INSA Lyon Universitat Passau

MASTER THESIS

Image Annotation Network

Author:
Mael Ogier

Supervisors:

Dr. David Coquil

Dr. Elöd Egyed-Zsigmond

This thesis submitted in fulfilment of the requirements for the degree of Master of Science

in the

 $\begin{array}{c} \hbox{Informatique - Information und Kommunikation (IFIK)} \\ \hbox{Double Master Program} \end{array}$

August 2015

"I have a dream for the Web [in which computers] become capable of analyzing all the data on the Web—the content, links, and transactions between people and computers. A "Semantic Web", which should make this possible, has yet to emerge, but when it does, the day-to-day mechanisms of trade, bureaucracy and our daily lives will be handled by machines talking to machines."

Tim Berners-Lee

Acknowledgements

This master thesis was done in the context of the double master degree Informatique - Information und Kommunikation (IFIK), which brings together two Master programs: a degree in computer engineering at the National Institute of Applied Sciences in Lyon (INSA Lyon) and a Master in Informatik (Schwerpunkt: Information und Kommunikationssysteme) at the University of Passau.

INSA LYON UNIVERSITAT PASSAU

Abstract

 $\begin{array}{c} \text{IFIK} \\ \text{Double Master Program} \end{array}$

Master of Science

Image Annotation Network

by Mael Ogier

The Thesis Abstract is written here (and usually kept to just this page). The page is kept centered vertically so can expand into the blank space above the title too...

Contents

Acknowledgements Abstract						
Li	st of	Figures	vi			
Li	st of	Tables v	ii			
A	bbre	iations vi	iii			
1	1.1 1.2 1.3 1.4	Motivation	1 1 2 3 3			
Ι	Sta	te of the Art	5			
2	Sen	antic Web Resources	6			
	2.1	DBpedia	6			
	2.2	GeoNames	6			
	2.3	WordNet	7			
	2.4	ImageNet	7			
3	Disambiguation					
	3.1	Section1	8			
4	Measures					
	4.1	Distance measures	6			
	4.2	Similarity measures	9			
5	Existing services					
	5.1	Web service 1				
	5.2	Web service 2				
	E 2	Apparation via state	1 1			

Contents

6	Con 6.1	Conclusion 5 6.1 Section 1 5 Contribution 5					
II	Co						
7	Proposed Methodology 7.1 Section 1						
8	Proposed Architecture 15						
	8.1	Techno	ology Choices	. 15			
		8.1.1	Java	. 15			
		8.1.2	Neo4j	. 15			
		8.1.3	NLP	16			
		8.1.4	DBpedia Spotlight	. 16			
		8.1.5	JAWS	. 16			
		8.1.6	JENA	. 16			
		8.1.7	JSoup	. 17			
	8.2	Graph	Structure	. 17			
		8.2.1	Pro-Cons	. 17			
		8.2.2	Vertexes	. 17			
		8.2.3	Edges	18			
9	Experiments 1						
	9.1		et				
	9.2 Code Explanation						
	9.3		s and Analysis				
	0.0	9.3.1	Evaluation methodology				
		9.3.2	Graph-based experiments				
		0.0.2	9.3.2.1 Direct Neighbors				
			9.3.2.2 Lists - WL				
			9.3.2.3 Lists - SL				
		9.3.3	Plain-text experiments	. 20			
			9.3.3.1 WikiLinks				
			9.3.3.2 WikiContent				
10	Con	clusio		21			
10		Section					
A	Арр	endix	Title Here	22			
Bibliography							

List of Figures

List of Tables

Abbreviations

LAH List Abbreviations Here

Introduction

1.1 Background

Image is a popular medium nowadays: it is easy to capture, can be really light on your computer and speaks to everyone without distinction of language.

In the all days life, people share their pictures on social networks in less than a blink of eye. In average, 70M of pictures are posted on Instagram each day and the users hit the "Like" button 2.5B times¹. Other services like Picasa or Flickr exists but aren't as used as Instagram which is the favorite in the eyes of the teen public.

Companies also produce a lot of media data. Industry companies need their products' pictures, marketing and advertising studios use a lot of images in order to create new stuff for their client, ... But the most consumer of media data are obviously mass media themselves: Newspapers, TV shows, news broadcasts are dealing with pictures at every moment of their day.

