Turin is the world's 78th richest city by purchasing power, and as of 2010 has been ranked by GaWC as a Gamma- world city. Turin is also home to much of the Italian automotive industry. Turin is well known as the home of the Shroud of Turin, the football teams Juventus F.C. and Torino F.C., the headquarters of automobile manufacturers FIAT, Lancia and Alfa Romeo, and as host of the 2006 Winter Olympics. (en)
dbo:categories	▪ Roman colonies Cities and towns in Piedmont Province of Turin,Capitals of former nations,Former national capitals,Turin Former capitals of Italy,Roman towns and cities in Italy
dbo:similarentitynames	▪ Castra Taurinorum Turin, Piedmont Türì TORINO capital of Piedmont region Augusta Pragelato 112 Turinese family Turin City Hall Turino Turin Film Festival (xsd:double)
dbo:relatedentitynames	▪ International Space Station Teatro Carignano Savoie List of political philosophers Haifa Parola, Carlo Residences of the Royal House of Savoy Eco, Umberto-1 Bagneux Anthropology The Frozen Autumn 1979 European Cup (athletics)

- ⟼ Names
- ⟼ Attributes
- ⟼ Categories
- ⟼ Similar Entity Names
- ⟼ Related Entity Names

Entity Retrieval from Knowledge Graph(s)



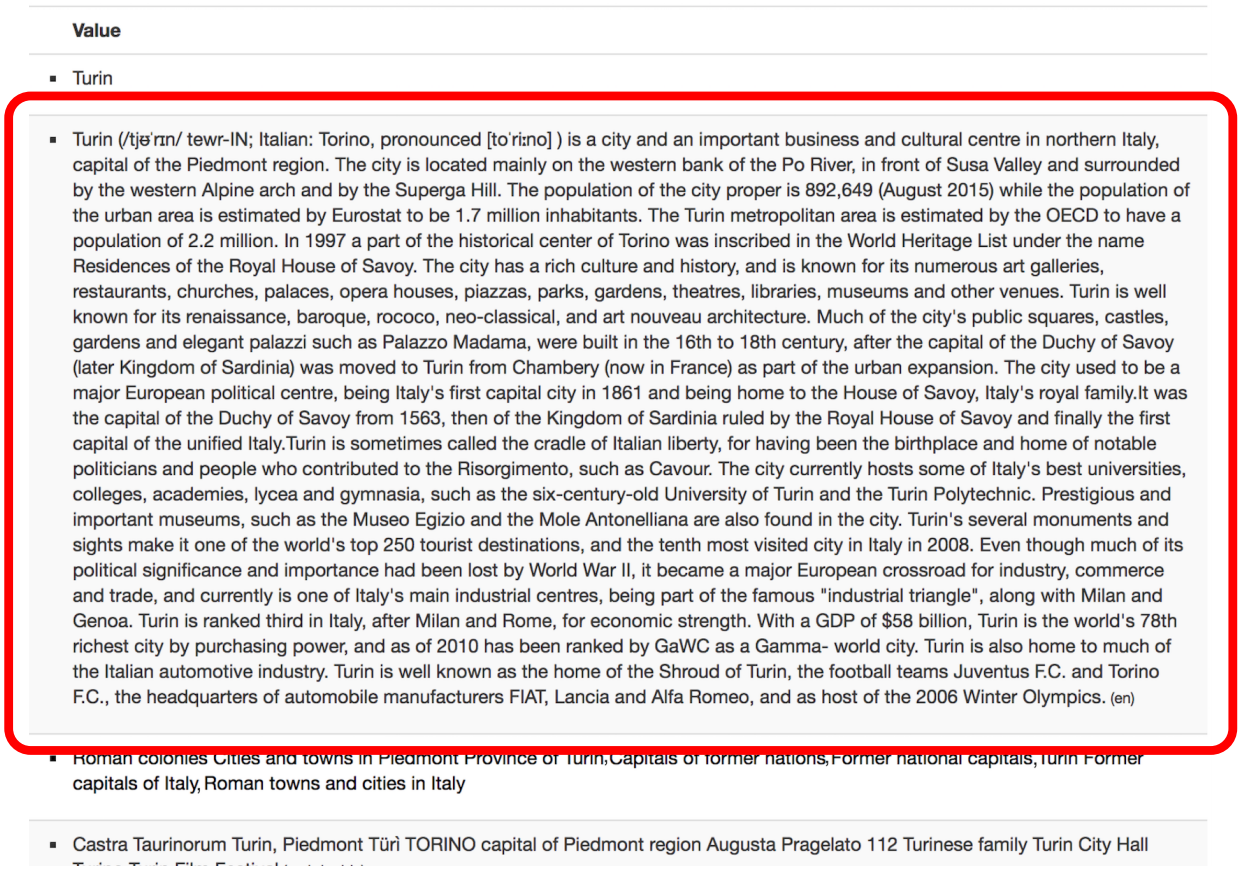
Value

- Turin

- Turin (/tjʊˈrɪn/ tewr-IN; Italian: Torino, pronounced [toˈriːno]) is a city and an important business and cultural centre in northern Italy, capital of the Piedmont region. The city is located mainly on the western bank of the Po River, in front of Susa Valley and surrounded by the western Alpine arch and by the Superga Hill. The population of the city proper is 892,649 (August 2015) while the population of the urban area is estimated by Eurostat to be 1.7 million inhabitants. The Turin metropolitan area is estimated by the OECD to have a population of 2.2 million. In 1997 a part of the historical center of Torino was inscribed in the World Heritage List under the name Residences of the Royal House of Savoy. The city has a rich culture and history, and is known for its numerous art galleries, restaurants, churches, palaces, opera houses, piazzas, parks, gardens, theatres, libraries, museums and other venues. Turin is well

- ⟼ Names
- ⟼ Attributes
- ⟼ Categories
- ⟼ Similar Entity Names
- ⟼ Related Entity Names

Entity Retrieval from Knowledge Graph(s)



