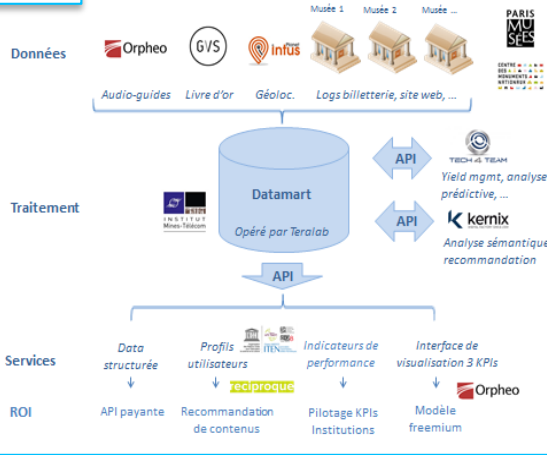


Nada Mimouni, Jean-Claude Moissinac, and Anh Tuan Vu
Télécom ParisTech, Institut Mines Télécom

Data&Musée Project



Data

- Gather data from institutions:**
- Collect data respecting privacy
 - Adopt homogeneous representation
 - Choose a model able to represent links between data

Rely on external data:

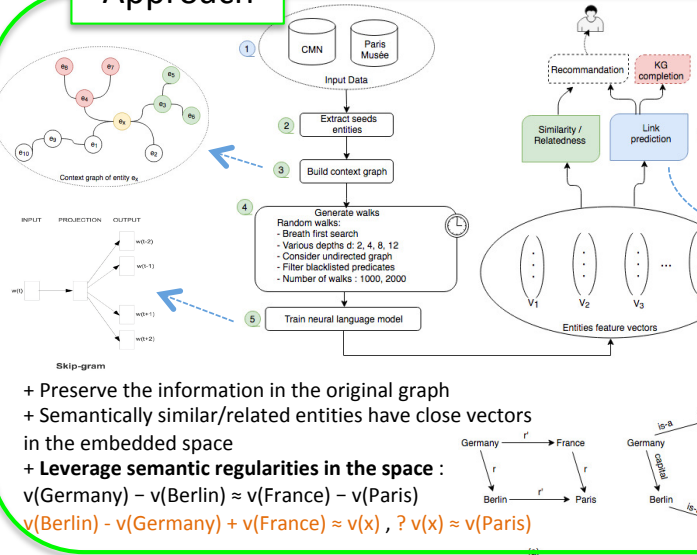
- DataTourisme, tourist office data on places and events
- OpenAgenda, and other event calendar
- Joconde database, and other cultural data
- General knowledge bases: DBpedia, Wikidata, ...
- Geolocation data: geonames, data on data.gouv.fr, ...

Objectives

- Create an interoperable platform for the **aggregation of museum data**
- Develop **steering indicators** for museums
- Build tools for **personalizing visitor experience**



Approach



- Adapt neural language model
- Transform RDF graph into sequences of entities and relations (sentences)
- Train the model and generate entity vectors

Evaluation

Dataset

- Build context graph for Paris Musées
- 11 sites: 448309 entities, 2285 relations, 5122879 triples
- Random walks with BFS: d={4,8}
- Add all direct neighbours: d=2
- Train Skip-Gram model

Ground Truth

- Analogy between entities in the KG (each entry corresponds to a parallelogram with one unobserved triple in the KG)
- For each museum entity, collect a list of well-known artists
- Use conventional metrics: Mean Reciprocal Rank (MRR) and Hits@{3,5,10}

Results

Museums for which we exploit known triples <museum, artist>

- For m_1 , find a_1 , such that $\langle m_1, a_1 \rangle$ a correct triple
- Where: $\langle m_1, a_1 \rangle$ exist

→ Longer path (D=8) gives better results

Artists for which we exploit one known triple <museum, artist>

- For m_1 , find a_i , such that $\langle m_1, a_i \rangle$ correct triple
- Where: $\langle m_1, a_i \rangle$ an existing triple (suggested by an expert)

Entity m_1	MRR		H@3%		H@5%		H@10%	
	D=4	D=8	D=4	D=8	D=4	D=8	D=4	D=8
Entity m_1	W=1000	2000	1000	1000	1000	1000	1000	1000
dbr:Musée_Carnavalet	0,41	0,06	0,27	0,45	0,00	0,36	0,55	0,09
dbr:Musée_de_Grenoble	0,77	0,76	0,61	0,73	0,73	0,82	0,91	0,82
dbr:Musée_Zadkine	0,39	0,42	0,70	0,36	0,64	0,82	0,36	0,64
dbr:Musée_Cernuschi	0,36	0,48	0,75	0,45	0,55	0,73	0,45	0,55
dbr:Petit_Palais	0,45	0,24	0,50	0,64	0,36	0,45	0,73	0,36
dbr:Maison_de_Balzac	0,47	0,61	0,66	0,55	0,64	0,64	0,64	0,73
dbr:Musée_Cognacq-Jay	0,35	0,50	0,70	0,45	0,64	0,82	0,73	0,82
dbr:Musée_d'art_moderne	0,46	0,34	0,50	0,64	0,45	0,73	0,73	0,55
dbr:Musée_de_la_Vie_Romantique	0,36	0,21	0,20	0,45	0,36	0,18	0,45	0,36
dbr:Palais_Galliera	0,53	0,82	0,58	0,73	0,91	0,82	0,73	0,91
dbr:Maison_de_Victor_Hugo	0,22	0,29	0,70	0,36	0,45	0,73	0,36	0,45
dbr:Musée_Bourdelle	0,43	0,34	0,21	0,45	0,36	0,27	0,55	0,55
All Paris Musées	0,43	0,42	0,53	0,52	0,51	0,61	0,60	0,57

Entity a_1	W=1000	MRR		H@3%		H@5%		H@10%	
		D=4	D=8	D=4	D=8	D=4	D=8	D=4	D=8
dbr:Antoine_Bourdelle	0,61	0,66	0,68	0,73	0,64	0,73	0,82	0,64	0,91
dbr:Israël_Silvestre	0,08	0,23	0,30	0,09	0,27	0,36	0,14	0,27	0,36
dbr:Gustave_Courbet	0,38	0,48	0,50	0,45	0,64	0,64	0,45	0,64	0,73
dbr:Ossip_Zadkine	0,67	0,62	0,75	0,82	0,73	0,82	0,91	0,73	0,91
dbr:Xu_Beihong	0,74	0,36	0,68	1,00	0,45	0,82	1,00	0,64	0,82
dbr:Honoré_de_Balzac	0,54	0,55	0,66	0,73	0,64	0,73	0,82	0,64	0,82
dbr:François_Boucher	0,66	0,58	0,67	0,73	0,64	0,73	0,82	0,73	0,82
dbr:Geer_Van_Velde	0,10	0,27	0,63	0,09	0,36	0,64	0,09	0,64	0,73
dbr:Ary_Scheffer	0,62	0,41	0,51	0,73	0,55	0,45	0,82	0,64	0,82
dbr:Jacques_Heim	0,19	0,18	0,22	0,18	0,27	0,27	0,45	0,27	0,64
dbr:Victor_Hugo	0,53	0,53	0,49	0,64	0,64	0,82	0,73	0,73	0,82
All artists	0,47	0,44	0,55	0,56	0,53	0,64	0,64	0,60	0,76

Comparison with RDF2vec Model :

- Trained on all DBpedia
- MRR = Hits@{3,5,10} = 0

→ Context Graph > DBpedia, for domain applications