This huge production and consumption of images implies the need of an efficient way to store and search for the relevant one when the time comes. The best illustration to this need is to think of the nice but long moments one had with its relatives searching for the good picture of the new-born nephew in the family pictures album.

Since an image itself doesn't have a natural plain-text representation the best way to

¹Stats from: https://instagram.com/press/

describe it is to add meta-data (data about the data) such as its date of creation, its dimensions or, and this is what this thesis is about, some tags.

There are a lot of ways if one wants to annotate pictures. We can do it manually, using our own words (like "Dad", "Home" ...), we can also analyze the raw picture, its pixel representation and compare some metrics (like the color histogram) to sample images in order to detect known concepts. Moreover, if the image already possesses annotations, we can enrich it semantically.

This field is so wide that it is impossible to speak about all the possibilities and technologies. In this study, we will focus on the last point and investigate the automation of the semantic enrichment. We will study the resources at our disposal and propose a solution keeping in mind the facts cited previously.

In the following section, we will present and discuss an application scenario to illustrate the motivation behind this thesis.

1.2 Motivation

NewsTV is a famous TV news channel which runs 24/7 and only speaks about the current news. It has lot of reporters worldwide, covering the important local news and sending their production to the main site in Paris, France.

The employees often need to consult older coverages in order to explain the context of the news, to make the necrology of a famous actor who recently died or to re-use common shots. Therefor, they need to query the central multimedia database management system using keywords they are familiar with like "Elections, France, 2007, José Bové". But sometimes, their research aren't so specific and they are looking for more generic pictures, let say "Land, Tree, Animal".

The first kind of keywords had been tagged by the former reporter who produced the coverage but he logically didn't think to add generic terms. NewsTV needs something to do it automatically when a picture, or any media, is first added to its system which a couple of initial tags.

Details about which kind of technology can be used to achieve this automatic tagging will come in the following sections. To summarize, the goal of this thesis is to propose a running prototype and evaluate different methods of tagging. The questions that we will try to answer during this study are described in the following section.

1.3 Thesis Objectives

Sed ullamcorper quam eu nisl interdum at interdum enim egestas. Aliquam placerat justo sed lectus lobortis ut porta nisl porttitor. Vestibulum mi dolor, lacinia molestie gravida at, tempus vitae ligula. Donec eget quam sapien, in viverra eros. Donec pellentesque justo a massa fringilla non vestibulum metus vestibulum. Vestibulum in orci quis felis tempor lacinia. Vivamus ornare ultrices facilisis. Ut hendrerit volutpat vulputate. Morbi condimentum venenatis augue, id porta ipsum vulputate in. Curabitur luctus tempus justo. Vestibulum risus lectus, adipiscing nec condimentum quis, condimentum nec nisl. Aliquam dictum sagittis velit sed iaculis. Morbi tristique augue sit amet nulla pulvinar id facilisis ligula mollis. Nam elit libero, tincidunt ut aliquam at, molestie in quam. Aenean rhoncus vehicula hendrerit.

1.4 Thesis Outline

The remainder of this thesis will be organized as follows:

- Chapter 2 Semantic Web Resources: presents different semantic web resources, their structures and how to browse them and how are they used in the literature.
- Chapter 3 Disambiguation: reviews the literature and assesses the most relevant ways to disambiguate a list of keywords which may be organize into sentences or not.

- Chapter 4 Measures: provides a solid background on semantic similarity and distance measures. We explore different metrics illustrating their pro/cons with examples.
- **Chapter 5 Existing Services:** describes existing image annotation services as well as their use-cases.
- Chapter 6 State of the Art Conclusion: summarizes the findings of the previous state of the art and opens the way to the presented contribution.
- Chapter 7 Proposed Methodology: presents the chosen methodology as well as some organizational points.
- Chapter 8 Proposed Architecture: details the technological choices by comparing them to their competitors and the chosen DBMS schema. Illustration figures will be presented.
- Chapter 9 Proposed Architecture: presents the chosen dataset, details some of the main algorithms and reviews the tests' results with the use of different evaluation methods.
- Chapter 10 Contribution Conclusion: summarize the findings of the presented research problem.