Value

- Turin

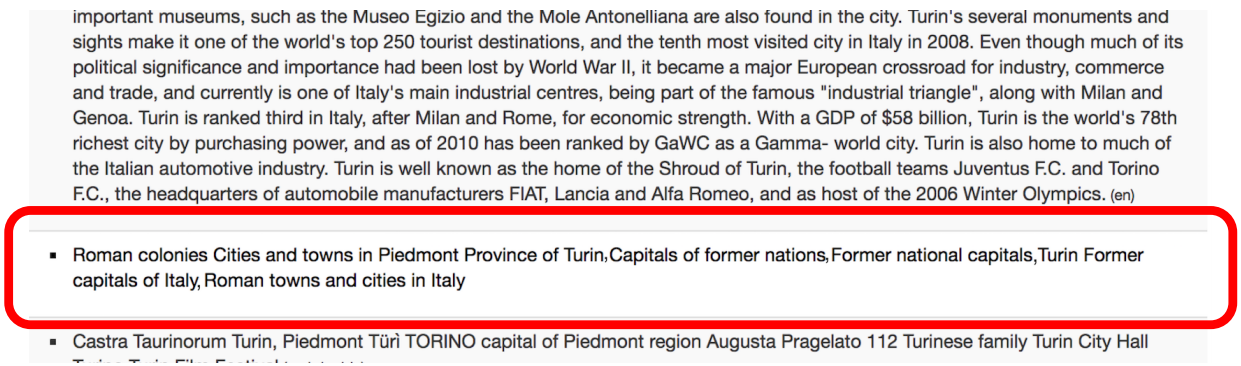
- Turin (/tjʊˈrɪn/ tewr-IN; Italian: Torino, pronounced [toˈriːno]) is a city and an important business and cultural centre in northern Italy, capital of the Piedmont region. The city is located mainly on the western bank of the Po River, in front of Susa Valley and surrounded by the western Alpine arch and by the Superga Hill. The population of the city proper is 892,649 (August 2015) while the population of the urban area is estimated by Eurostat to be 1.7 million inhabitants. The Turin metropolitan area is estimated by the OECD to have a population of 2.2 million. In 1997 a part of the historical center of Torino was inscribed in the World Heritage List under the name Residences of the Royal House of Savoy. The city has a rich culture and history, and is known for its numerous art galleries, restaurants, churches, palaces, opera houses, piazzas, parks, gardens, theatres, libraries, museums and other venues. Turin is well known for its renaissance, baroque, rococo, neo-classical, and art nouveau architecture. Much of the city's public squares, castles, gardens and elegant palazzi such as Palazzo Madama, were built in the 16th to 18th century, after the capital of the Duchy of Savoy (later Kingdom of Sardinia) was moved to Turin from Chambery (now in France) as part of the urban expansion. The city used to be a major European political centre, being Italy's first capital city in 1861 and being home to the House of Savoy, Italy's royal family.It was the capital of the Duchy of Savoy from 1563, then of the Kingdom of Sardinia ruled by the Royal House of Savoy and finally the first capital of the unified Italy.Turin is sometimes called the cradle of Italian liberty, for having been the birthplace and home of notable politicians and people who contributed to the Risorgimento, such as Cavour. The city currently hosts some of Italy's best universities, colleges, academies, lycea and gymnasia, such as the six-century-old University of Turin and the Turin Polytechnic. Prestigious and important museums, such as the Museo Egizio and the Mole Antonelliana are also found in the city. Turin's several monuments and sights make it one of the world's top 250 tourist destinations, and the tenth most visited city in Italy in 2008. Even though much of its political significance and importance had been lost by World War II, it became a major European crossroad for industry, commerce and trade, and currently is one of Italy's main industrial centres, being part of the famous "industrial triangle", along with Milan and Genoa. Turin is ranked third in Italy, after Milan and Rome, for economic strength. With a GDP of $58 billion, Turin is the world's 78th richest city by purchasing power, and as of 2010 has been ranked by GaWC as a Gamma- world city. Turin is also home to much of the Italian automotive industry. Turin is well known as the home of the Shroud of Turin, the football teams Juventus F.C. and Torino F.C., the headquarters of automobile manufacturers FIAT, Lancia and Alfa Romeo, and as host of the 2006 Winter Olympics. (en)

- Roman colonies Cities and towns in Piedmont Province of Turin,Capitals of former nations,Former national capitals,Turin Former capitals of Italy,Roman towns and cities in Italy

- Castra Taurinorum Turin, Piedmont Türì TORINO capital of Piedmont region Augusta Pragelato 112 Turinese family Turin City Hall

- ⟼ Names
- ⟼ Attributes
- ⟼ Categories
- ⟼ Similar Entity Names
- ⟼ Related Entity Names

Entity Retrieval from Knowledge Graph(s)



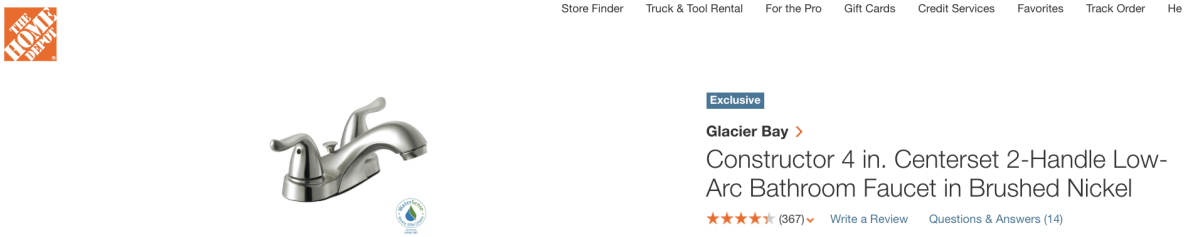
important museums, such as the Museo Egizio and the Mole Antonelliana are also found in the city. Turin's several monuments and sights make it one of the world's top 250 tourist destinations, and the tenth most visited city in Italy in 2008. Even though much of its political significance and importance had been lost by World War II, it became a major European crossroad for industry, commerce and trade, and currently is one of Italy's main industrial centres, being part of the famous "industrial triangle", along with Milan and Genoa. Turin is ranked third in Italy, after Milan and Rome, for economic strength. With a GDP of $58 billion, Turin is the world's 78th richest city by purchasing power, and as of 2010 has been ranked by GaWC as a Gamma- world city. Turin is also home to much of the Italian automotive industry. Turin is well known as the home of the Shroud of Turin, the football teams Juventus F.C. and Torino F.C., the headquarters of automobile manufacturers FIAT, Lancia and Alfa Romeo, and as host of the 2006 Winter Olympics. (en)

