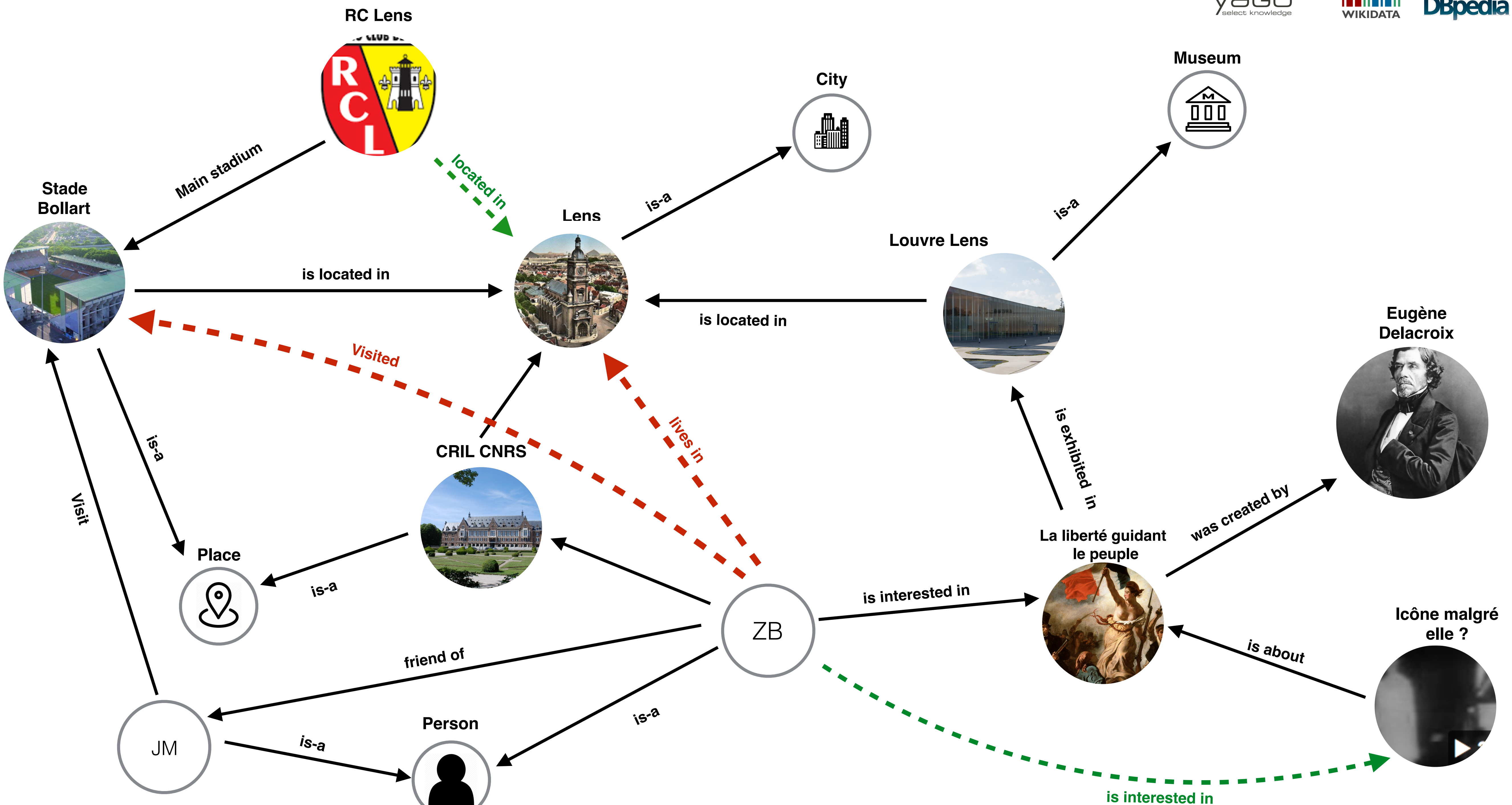
Inductive Reasoning with Conceptual Space Representations

Zied Bouraoui
(joint work with Steven Schockeart)

Université d'Artois
CRIL-CNRS UMR 8188
bouraoui@cril.univ-artois.fr
http://www.cril.univ-artois.fr/~bouraoui/

FCA4AI 2018 Workshop at IJCAI-ECAI 2018 Stockholm (Sweden)

STRUCTURED KNOWLEDGE AND THE WEB



- Require KBs with a wide coverage, even if that means accepting some inaccuracies
- Deductive reasoning is often too limited in this setting

Plausible Inference Patterns

PLAUSIBLE REASONING

Similarity-based reasoning

Marion enjoys hiking in the Alps

The Alps **are similar** to the Pyrenees

Marion enjoys hiking in the Pyrenees

Category-based reasoning

Athletics **is regulated** by the International Olympic Committee

Swimming **is regulated** by the International Olympic Committee

Athletics and Swimming are **representative examples** of Olympic games

---

All Olympic games **are regulated** by International Olympic Committee

PLAUSIBLE REASONING

Betweenness

Bars in France are required to display alcoholic beverage license

Restaurants in France are required to display alcoholic beverage license

Brasseries are **conceptually between** Bars and Restaurants

Brasseries in France are required to display alcoholic beverage license

Extrapolation

A beef steak **pairs well** with Médoc

A beef tartare **pairs well** with Dolcetto

Poached salmon **pairs well** with Chardonnay

Pinot Gris **is dryer and lighter** than Chardonnay, in the same way as Dolcetto is dryer and lighter than Médoc