Part I

State of the Art

Semantic Web Resources

2.1 DBpedia

DBpedia is a project originally launched by two German universities (Berlin and Leipzig) and backed by an important community. It explore and extract information from Wikipedia and then semantically format it. Each encyclopedic document has his own page filled with data store in RDF (Resource Description Framework) triplets.

DBpedia's hierarchical organization is based on classes and categories. Classes have super-classes and sub-classes, the root element being "Thing".

As well as any RDF-structured dataset, DBpedia can be requesting with SPARQL (SPARQL Protocol and RDF Query Language) queries. We can also use online applications according to our needs. For example, DBpedia Spotlight detects DBpedia entities/classes in a text which can be really useful if we possess a description.

2.2 GeoNames

GeoNames is a geographical database which contains more than 6,5 million places. It's a collaborative tool: users can add data or edit the existing ones. GeoNames' data are

link to other RDF resources (mainly DBpedia).

Semantically speaking, each GeoNames' resource is link to an URI. This URI lead the user to an HTML page or to a RDF description. We can then write SPARQL queries using specific namespaces and request GeoNames. This resource is a key in our process, it will give us our first information about the picture given its GPS coordinates.

2.3 WordNet

WordNet is a lexical database created by Priceton University. Its structure is based on the "synset" (synonym set), a set of words with the same meaning. If DBpedia's architecture uses classes, Wordnet use hyponyms and hyperonyms. The root element is "entity".

This resource is useful if we are searching for entities (tsetse-fly, jaguar, ...) but it also contains more general concepts (vehicle, animal, ...). It also have interesting links, like meronyms ("HAS-PART" relation), holonyms (the opposite), synonyms or antonyms.

2.4 ImageNet

ImageNet is an image dataset based on the WordNet hierarchy. For each synset, ImageNet provide a set of images depicting it.

Disambiguation

3.1 Section1

Measures

4.1 Distance measures

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aliquam ultricies lacinia euismod. Nam tempus risus in dolor rhoncus in interdum enim tincidunt. Donec vel nunc neque. In condimentum ullamcorper quam non consequat. Fusce sagittis tempor feugiat. Fusce magna erat, molestie eu convallis ut, tempus sed arcu. Quisque molestie, ante a tincidunt ullamcorper, sapien enim dignissim lacus, in semper nibh erat lobortis purus. Integer dapibus ligula ac risus convallis pellentesque.

4.2 Similarity measures

Existing services

5.1 Web service 1

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aliquam ultricies lacinia euismod. Nam tempus risus in dolor rhoncus in interdum enim tincidunt. Donec vel nunc neque. In condimentum ullamcorper quam non consequat. Fusce sagittis tempor feugiat. Fusce magna erat, molestie eu convallis ut, tempus sed arcu. Quisque molestie, ante a tincidunt ullamcorper, sapien enim dignissim lacus, in semper nibh erat lobortis purus. Integer dapibus ligula ac risus convallis pellentesque.

5.2 Web service 2

5.3 Annotation via stats

Conclusion

6.1 Section 1

Part II

Contribution

Proposed Methodology

7.1 Section 1

Proposed Architecture

8.1 Technology Choices

8.1.1 Java

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aliquam ultricies lacinia euismod. Nam tempus risus in dolor rhoncus in interdum enim tincidunt. Donec vel nunc neque. In condimentum ullamcorper quam non consequat. Fusce sagittis tempor feugiat. Fusce magna erat, molestie eu convallis ut, tempus sed arcu. Quisque molestie, ante a tincidunt ullamcorper, sapien enim dignissim lacus, in semper nibh erat lobortis purus. Integer dapibus ligula ac risus convallis pellentesque.

8.1.2 Neo4j

8.1.3 NLP

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aliquam ultricies lacinia euismod. Nam tempus risus in dolor rhoncus in interdum enim tincidunt. Donec vel nunc neque. In condimentum ullamcorper quam non consequat. Fusce sagittis tempor feugiat. Fusce magna erat, molestie eu convallis ut, tempus sed arcu. Quisque molestie, ante a tincidunt ullamcorper, sapien enim dignissim lacus, in semper nibh erat lobortis purus. Integer dapibus ligula ac risus convallis pellentesque.