- Roman colonies Cities and towns in Piedmont Province of Turin, Capitals of former nations, Former national capitals, Turin Former capitals of Italy, Roman towns and cities in Italy

- Castra Taurinorum Turin, Piedmont Türì TORINO capital of Piedmont region Augusta Pragelato 112 Turinese family Turin City Hall

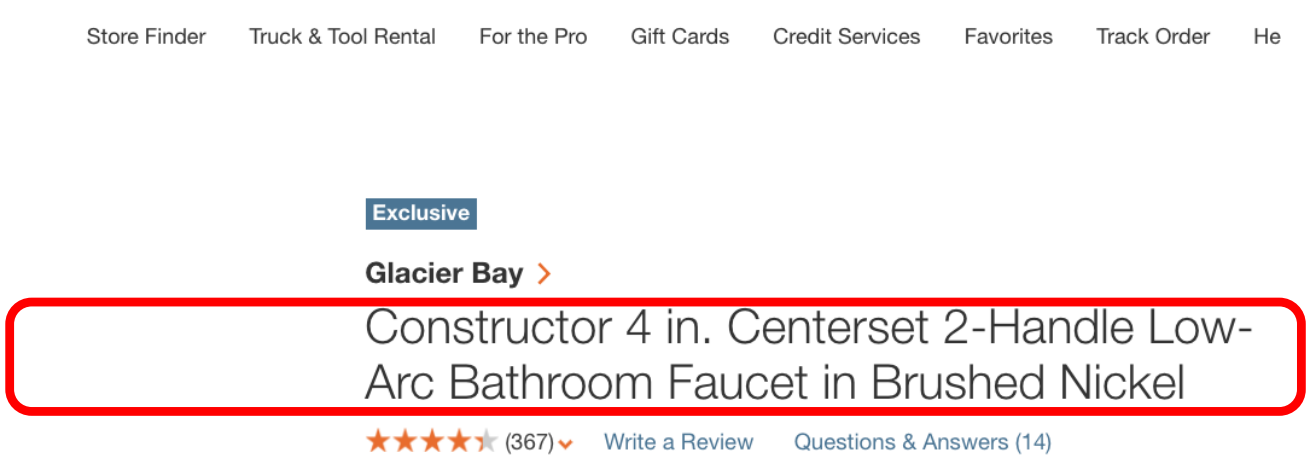
- ⟼ Names
- ⟼ Attributes
- ⟼ Categories
- ⟼ Similar Entity Names
- ⟼ Related Entity Names

Product Search



- ⟶ Title
- ⟶ Description
- ⟶ Attributes

Product Search



- ⟼ Title
- ⟼ Description
- ⟼ Attributes

Product Search

Description

The Glacier Bay Constructor 4 in. 2-Handle Low-Arc Bathroom Faucet in Brushed Nickel has an elegant, curved design that will complement a wide variety of bathroom or powder room decor. WaterSense certified, this faucet has a 1.2 GPM rate to help reduce water use and washerless cartridges to prevent drips. Easy to use with easy control metal lever handles, the sleek low-arc spout allows plenty of sink space. This faucet features the Glacier Bay Exclusive ClickInstall drain assembly for quick and easy installation. Pair with other pieces from the Constructor Collection for a complete and polished look.

- Brushed nickel finish uses a physical vapor deposition process that ensures the finish is a durable lifetime finish
- Washerless cartridges help prevent dripping
- Metal construction for durability and reliability
- Easy control metal lever handles are ADA complaint
- WaterSense certified to reduce water use without sacrificing performance
- Maximum flow rate: 1.2 GPM at 60 PSI water saving
- Faucet features the Glacier Bay exclusive ClickInstall drain for quick and simple installation
- 3-installation holes required
- Limited lifetime warranty

⟶ Title
⟶ Description
⟶ Attributes

Product Search

- Easy control metal lever handles are ADA complaint
- WaterSense certified to reduce water use without sacrificing performance
- Maximum flow rate: 1.2 GPM at 60 PSI water saving
- Faucet features the Glacier Bay exclusive ClickInstall drain for quick and simple installation
- 3-installation holes required
- Limited lifetime warranty

Attributes

Dimensions

Connection size (in.)	1/2 In.	Faucet Height (in.)

Details

Color Family	Nickel	Installation Type
Color/Finish	Brushed Nickel	Maximum Deck Thickness (in.)
Drain Included	Drain Included with Purchase	Minimum Sink Holes Required

⟼ Title
⟼ Description
⟼ Attributes

the WHITE HOUSE
PRESIDENT BARACK OBAMA

The Cabinet

Established in Article II, Section 2 of the Constitution, the Cabinet's role is to advise the President on any subject he may require relating to the duties of each member's respective office.

The tradition of the Cabinet dates back to the beginnings of the Presidency itself. Established in Article II, Section 2, of the Constitution, the Cabinet's role is to advise the President on any subject he may require relating to the duties of each member's respective office.

The Cabinet includes the Vice President and the heads of 15 executive departments — the Secretaries of Agriculture, Commerce, Defense, Education, Energy, Health and Human Services, Homeland Security, Housing and Urban Development, Interior, Labor, State, Transportation, Treasury, and Veterans Affairs, as well as the Attorney General.