---

Salmon Carpaccio **pairs well** with Pinot Gris

HUMAN REASONING VS CLASSICAL LOGIC

- Human reasoning captures **statistical regularities**, rather than tautologies
- **Learning from examples** is essential for building human knowledge
- **Natural Language** is the central in human reasoning

…but, these observations (and others) cannot always be fully captured in logic

CONCEPTUAL SPACE

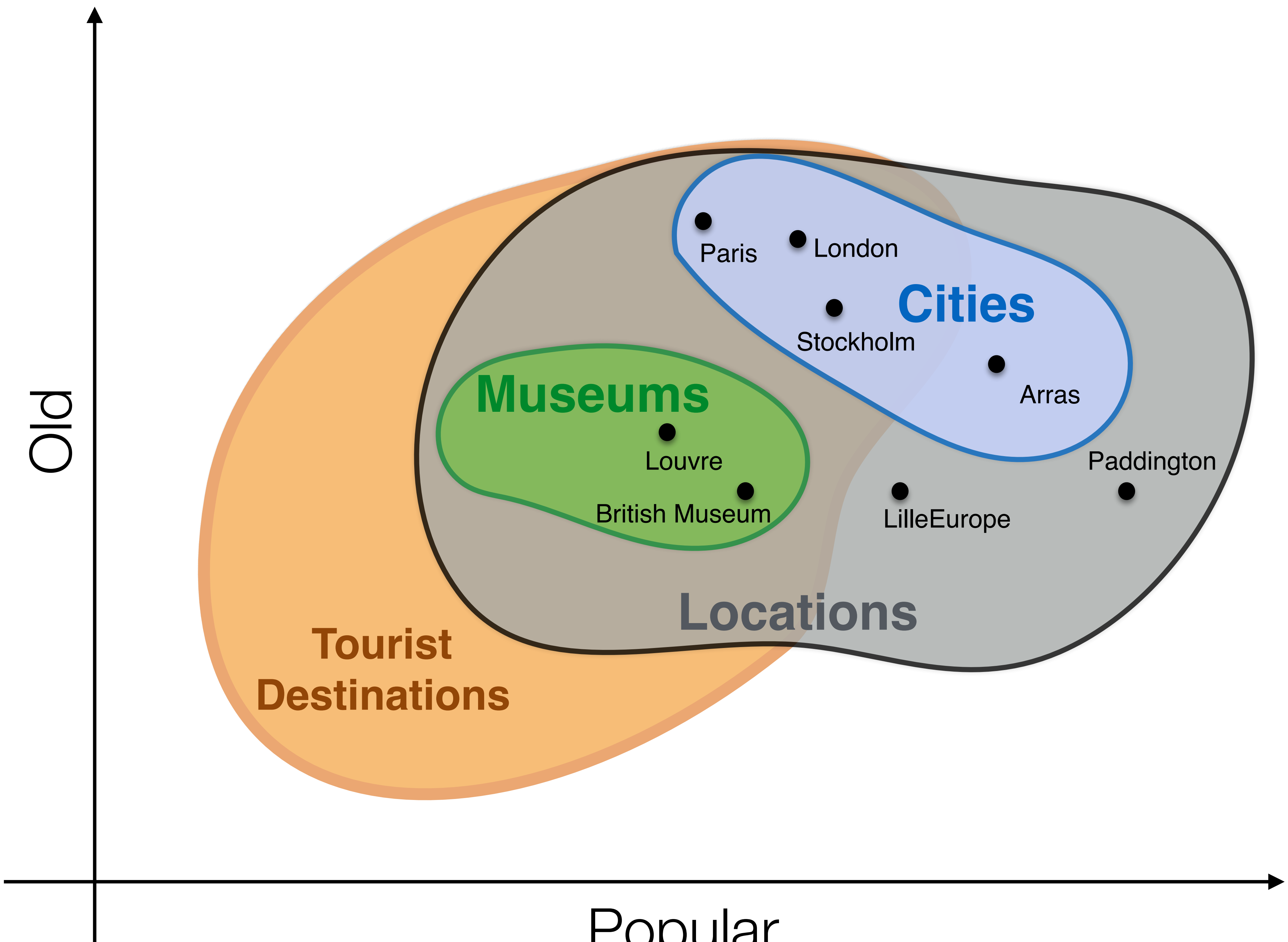
Symbolic — Propositional representation

Conceptual — Geometric representation

Associationist — Connectionist representation

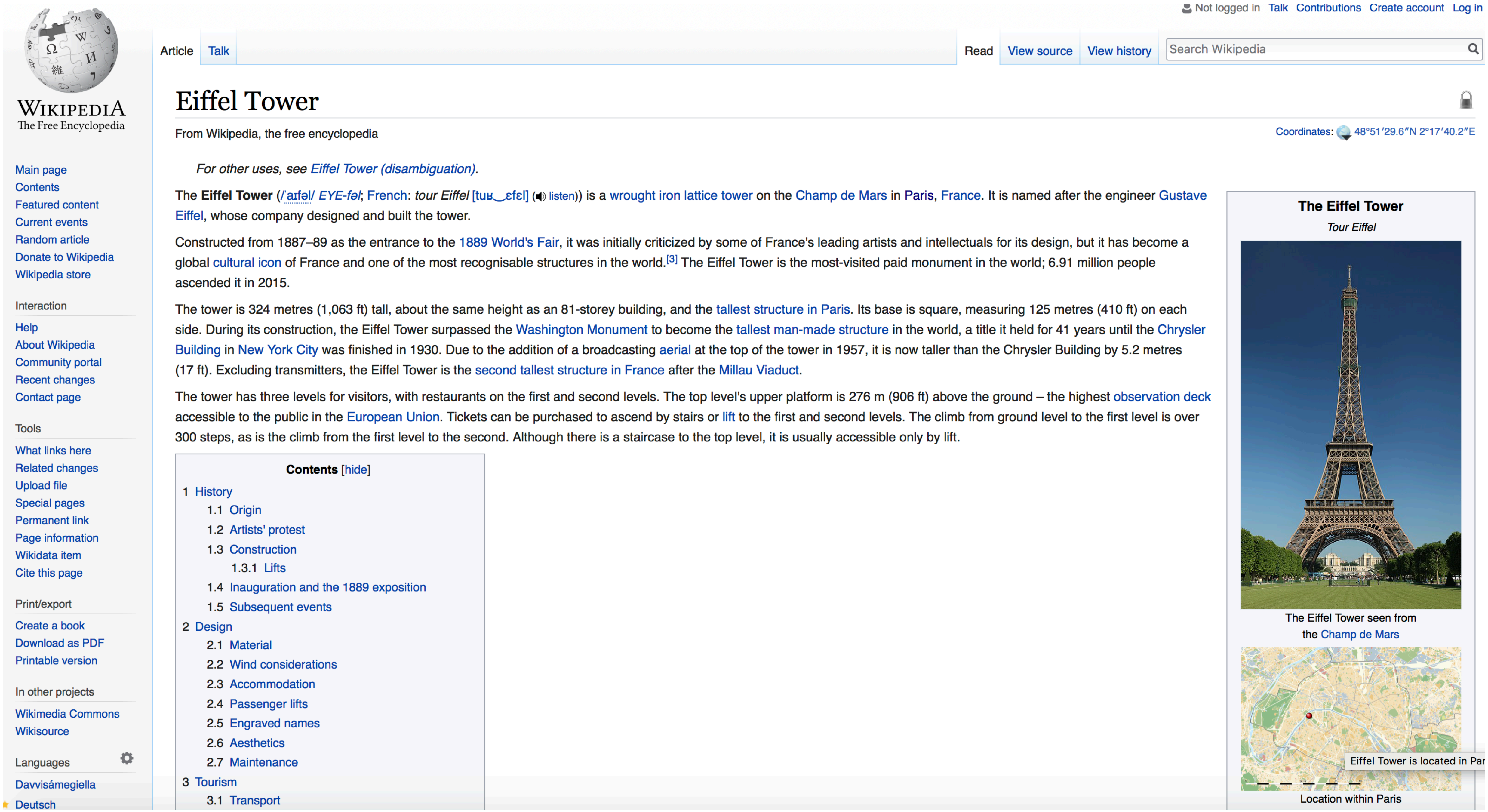
CONCEPTUAL SPACES

A conceptual space is defined as the Cartesian product of a number of so-called quality dimensions



Learning Object Representations From Data

DATA SOURCE: TEXT DESCRIPTIONS



Not logged in Talk Contributions Create account Log in

Article Talk | Read View source View history | Search Wikipedia

WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikipedia store

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact page

Tools
What links here
Related changes
Upload file
Special pages
Permanent link
Page information
Wikidata item
Cite this page

Print/export
Create a book
Download as PDF
Printable version

In other projects
Wikimedia Commons
Wikisource

Languages
Davvisámegiella
Deutsch

Eiffel Tower

From Wikipedia, the free encyclopedia

Coordinates: 48°51′29.6″N 2°17′40.2″E

For other uses, see Eiffel Tower (disambiguation).

The **Eiffel Tower** (/ˈaɪfəl/ *EYE-fəl*; French: *tour Eiffel* [tuʁ‿ɛfɛl] (listen)) is a wrought iron lattice tower on the Champ de Mars in Paris, France. It is named after the engineer Gustave Eiffel, whose company designed and built the tower.