8.1.4 DBpedia Spotlight

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aliquam ultricies lacinia euismod. Nam tempus risus in dolor rhoncus in interdum enim tincidunt. Donec vel nunc neque. In condimentum ullamcorper quam non consequat. Fusce sagittis tempor feugiat. Fusce magna erat, molestie eu convallis ut, tempus sed arcu. Quisque molestie, ante a tincidunt ullamcorper, sapien enim dignissim lacus, in semper nibh erat lobortis purus. Integer dapibus ligula ac risus convallis pellentesque.

8.1.5 JAWS

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aliquam ultricies lacinia euismod. Nam tempus risus in dolor rhoncus in interdum enim tincidunt. Donec vel nunc neque. In condimentum ullamcorper quam non consequat. Fusce sagittis tempor feugiat. Fusce magna erat, molestie eu convallis ut, tempus sed arcu. Quisque molestie, ante a tincidunt ullamcorper, sapien enim dignissim lacus, in semper nibh erat lobortis purus. Integer dapibus ligula ac risus convallis pellentesque.

8.1.6 JENA

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aliquam ultricies lacinia euismod. Nam tempus risus in dolor rhoncus in interdum enim tincidunt. Donec vel nunc neque. In condimentum ullamcorper quam non consequat. Fusce sagittis tempor feugiat. Fusce magna erat, molestie eu convallis ut, tempus sed arcu. Quisque molestie,

ante a tincidunt ullamcorper, sapien enim dignissim lacus, in semper nibh erat lobortis purus. Integer dapibus ligula ac risus convallis pellentesque.

8.1.7 JSoup

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aliquam ultricies lacinia euismod. Nam tempus risus in dolor rhoncus in interdum enim tincidunt. Donec vel nunc neque. In condimentum ullamcorper quam non consequat. Fusce sagittis tempor feugiat. Fusce magna erat, molestie eu convallis ut, tempus sed arcu. Quisque molestie, ante a tincidunt ullamcorper, sapien enim dignissim lacus, in semper nibh erat lobortis purus. Integer dapibus ligula ac risus convallis pellentesque.

8.2 Graph Structure

8.2.1 Pro-Cons

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aliquam ultricies lacinia euismod. Nam tempus risus in dolor rhoncus in interdum enim tincidunt. Donec vel nunc neque. In condimentum ullamcorper quam non consequat. Fusce sagittis tempor feugiat. Fusce magna erat, molestie eu convallis ut, tempus sed arcu. Quisque molestie, ante a tincidunt ullamcorper, sapien enim dignissim lacus, in semper nibh erat lobortis purus. Integer dapibus ligula ac risus convallis pellentesque.

8.2.2 Vertexes

8.2.3 Edges

Experiments

9.1 Dataset

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aliquam ultricies lacinia euismod. Nam tempus risus in dolor rhoncus in interdum enim tincidunt. Donec vel nunc neque. In condimentum ullamcorper quam non consequat. Fusce sagittis tempor feugiat. Fusce magna erat, molestie eu convallis ut, tempus sed arcu. Quisque molestie, ante a tincidunt ullamcorper, sapien enim dignissim lacus, in semper nibh erat lobortis purus. Integer dapibus ligula ac risus convallis pellentesque.

9.2 Code Explanation

9.3 Results and Analysis

9.3.1 Evaluation methodology

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aliquam ultricies lacinia euismod. Nam tempus risus in dolor rhoncus in interdum enim tincidunt. Donec vel nunc neque. In condimentum ullamcorper quam non consequat. Fusce sagittis tempor feugiat. Fusce magna erat, molestie eu convallis ut, tempus sed arcu. Quisque molestie, ante a tincidunt ullamcorper, sapien enim dignissim lacus, in semper nibh erat lobortis purus. Integer dapibus ligula ac risus convallis pellentesque.

9.3.2 Graph-based experiments

- 9.3.2.1 Direct Neighbors
- 9.3.2.2 Lists WL
- 9.3.2.3 Lists SL
- 9.3.3 Plain-text experiments
- 9.3.3.1 WikiLinks
- 9.3.3.2 WikiContent

Conclusion

10.1 Section 1

Appendix A

Appendix Title Here

Write your Appendix content here.

Bibliography