In order of succession to the Presidency:

Vice President of the United States
Joseph R. Biden

Department of State
Secretary John Kerry
state.gov

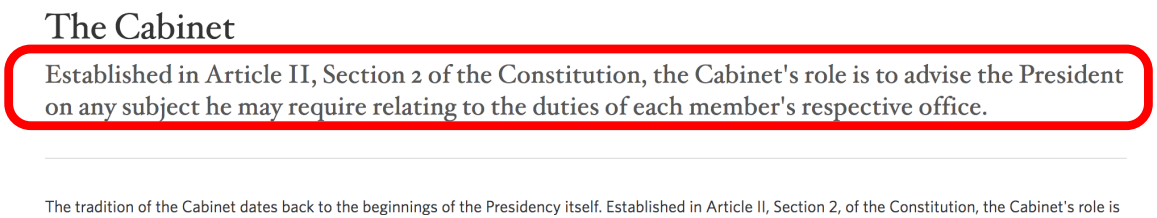
Department of the Treasury
Secretary Jack Lew
treasury.gov

Department of Defense
Secretary Ashton Carter
defense.gov

Department of Justice
Attorney General Loretta E. Lynch

- ⟼ Title
- ⟼ Texts in Large Font
- ⟼ Contents
- ⟼ Incoming Hyper-links
- ⟼ Document Meta-data
- ⟼ Alternative Texts for Images

- ⟼ Title
- ⟼ Texts in Large Font
- ⟼ Contents
- ⟼ Incoming Hyper-links
- ⟼ Document Meta-data
- ⟼ Alternative Texts for Images



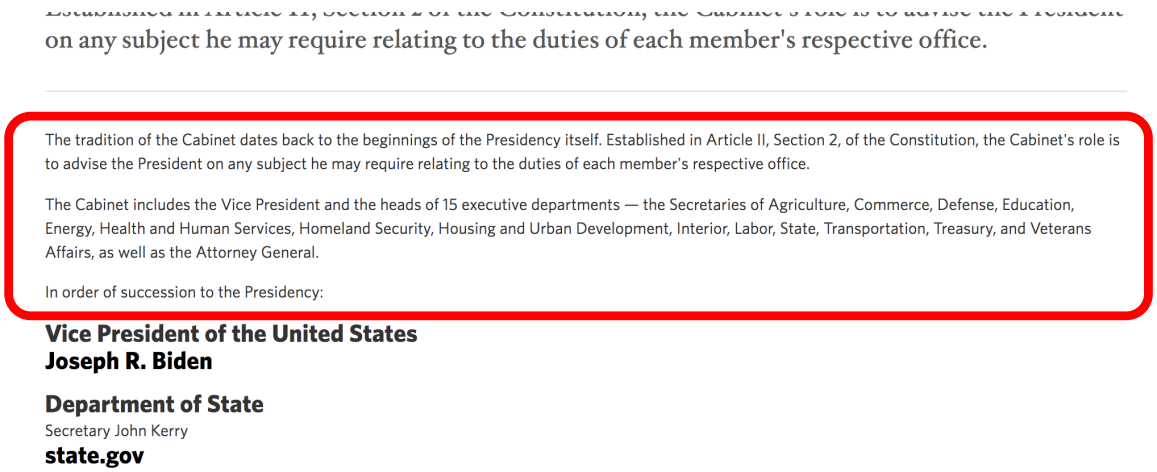
The Cabinet

Established in Article II, Section 2 of the Constitution, the Cabinet's role is to advise the President on any subject he may require relating to the duties of each member's respective office.

The tradition of the Cabinet dates back to the beginnings of the Presidency itself. Established in Article II, Section 2, of the Constitution, the Cabinet's role is

- ⟼ Title
- ⟼ Texts in Large Font
- ⟼ Contents
- ⟼ Incoming Hyper-links
- ⟼ Document Meta-data
- ⟼ Alternative Texts for Images

Web Search



on any subject he may require relating to the duties of each member's respective office.

The tradition of the Cabinet dates back to the beginnings of the Presidency itself. Established in Article II, Section 2, of the Constitution, the Cabinet's role is to advise the President on any subject he may require relating to the duties of each member's respective office.

The Cabinet includes the Vice President and the heads of 15 executive departments — the Secretaries of Agriculture, Commerce, Defense, Education, Energy, Health and Human Services, Homeland Security, Housing and Urban Development, Interior, Labor, State, Transportation, Treasury, and Veterans Affairs, as well as the Attorney General.

In order of succession to the Presidency:

Vice President of the United States
Joseph R. Biden

Department of State
Secretary John Kerry
state.gov

- ⟼ Title
- ⟼ Texts in Large Font
- ⟼ Contents
- ⟼ Incoming Hyper-links
- ⟼ Document Meta-data
- ⟼ Alternative Texts for Images

Document vs. Structured Document Retrieval

Document Retrieval

- relevance is quantified by aggregating heuristics calculated at the **document or collection level** (# of occurrences and proximity of query terms, IDF, document length)

Structured Document Retrieval

- requires strategies for aggregating heuristics calculated at the **level of document fields** into the matching score of an entire document

- effective for retrieving documents with **lexically similar, but semantically diverse fields**

Importance of Document Fields

Aggregation of field-level statistics of query terms in structured document retrieval is informed by a relative importance of document fields, which depends on:

- **properties or semantics of document fields**: e.g. a query term matched in a section of a Web page, which is in larger font, should have a different importance than a query term matched in other sections
- **query intent**: e.g. in the query *"attractive outdoor light with security features"* *"attractive"* refers to product description, *"outdoor light"* to product name and *"security features"* to product attributes

Mixture of Language Models (MLM)

[Ogilvie and Callan, SIGIR'03]

Document D with F fields is ranked w.r.t query Q according to:

$$P(Q|D) \overset{rank}{=} \prod_{q_i \in Q} P(q_i|\theta_D)^{n(q_i,Q)}$$

where

$$P(q_i|\theta_D) = \sum_{j=1}^{F} w^j P(q_i|\theta^j)$$

Fielded Sequential Dependence Model (FSDM)

[Zhiltsov et al., SIGIR'15]

- Extends SDM to the case of structured document retrieval (i.e. accounts for both unigram and sequential bigram concepts in a query and document structure)
- Document D with F fields is ranked w.r.t query Q according to:

$$P(D|Q) \overset{rank}{=} \lambda_T \sum_{q_i \in Q} \tilde{f}_T(q_i, D) + \lambda_O \sum_{q_i \in Q} \tilde{f}_O(q_i, q_{i+1}, D) + \lambda_U \sum_{q_i \in Q} \tilde{f}_U(q_i, q_{i+1}, D)$$

- Potential function for query unigram q_i:

$$\tilde{f}_T(q_i, D) = \log \sum_{j=1}^{F} w^j P(q_i|\theta^j)$$

Challenges of Structured Document Retrieval

Methods for structured document retrieval (SDR) face three major challenges:

- identifying the key concepts (words or phrases) in keyword queries
- semantic matching of the key query concepts in different fields of structured documents
- aggregating the scores of the matched query phrases into the overall score of a structured document