Constructed from 1887–89 as the entrance to the 1889 World's Fair, it was initially criticized by some of France's leading artists and intellectuals for its design, but it has become a global cultural icon of France and one of the most recognisable structures in the world.[3] The Eiffel Tower is the most-visited paid monument in the world; 6.91 million people ascended it in 2015.

The tower is 324 metres (1,063 ft) tall, about the same height as an 81-storey building, and the tallest structure in Paris. Its base is square, measuring 125 metres (410 ft) on each side. During its construction, the Eiffel Tower surpassed the Washington Monument to become the tallest man-made structure in the world, a title it held for 41 years until the Chrysler Building in New York City was finished in 1930. Due to the addition of a broadcasting aerial at the top of the tower in 1957, it is now taller than the Chrysler Building by 5.2 metres (17 ft). Excluding transmitters, the Eiffel Tower is the second tallest structure in France after the Millau Viaduct.

The tower has three levels for visitors, with restaurants on the first and second levels. The top level's upper platform is 276 m (906 ft) above the ground – the highest observation deck accessible to the public in the European Union. Tickets can be purchased to ascend by stairs or lift to the first and second levels. The climb from ground level to the first level is over 300 steps, as is the climb from the first level to the second. Although there is a staircase to the top level, it is usually accessible only by lift.

