Answering Visuo-semantic Queries with IMGpedia

Sebastián Ferrada, Benjamin Bustos, and Aidan Hogan

Center for Semantic Web Research
Department of Computer Science, Universidad de Chile
{sferrada,bebustos,ahogan}@dcc.uchile.cl

Abstract. IMGpedia is a linked dataset that provides a public SPARQL endpoint where users can answer queries that combine the visual similarity of images from Wikimedia Commons and semantic information from existing knowledge-bases. Our demo will show example queries that capture the potential of the current data stored in IMGpedia. We also plan to discuss potential use-cases for the dataset and ways in which we can improve the quality of the information it captures and the expressiveness of its queries.

1 Introduction

Wikimedia Commons[1] is a large-scale dataset that contains about 30 million freely usable media files (image, audio and video), many of which are used within Wikipedia articles and galleries; it also contains meta-data about each file, such as its author, licensing, and the articles where the file is used. Using this information, DBpedia Commons [5] automatically extracts the meta-data of the media files of Wikimedia Commons pages and presents the resulting corpus as a linked dataset.

To compliment DBpedia Commons, we have created IMGpedia [2]: a linked dataset that contains different feature descriptors for 14.8 million images from the Wikimedia Commons. IMGpedia also provides similarity relations among the images, as well as references to DBpedia [3] if the image is used on a Wikipedia article of the entity. This dataset thus enables people to perform visuo-semantic queries, that is, queries that combine image similarity and semantic criteria.

We first introduce IMGpedia. We then show examples of queries that IMGpedia supports, as will be shown in the demo session. Finally we address the challenges and the future directions of the project, which we also plan to discuss in the session.

2 IMGpedia

IMGpedia is a linked dataset [2] that contains three different visual descriptors for each of the 14.8 million images of Wikimedia Commons. These descriptors capture the following features of each image as high-dimensional vectors: brightness distribution, border orientations and color layout. IMGpedia provides static similarity relations among the images: for each image and for each descriptor, the dataset contains the 10 nearest neighbors—that is, the 10 most similar images according to how close they are in the Manhattan distance between their descriptors.

This information is described in RDF using a custom vocabulary that combines novel terms with terms from established vocabularies and appropriate RDFS/OWL definitions (see our extended paper accepted for the ISWC Resources Track for

[1] http://commons.wikimedia.org

more details [2]). Additionaly, IMGpedia contains links to DBpedia entities and to DBpedia Commons in order to obtain further metadata related to the image. Currently we provide ∼12 million links to DBpedia: an image is linked to an entity if the image appears in the Wikipedia article of which the entity is about.

IMGpedia is publicly available as an RDF dump[2] and as a SPARQL endpoint[3].

3 Visuo-semantic Queries

Using SPARQL federation over the IMGpedia and DBpedia datasets, we are able to answer *visuo-semantic* queries—that is, queries that combine visual similarity (e.g. images similar to a given picture of La Moneda Palace, in Santiago) with queries about semantic facts (e.g. obtain a list of governmental palaces in Europe). Hence, an example of a visuo-semantic query would be to *obtain the depictions of the European governmental palaces that are similar to La Moneda Palace*. In this section we show some examples of queries that can be answered using IMGpedia.

First, IMGpedia can answer image similarity queries, since it provides static similarity relations among them. In our extended paper [2], an example of this kind of query – looking for images similar to one of Hopsten Marktplatz in Germany – and the respective results can be found.

We can also perform semantic image retrieval[4]. In Listing 1 we request the images of the paintings made in the 16$^{\text{th}}$ century that are currently being displayed at the Louvre. In Figure 1 we show the results.

Listing 1: Query to retrieve images of paintings from the 16$^{\text{th}}$ century that are displayed at the Louvre.

```
SELECT ?url ?label WHERE {
  SERVICE <http://dbpedia.org/sparql> {
    ?res a yago:Wikicat16th-centuryPaintings ;
      dcterms:subject dbc:Paintings_of_the_Louvre ; rdfs:label ?label .
    FILTER(LANG(?label)='en')
  }
  ?img imo:appearsIn ?res ; imo:fileURL ?url . }
```

Finally, IMGpedia can answer visuo-semantic queries. In our extended paper [2] we show a visuo-semantic query that requests the images of museums that are similar to any image of an European cathedral on Wikipedia. In Listing 2 we show a SPARQL query that requests the museums that are similar to images that appear on articles categorized as Roman Catholic cathedrals in Europe, using the property path `dcterms:subject/skos:broader*` to navigate sub-categories. In Figure 2 we show a sample of the retrieved results.

Listing 2: Federated visuo-semantic query requesting images of museums that are similar to images related to European cathedrals

```
SELECT  DISTINCT ?urls ?urlt WHERE {
  SERVICE <http://dbpedia.org/sparql> {
    ?sres dcterms:subject/skos:broader* dbc:Roman_Catholic_cathedrals_in_Europe}
  ?source imo:appearsIn ?sres ; imo:similar ?target ; imo:fileURL ?urls .
  ?target imo:appearsIn ?tres ; imo:fileURL ?urlt .
  SERVICE <http://dbpedia.org/sparql> {
    ?tres dcterms:subject ?sub . FILTER(CONTAINS(STR(?sub), "Museum")) } }
```

[2] http://imgpedia.dcc.uchile.cl/dumps

[3] http://imgpedia.dcc.uchile.cl/sparql

[4] Currently this cannot be done in DBpedia Commons since they do not extract the `appearsIn` relation

Fig. 1: Images of the WIKIPEDIA articles about paintings from the 16$^{\text{th}}$ century displayed at the Louvre.