Key limitation: all previously proposed SDR methods are based on direct matching of concepts in queries and document fields $\rightarrow$ *lexical gap*

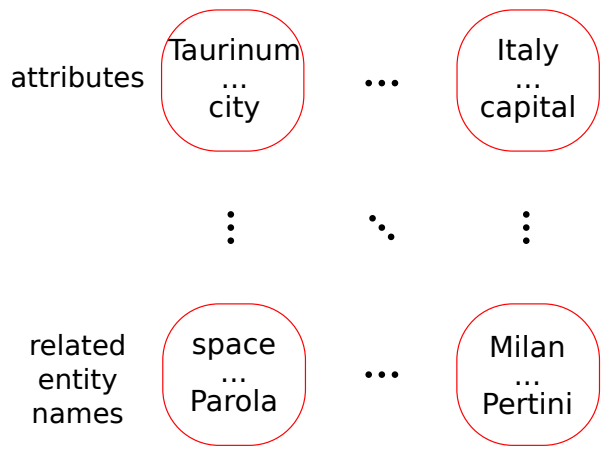
Attention-based **N**eural Architecture for Ad-hoc **S**tructured Document **R**etrieval (**ANSR**):

- **Input:** embeddings of words in a query and document fields
- **Pooling layers:** create compressed interaction matrices of the same dimensions between unigram- and bigram-based query and document field phrases
- **Matching score aggregation layers:** combine the matching scores of query phrases in different document fields into the overall document relevance score by taking into account relative importance of query phrases and document fields
- **Document field attention layers:** calculate relative importance of document fields
- **Query phrase attention layers:** calculate relative importance of query phrases

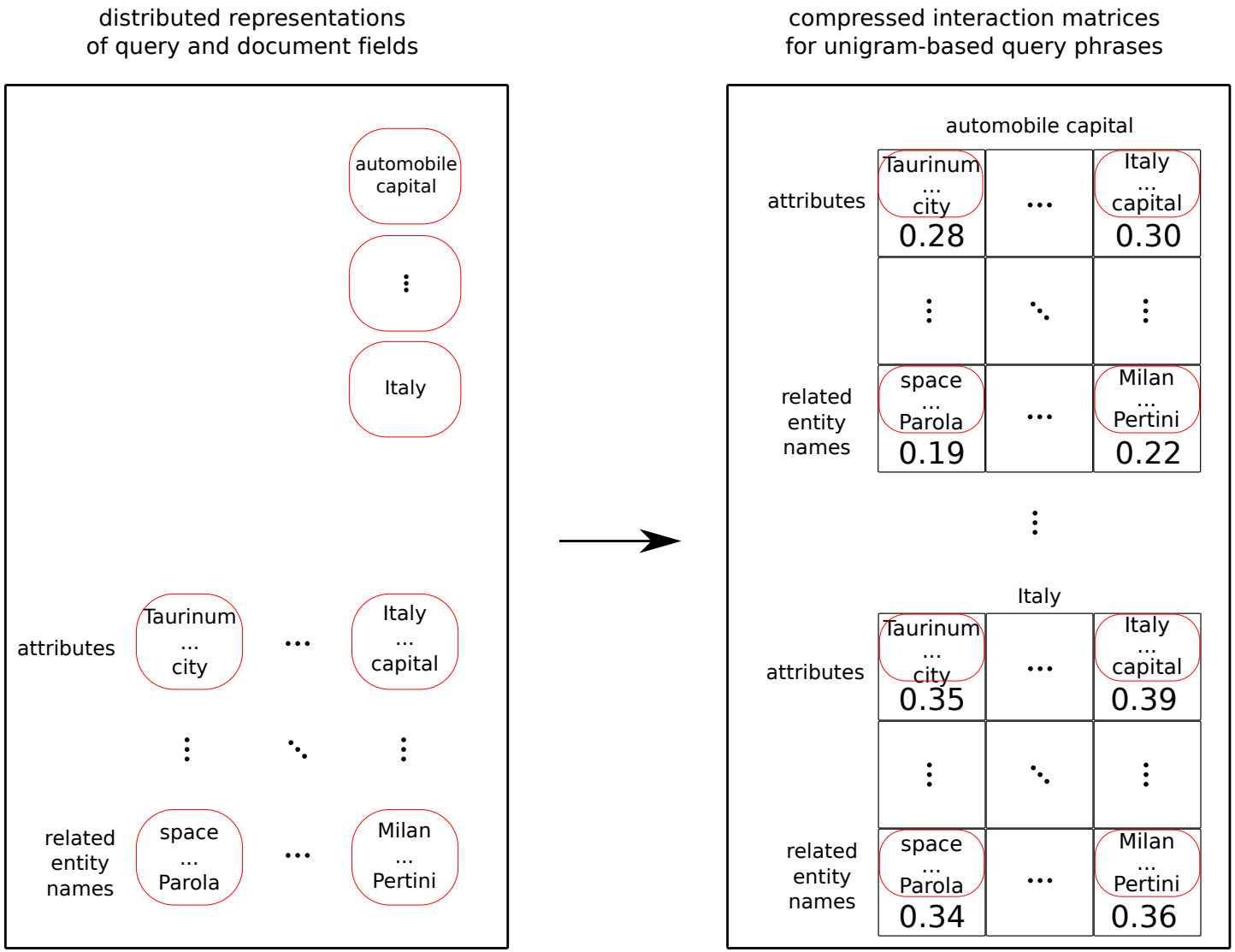
Step 1: create distributed representations of a query and each document field

Query: automobile capital and the Detroit of Italy
Document: `http://dbpedia.org/page/Turin`

attributes	Taurinum Turin is an important business and cultural center in northern Italy, capital city of the Piedmont region located mainly on the left bank of the Po River · · · · · · · · · · · · · · · Susa Valley Italy it is also dubbed la capitale Sabauda Savoyard capital
· · ·	· · ·
related entity names	Space Station Teatro Carignano Savoie List of political philosophers Haifa Parola, Carlo Residences of the Royal House of Savoy Eco · · · , · · · Duchy of Milan Mezzo-soprano Genoa Ginzburg Alessandro Pertini