Contents [hide]

1 History
1.1 Origin
1.2 Artists' protest
1.3 Construction
1.3.1 Lifts
1.4 Inauguration and the 1889 exposition
1.5 Subsequent events
2 Design
2.1 Material
2.2 Wind considerations
2.3 Accommodation
2.4 Passenger lifts
2.5 Engraved names
2.6 Aesthetics
2.7 Maintenance
3 Tourism
3.1 Transport

The Eiffel Tower

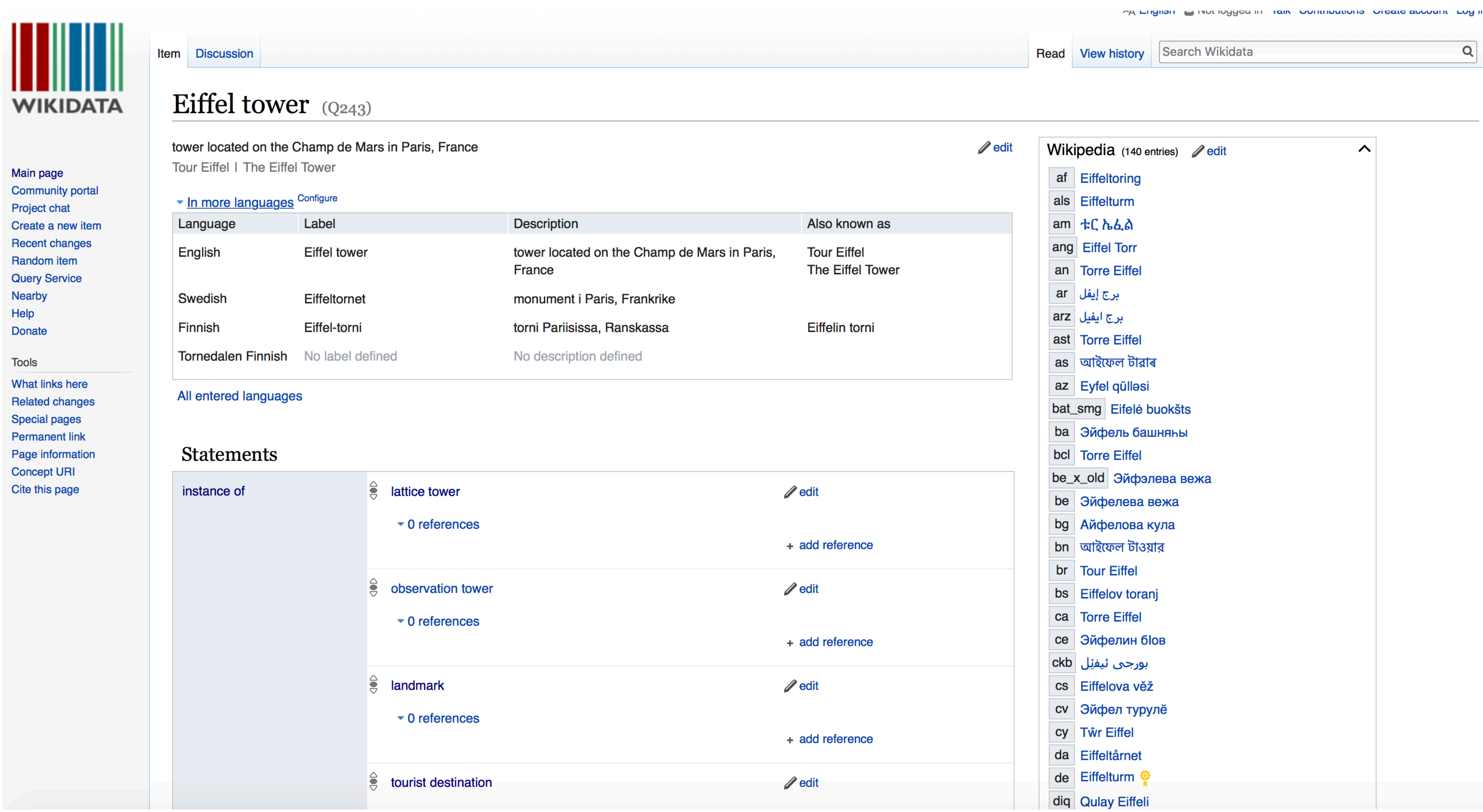
Tour Eiffel

The Eiffel Tower seen from the Champ de Mars