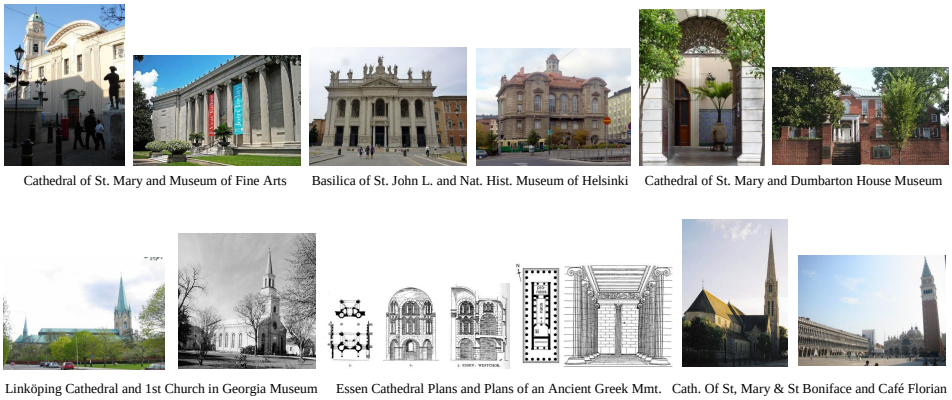

Fig. 2: Images of Roman Catholic cathedrals in Europe that have a similar image relating to a museum.

4 Future Extensions

IMGPEDIA is a novel resource. We plan to demo the first release of the dataset by showing the different kinds of queries that it is able to answer. However, we also wish to discuss plans to extend and improve the dataset and are interested to collect feedback from the ISWC community.[5] We are currently working on the following tasks towards improving the quality and usability of the data:

[5] An issue-tracker is also available at `http://github.com/scferrada/imgpedia/issues` for feedback, feature requests, suggestions, etc.

– **Provide links to Wikidata:** Categories on DBpedia are not flexible enough for some visuo-semantic queries. We are interested in creating links with Wikidata [6] to see if this would enable new/better visuo-semantic queries.
– **Compare similarity methods:** IMGpedia was built using FLANN [4] to compute the similarity relations. However, other approximated algorithms or indexing techniques can be used. Hence we are studying and comparing the different ways to provide the similarity links.
– **Include modern descriptors:** The visual descriptors used in IMGpedia are rather classic techniques. We want to explore how image similarity would behave using more modern descriptors. One such descriptor is DeCAF7 [1], which is based on the neural network classification of the image.
– **Explore more relations among images:** IMGpedia currently only provides similarity relations between images. We will explore of there are other relations that are worth including, such as `contains` if one image forms part of another, or `sameObject` if two images capture the same object but with different perspectives or scales.
– **Provide user-interfaces:** Currently IMGpedia can be accessed through a dump, a SPARQL endpoint, or through dereferencing Linked Data IRIs. We also plan to investigate interfaces that will help users interact more intuitively with the IMGpedia dataset.

Aside from extensions and improvements to IMGpedia, we are interested to find additional use-cases for the dataset. We believe that many applications can be built upon IMGpedia. We can use pre-trained neural networks to classify the dataset's images and provide the results as further context. We can also train our own network using the classes or categories of the related DBpedia/Wikidata resources to label the images and see if these provide an improved classification. More generally, we hope that IMGpedia may become a test-bed dataset for further works in the intersection of the Semantic Web and Multimedia.

Acknowledgments This work was supported by the Millennium Nucleus Center for Semantic Web Research, Grant № NC120004 and Fondecyt, Grant № 11140900.

References

1. Donahue, J., Jia, Y., Vinyals, O., Hoffman, J., Zhang, N., Tzeng, E., Darrell, T.: Decaf: A deep convolutional activation feature for generic visual recognition. In: International conference on machine learning. pp. 647–655 (2014)
2. Ferrada, S., Bustos, B., Hogan, A.: IMGpedia: a linked dataset with content-based analysis of Wikimedia images. In: The Semantic Web-ISWC 2017 (to appear)
3. Lehmann, J., Isele, R., Jakob, M., Jentzsch, A., Kontokostas, D., Mendes, P.N., Hellmann, S., Morsey, M., van Kleef, P., Auer, S., Bizer, C.: DBpedia - A Large-scale, Multilingual Knowledge Base Extracted from Wikipedia. Semantic Web Journal (2014)
4. Muja, M., Lowe, D.G.: Fast approximate nearest neighbors with automatic algorithm configuration. In: VISSAPP. pp. 331–340. INSTICC Press (2009)
5. Vaidya, G., Kontokostas, D., Knuth, M., Lehmann, J., Hellmann, S.: DBpedia Commons: Structured multimedia metadata from the Wikimedia Commons. In: The Semantic Web-ISWC 2015, pp. 281–289. Springer (2015)
6. Vrandečić, D., Krötzsch, M.: Wikidata: A free collaborative knowledgebase. Comm. ACM 57, 78–85 (2014)