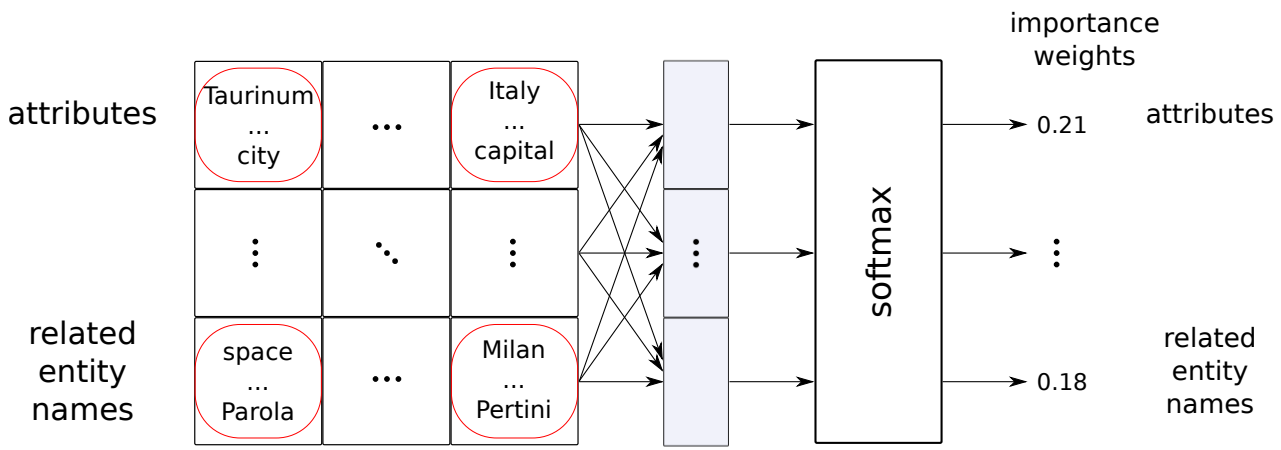
Step 2: create document fields interaction matrix for each query phrase



Document Field Attention Layers

Goal: compute the importance weights of document fields for aggregating the matching scores of query phrases

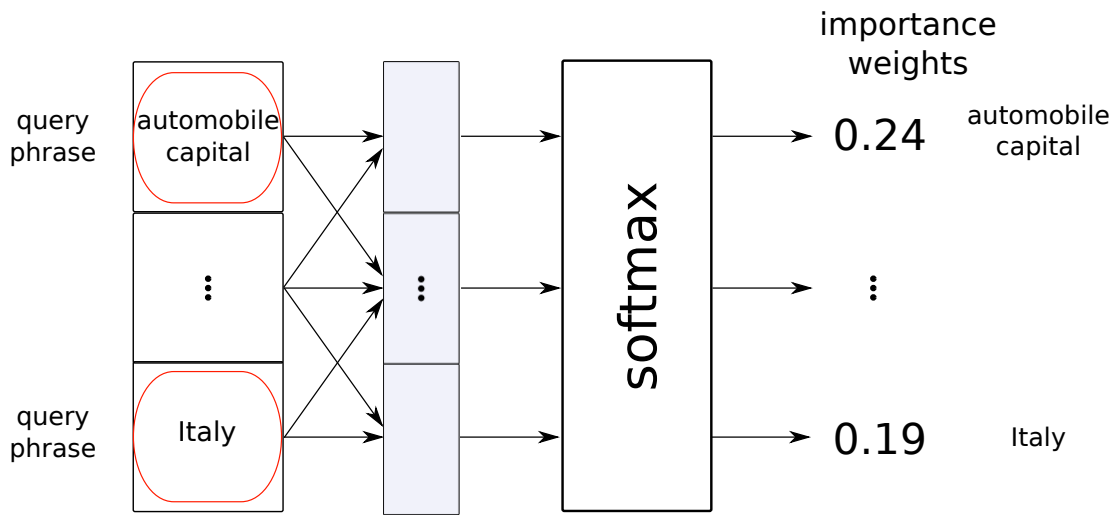
Document: `http://dbpedia.org/page/Turin`