Eiffel Tower is located in Par

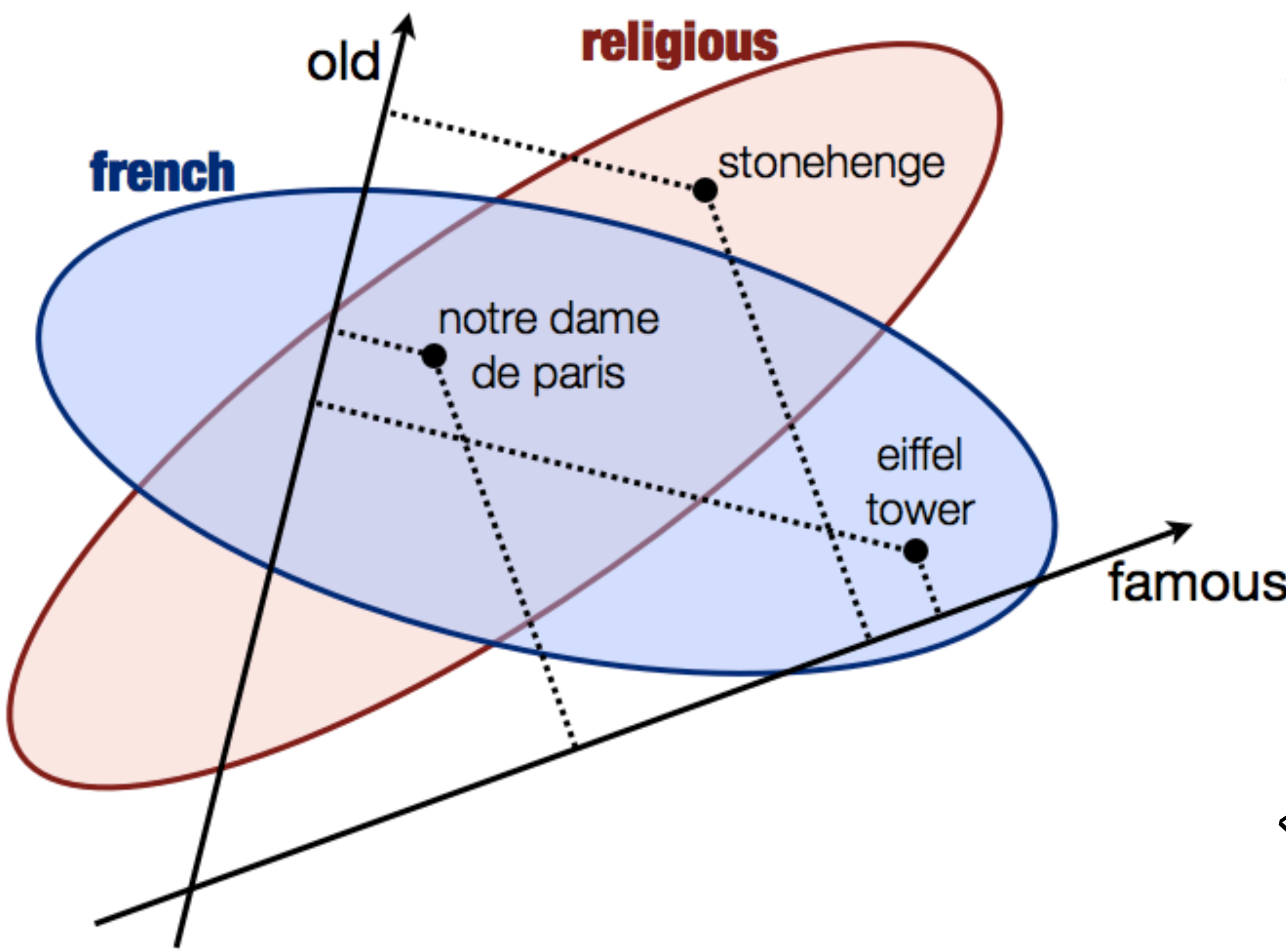
Location within Paris

DATA SOURCE: KNOWLEDGE GRAPHS



LEARNING CONCEPTUAL SPACES

Low-dimensional vector space embedding ⇔ structured knowledge + bag-of-words representations



300 dimensions
WikiData + Wikipedia

« most popular french tourist destinations »

LEARNING CONCEPTUAL SPACES

$$J = \alpha(J_{text} + J_{glove}) + (1-\alpha)(J_{type} + J_{rel}) + \beta J_{reg}$$