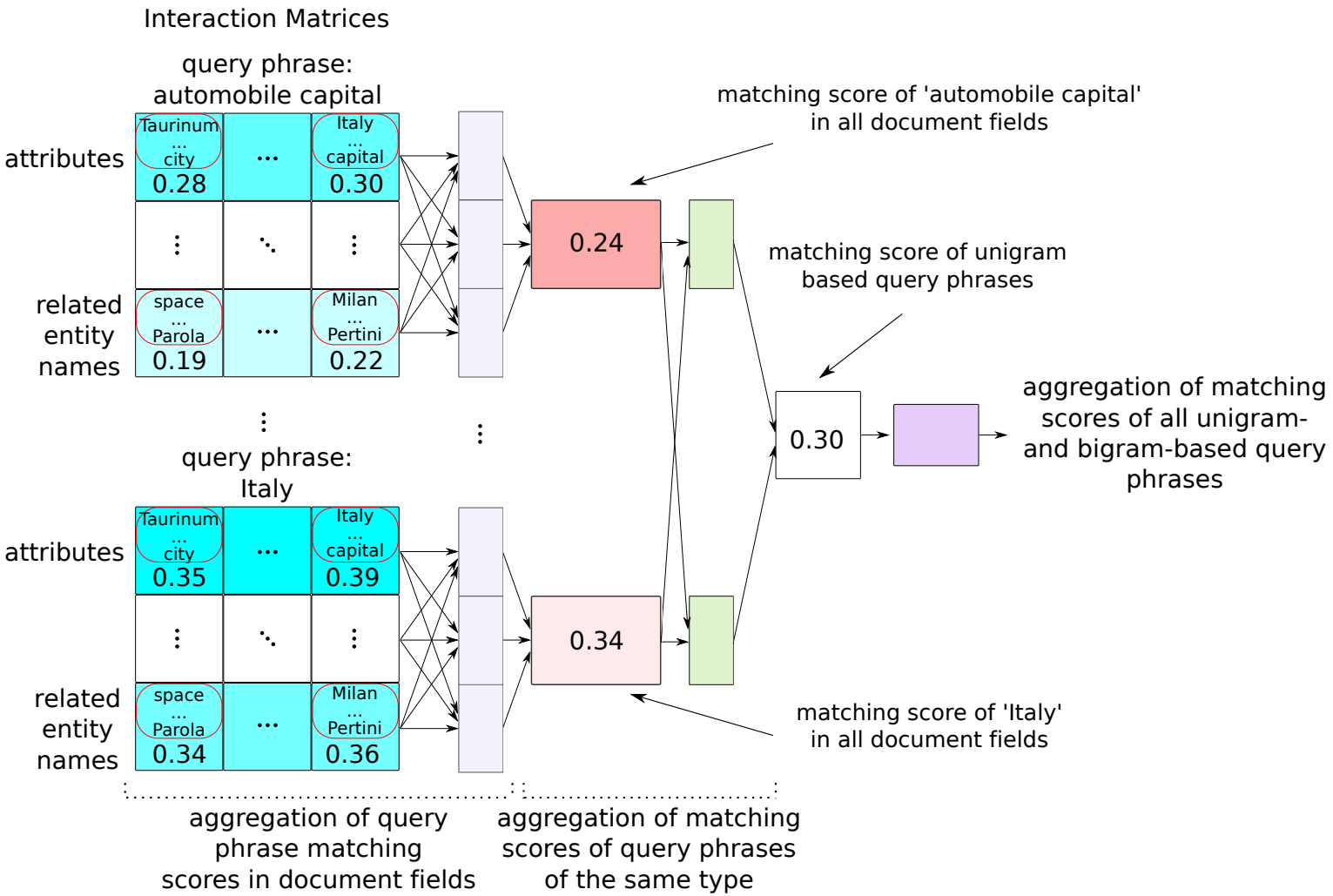
Query Phrase Attention Layers

Goal: compute the importance weights of query phrases for aggregating the matching scores of query phrases of the same type

Query: automobile capital and the Detroit of Italy



Matching Score Aggregation Layers



ANSR is trained to minimize contrastive max-margin loss, given a collection of triplets $<q, d^{\mathrm{n}}, d^{\mathrm{r}}>$ consisting of relevant d^{r} and non-relevant d^{n} documents for query q:

$$\min_{\mathcal{W}} \left(\sum_{<q,d^{\mathrm{n}},d^{\mathrm{r}}>\in\mathcal{T}} \max(0, \zeta - s(q, d^{\mathrm{r}}) + s(q, d^{\mathrm{n}})) + \frac{\gamma}{2}||\mathcal{W}||_2^2 \right)$$

- Language modeling and probabilistic baselines:
  - PRMS (Probabilistic Retrieval Model for Semistructured Data) [Kim, Xue and Croft, ECIR'09]
  - MLM (Mixture of Language Models) [Ogilvie and Callan, SIGIR'03]
  - BM25F [Robertson, Zaragoza and Taylor, CIKM'04]
  - FSDM (Fielded Sequential Dependence Model) [Zhiltsov, Kotov and Nikolaev, SIGIR'15]
- Neural baselines:
  - DRMM (Deep Relevance Matching Model) [Guo, Fan, Ai and Croft, CIKM'16]
  - DESM (Dual Embedding Space Model) [Nalisnick, Mitra, Craswell and Caruanan, WWW'16]
  - NRM-F (Neural Ranking Model with Multiple Document Fields) [Zamani, Mitra, Song, Craswell and Tiwary, WSDM'18]

Performance of ANSR and the baselines

GOV2 collection

	MAP	P@10	NDCG@10
PRMS	0.1964 (-39.49%)	0.4058 (-32.62%)	0.3448 (-30.16%)
MLM	0.2908 (-10.41%)	0.5648 (-6.23%)	0.4729 (-4.21%)
BM25F	0.2954 (-9.00%)	0.5478 (-9.05%)	0.4556 (-7.72%)
FSDM	0.3012 (-7.21%)	0.5817 (-3.42%)	0.4789 (-3.00%)
DESM	0.2968 (-8.56%)	0.5714 (-5.13%)	0.4575 (-7.33%)
DRMM	0.3113 (4.10%)	0.5880 (-2.37%)	0.4722 (-4.35%)
NRM-F*	0.1491 (-54.07%)	0.2903 (-51.80%)	0.2132 (-56.82%)
ANSR	0.3246	0.6023	0.4937

ANSR achieved 7.21% and 3% improvement over FSDM in terms of MAP and NDCG@10 and 4.35% improvement over DRMM in terms of NDCG@10

Performance of ANSR and the baselines

HomeDepot collection

	MAP	P@10	NDCG@10
PRMS	0.2287 (-19.64%)	0.1080 (-21.57%)	0.2641 (-17.57%)
MLM	0.2476 (-13.00%)	0.1183 (-14.09%)	0.2893 (-9.71%)
BM25F	0.2537 (-10.86%)	0.1201 (-12.78%)	0.2952 (-7.87%)
FSDM	0.2591 (-8.96%)	0.1206 (-12.42%)	0.3024 (-5.62%)
DESM	0.2349 (-17.46%)	0.1107 (-19.61%)	0.2943 (-8.15%)
DRMM	0.2484 (-12.72%)	0.1131 (-17.86%)	0.2952 (-7.87%)
NRM-F*	0.1536 (-46.03%)	0.0723 (-47.49%)	0.1832 (-42.82%)
ANSR	0.2846	0.1377	0.3204

ANSR achieved 8.96% and 5.62% improvement over FSDM as well as 12.72% and 7.87% improvement over DRMM in terms of MAP and NDCG@10

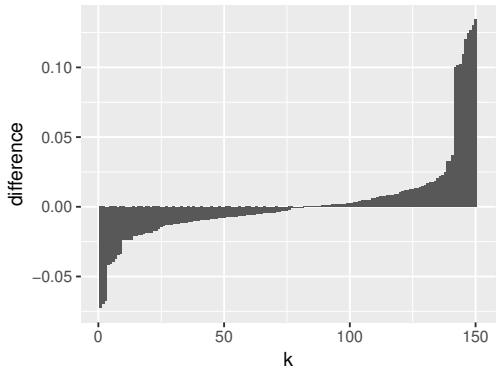
Performance of ANSR and the baselines

DBPedia-v2 collection

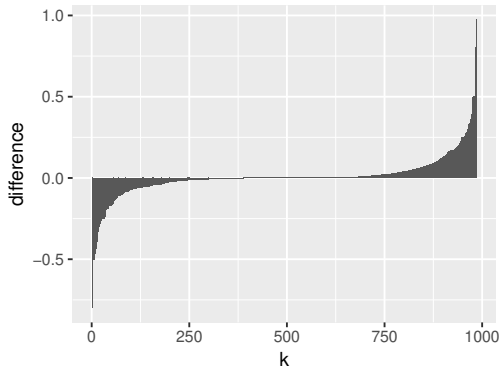
	MAP	P@10	NDCG@10
PRMS	0.2934 (-26.50%)	0.3594 (-15.55%)	0.4126 (-14.26%)
MLM	0.3467 (-13.15%)	0.3887 (-8.67%)	0.4365 (-9.29%)
BM25F	0.3799 (-4.83%)	0.4077 (-4.21%)	0.4605 (-4.30%)
FSDM	0.3679 (-7.84%)	0.4073 (-4.30%)	0.4524 (-5.99%)
DESM	0.3523 (-11.75%)	0.3894 (-8.51%)	0.4527 (-5.92%)
DRMM	0.3682 (-7.77%)	0.4012 (-5.73%)	0.4515 (-6.17%)
NRM-F*	0.1878 (-52.96%)	0.2092 (-50.85%)	0.2402 (-50.08%)
ANSR	0.3992	0.4256	0.4812

ANSR achieved 4.83% and 4.30% improvement over BM25F as well as 7.77% and 6.17% improvement over DRMM in terms of MAP and NDCG@10

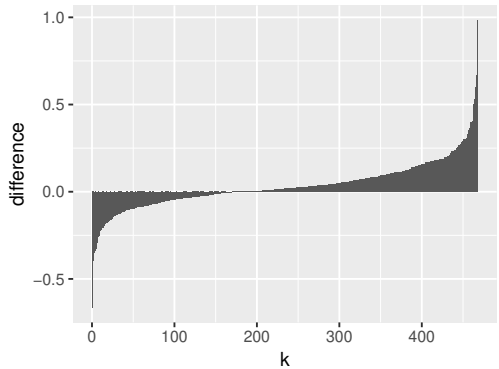
Topic-level difference in retrieval accuracy between ANSR and FSDM



(a) GOV2



(b) HomeDepot

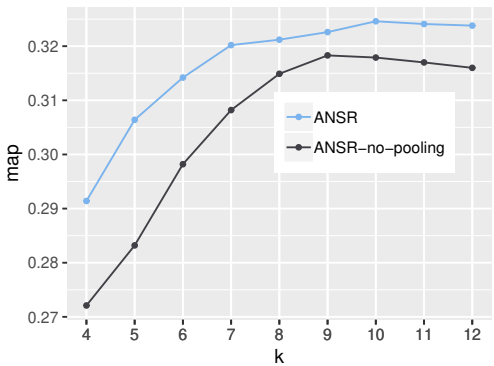


(c) DBpedia-v2

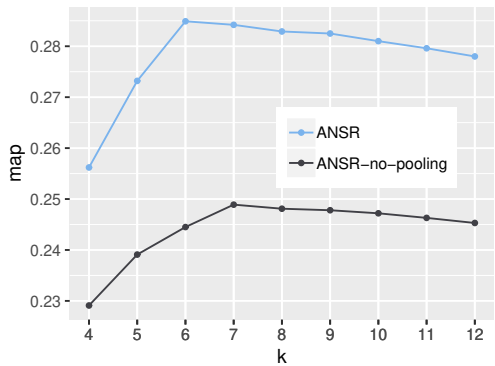
- ANSR has higher average precision than FSDM for 58.88% of the queries in HomeDepot collection
- In GOV2, the magnitude of improvements in average precision is 1.66 times greater than the magnitude of reductions
- Superior ability of ANSR to deal with long field documents, due to utilization of compressed representations and explicit correction of the pooling bias

The effect of the pooling size (k) on the performance of ANSR and ANSR-no-pooling

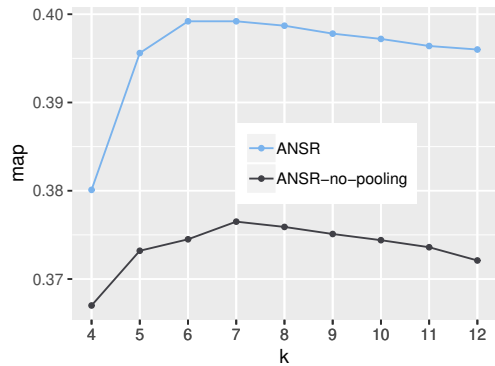
ANSR-no-pooling: select the first k terms in each document field instead of pooling



(a) GOV2



(b) HomeDepot



(c) DBpedia-v2

- ANSR has substantially better retrieval accuracy in terms of MAP than ANSR-no-pooling
- The optimal value of k depends on the collection and the retrieval task: ANSR has the best performance on GOV2, DBpedia-v2 and HomeDepot collections when $k = 10$, $k = 6$ and $k = 6$, respectively

Best performing queries in comparison to FSDM

- The best performing query is “single lever hole bathroom sink faucet”:
  - only one relevant document with the title *“Belle Foret Single Hole 1-Handle High Arc Bathroom Vessel Faucet in Chrome with Metal Lever Handles”* in relevance judgments
  - This document has longer fields than the average field length in this collection

Worst performing queries in comparison to FSDM

- The worst performing query is "popular":
  - only one relevant document with the title *"Bloomsz Most Popular Water Plant Collection (8-Pack)"* in relevance judgments
  - ANSR ranked the document with the title *"South Shore Furniture Popular Twin Mates Bed in Mocha"* as the top-ranked document, since it has more words that are semantically similar to the query term *"popular"*
  - This can be a consequence of using word embeddings by ANSR, which can cause topic drift for very short queries

Summary

- ANSR utilizes **pooling** to generate fixed-size interactions matrices between representations of phrases in a query and document fields and employs an **attention mechanism** to focus on the most important document fields and query phrases
- ANSR includes the layers to compute and aggregate the relevance score of a structured document at **different levels of granularity**
- ANSR outperforms state-of-the-art LM and neural baselines in different SDR tasks, such as **Web search**, **product search** and **entity retrieval from a knowledge graph**.

Thank you! Questions?