ordinal SVM regression with quadratic kernel

+

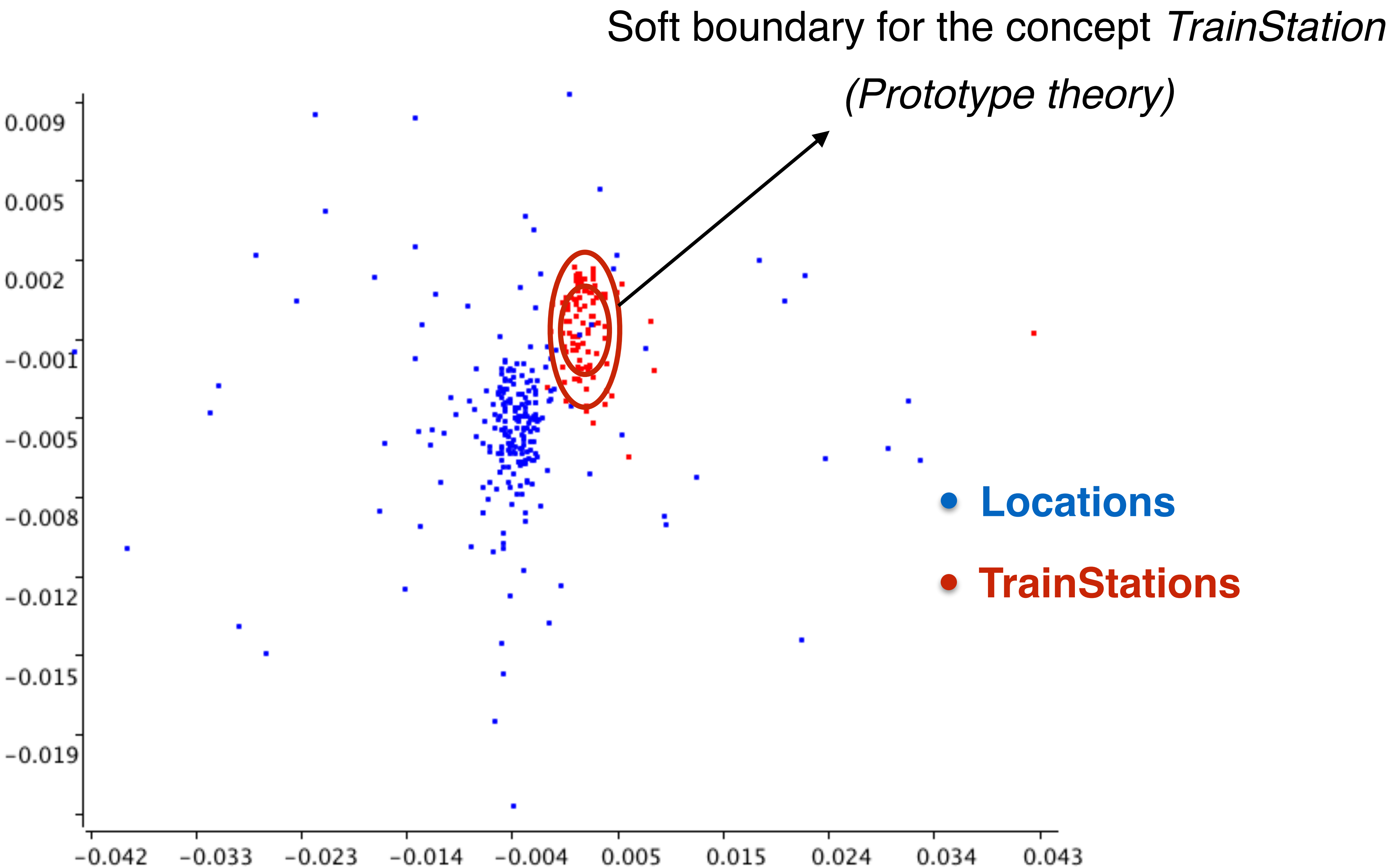
nuclear norm regularisation

+

Constrain representations based on knowledge graph triples

Learning Concept Representations

MAIN IDEA



Soft boundary for the concept *TrainStation*
(Prototype theory)
0.009
0.005
0.002
−0.001
−0.005
−0.008
−0.012
−0.015
−0.019
−0.042
−0.033
−0.023
−0.014
−0.004
0.005
0.015
0.024
0.034
0.043
Locations
TrainStations

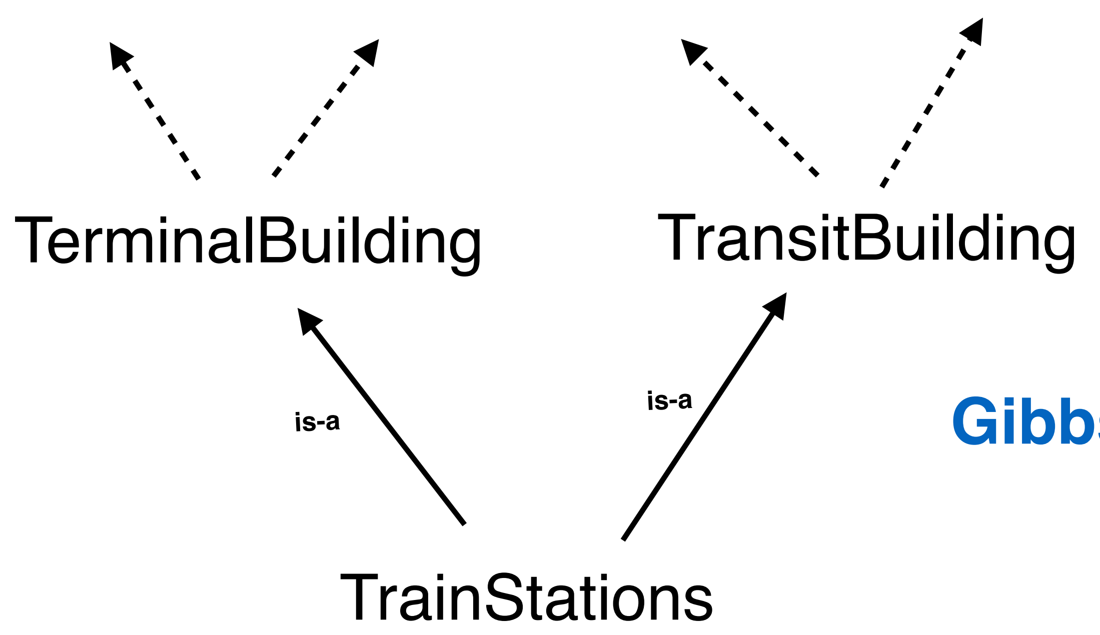
CONCEPT REPRESENTATION

Control how common the instances are

$$P(C|v_a) = \lambda_C \cdot G_C(v_a)$$

The variance of this Gaussian encodes how much the instances are dispersed across the space

Parameters of the Gaussians…



Gibbs Sampling to cope with cyclic dependencies

HANDLING CYCLIC DEPENDENCIES

Generate sequences of parameters μ_{C0}, μ_{C1},… and Σ_{C0}, Σ_{C1},… for each concept

Steps:

- Init parameters μ_{C0} and Σ_{C0}
- repeatedly iterate over all concepts and in the i^{th} iteration, choose the next samples μ_{Ci} and Σ_{Ci} for each concept C

Use known dependencies between concepts to construct informative priors on μ_{Ci} and Σ_{Ci}

PRIORS ON THE MEAN

It depends whether the concept is atomic or complex (Description Logic is used)

Used information:

1. If $A \sqsubseteq C$ holds then μ_A should correspond to a plausible instance of C. In particular, we would expect the probability $G^*_C(\mu_A)$ to be high

2. Vector representation v_A of A: Suppose $B_1 \sqsubseteq C$, $B_2 \sqsubseteq C$, …, $B_r \sqsubseteq C$, then $v_{B1} - \mu_{B1}$, $v_{B2} - \mu_{B2}, \ldots, v_{Br} - \mu_{Br}$ should be similar to $v_A - \mu_A$.

3. We do not have vector representation, but we have more logical structure

PRIORS ON THE VARIANCE

It depends whether the concept is atomic or complex

Used information :

- If A $\sqsubseteq$ C holds then $\Sigma_A \leqslant \Sigma_C$ should hold
- If $B_1 \sqsubseteq C$, $B_2 \sqsubseteq C$, …, $B_r \sqsubseteq C$, then one can consider Σ_A as the average of Σ_{B1} , Σ_{B2} ,…, Σ_{B3} (use most similar siblings, i.e closest in terms of Euclidean distance)

KNOWLEDGE BASE COMPLETION

Average over the Gibbs samples

$$P(C|v) = \frac{\lambda_C}{N} \sum_{i=1}^{N} p(v; \mu_C^i, \Sigma_C^i)$$

maximizing the likelihood to obtain estimates of the scaling parameters λ

$$\sum_{i=1}^{s} \log(\lambda_C P(v_i|C)) + \sum_{i=1}^{r} \log(1 - \lambda_C P(u_i|C))$$

Induction with Conceptual Space representations

FACT INDUCTION - INTUITIONS

Locations

Nancy, Louvre,

Gare du Nord,

Lille Flandre,

…

is-a

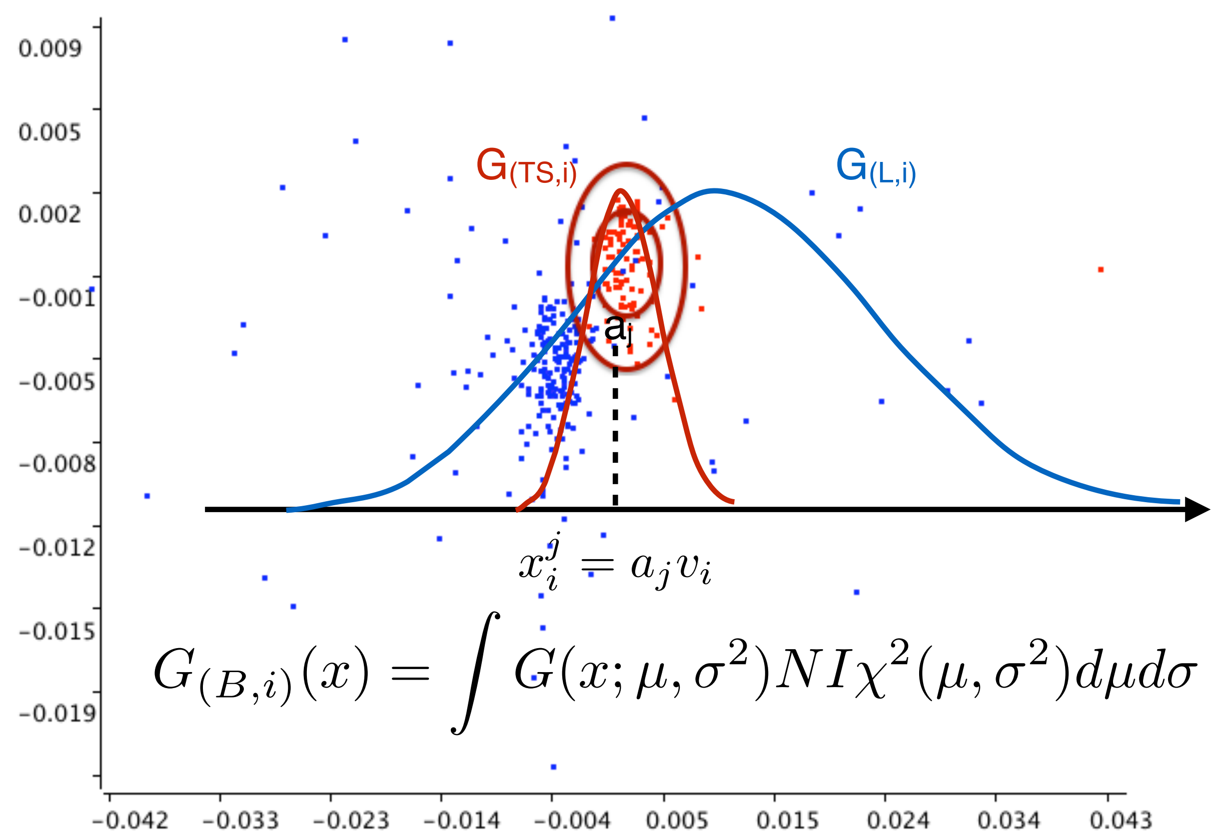
TrainStations

Gare du Nord,

Lille Europe,

Gare de Nancy,

…



Knowing that Gare du Nord, Lille Europe, Gare de Nancy, … all have some property P, can we conclude that some other entities (e.g. Lille Flandre) has property P ?

Unknown

$$p(T \mid x_i^j, L) = \frac{p(x_i^j \mid T, L) \cdot p(T \mid L)}{p(x_i^j \mid L)} \propto \prod_i \frac{G_{(T,i)}(x_i^j)}{G_{(L,i)}(x_i^j)}$$

RULE INDUCTION - INTUITIONS

Finding missing rules from a given (existential) knowledge base

Interpolation

$$\left.\begin{aligned} r_1(X) \wedge orange(X) \rightarrow r_2(X) \\ r_1(X) \wedge lemon(X) \rightarrow r_2(X) \\ r_1(X) \wedge grapefruit(X) \rightarrow r_2(X) \end{aligned}\right\} \; r_1(X) \wedge lime(X) \rightarrow r_2(X)$$

Analogy

$$\left.\begin{aligned} r_1(X,Y) \wedge bat(X) \rightarrow cave(Y) \\ r_1(X,Y) \wedge duck(X) \rightarrow pond(Y) \\ r_1(X,Y) \wedge dolphin(X) \rightarrow sea(Y) \end{aligned}\right\} \; r_1(X,Y) \wedge trout(X) \rightarrow river(Y)$$

RULE INDUCTION

Unary templates: Probability that a given template satisfies a relation r, knowing that it satisfies the relations $r_1, \ldots, r_n$.

$$P(\tau(r) \mid v_r) = \lambda_\tau \cdot \frac{f(v_r \mid \tau(r))}{f(v_r)}$$

Binary templates: The probability that a relation pair (r,s) satisfies ta given binary template

$$\begin{aligned} P(\tau(r,s) \mid v_r, v_s, u_{r,s}) = & \lambda_\tau \cdot \frac{f(v_r|\tau(r,\bullet))}{f(v_r)} \cdot \frac{f(v_s|\tau(\star,s))}{f(v_s)} \\ & \cdot \frac{f(v_s - v_r|\tau(r,s))}{f(v_s - v_r|\tau(r,\bullet),\tau(\star,s))} \\ & \cdot f(u_{r,s}|\tau(r,s)) \end{aligned}$$

EXPERIMENTAL RESULTS - FACT INDUCTION

	SVM-Linear				SVM-Quad				Gibbs			
	Pr	Rec	F1	AP	Pr	Rec	F1	AP	Pr	Rec	F1	AP
$1 \leq \lvert X \rvert \leq 5$	0.033	0.509	0.062	0.055	0.086	0.046	0.060	0.144	0.258	0.508	0.343	0.328
$5 < \lvert X \rvert \leq 10$	0.084	0.922	0.154	0.067	0.116	0.404	0.180	0.163	0.202	0.474	0.283	0.340
$10 < \lvert X \rvert \leq 50$	0.111	0.948	0.199	0.081	0.151	0.382	0.216	0.247	0.242	0.886	0.380	0.276
$\lvert X \rvert > 50$	0.153	0.217	0.180	0.230	0.224	0.721	0.342	0.260	0.361	0.678	0.471	0.404

Table 1: Results of the proposed model and the baselines.

	Gibbs-flat				Gibbs-emb				Gibbs-DL			
	Pr	Rec	F1	AP	Pr	Rec	F1	AP	Pr	Rec	F1	AP
$1 \leq \lvert X \rvert \leq 5$	0.212	0.416	0.281	0.290	0.201	0.540	0.293	0.262	0.226	0,498	0.311	0.304
$5 < \lvert X \rvert \leq 10$	0.186	0.368	0.247	0.273	0.173	0.357	0.233	0.262	0.417	0.192	0.263	0.328
$10 < \lvert X \rvert \leq 50$	0.199	0.496	0.284	0.210	0.207	0.513	0.295	0.233	0.218	0.670	0.329	0.251
$\lvert X \rvert > 50$	0.316	0.312	0.314	0.328	0.321	0.373	0.345	0.321	0.344	0.450	0.390	0.369

Table 2: Results for the variants of the proposed model.

EXPERIMENTAL RESULTS - RULE INDUCTION

		SUMO	OpenCyc	Wine
AS	Pr	0.364	0.423	0.547
AS	Rec	0.457	0.539	0.615
AS	F1	0.405	0.474	0.579
AS	Pr@10	0.517	0.589	0.602
AS	Pr@100	0.402	0.445	n/a
VS	Pr	0.426	0.506	0.582
VS	Rec	0.514	0.613	0.659
VS	F1	0.465	0.554	0.618
VS	Pr@10	0.583	0.657	0.629
VS	Pr@100	0.479	0.573	n/a
RI$-\mathcal{R}$	Pr	0.616	0.639	0.734
RI$-\mathcal{R}$	Rec	0.483	0.512	0.611
RI$-\mathcal{R}$	F1	0.541	0.568	0.666
RI$-\mathcal{R}$	Pr@10	0.734	0.741	0.792
RI$-\mathcal{R}$	Pr@100	0.712	0.723	n/a
RI-*word*	Pr	0.642	0.703	0.782
RI-*word*	Rec	0.451	0.528	0.636
RI-*word*	F1	0.529	0.603	0.701
RI-*word*	Pr@10	0.755	0,789	0.811
RI-*word*	Pr@100	0.727	0.765	n/a
RI	Pr	0.692	0.745	0.813
RI	Rec	0.534	0.586	0.672
RI	F1	0.602	0.656	0.735
RI	Pr@10	0.802	0.834	0.834
RI	Pr@100	0.788	0.811	n/a

REFERENCES

- Zied Bouraoui and Steven Schockaert, Learning Conceptual Space Representations of Interrelated Concepts, 27th International Joint Conference on Artificial Intelligence and the 23rd European Conference on Artificial Intelligence **(IJCAI 2018)**.

- Zied Bouraoui, Shoaib Jameel, Steven Schockaert, Relation induction in word embeddings revisited, Proceedings of the 27th International Conference on Computational Linguistics **(COLING 2018)**.

- Shoaib Jameel, Zied Bouraoui, Steven Schockaert: Unsupervised Learning of Distributional Relation Vectors . In 56th Annual Meeting of the Association for Computational Linguistics **(ACL 2018).**

- Zied Bouraoui, Shoaib Jameel and Steven Schockaert, Inductive Reasoning about Ontologies Using Conceptual Spaces.P roceedings of the thirty-first AAAI Conference on Artificial Intelligence **(AAAI 2017)**.

- Shoaib Jameel, Zied Bouraoui, Steven Schockaert: MEmbER: Max-Margin Based Embeddings for Entity Retrieval. In 40th ACM SIGIR conference on research and development in Information Retrieval **(SIGIR 2017).**