Usage based Tag enhancement of Images

PROJECT TAKEN AT

BIG DATA EXPERIENCE LAB

ADOBE SYSTEMS

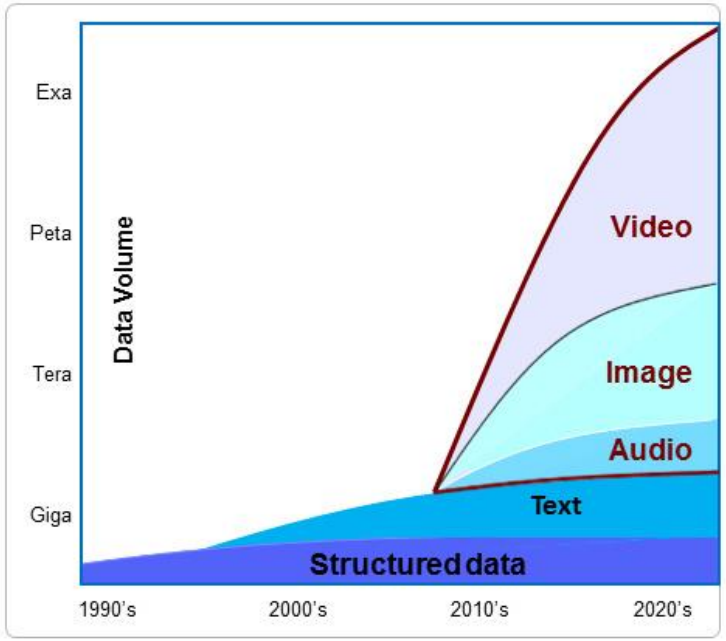
by

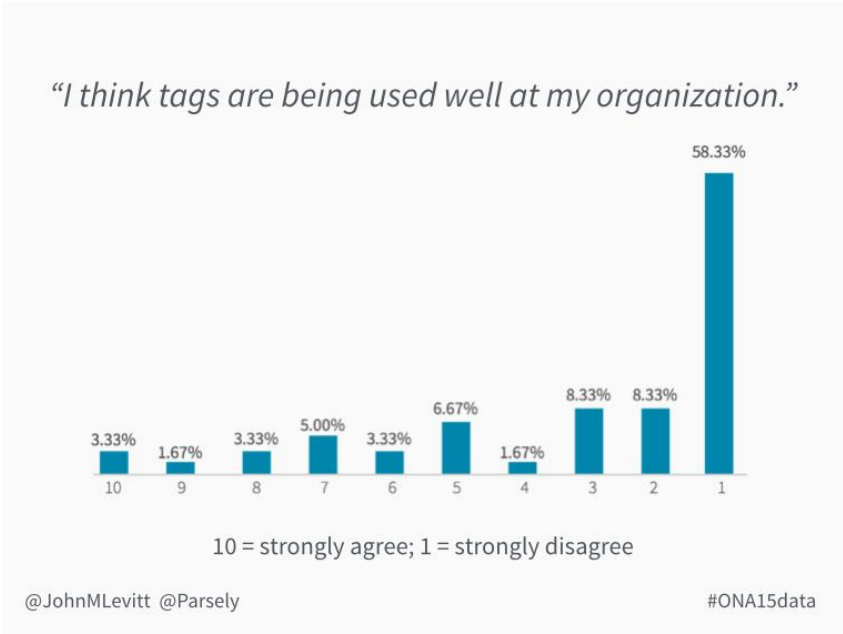
- **Saurabh Verma**
- Enrollment No. 13114057

Acknowledgements

- I would like to take this opportunity to sincerely thank my team members, Balaji Vasan(BEL Adobe), Noman Sheikh(IIT Delhi) and Roshan Kumar(IIT Kanpur) for their collaboration in the project.
- I am also thankful to Lab Members for their invaluable ideas and suggestions.
- I cannot imagine a successful internship without Akash Gupta and Sai Varun who were always available.
- I am highly indebted to Prof. Niloy Ganguly(IIT Kharagpur) for making the project reach these heights.
- This internship would not have been possible without Smitha, Anand and Ameena, internship coordinators.

What is tagging? Why do we need it?





https://avindh2014.wordpress.com/background/

http://blog.parsely.com/post/2565/ona15data-a-year-in-data-from-parse-ly/parse-ly/

What if there was a system that could bridge the gap between what images contain and how they are perceived?

What if there was a system that can give a set of tags that maximizes their significance in retrieval?

“Problem statement: Given an image and set of associated content, enhance the set of tags based on it's usage”

- Develop a semantic understanding of the image content
- Enhance a given set of tags using background knowledge
- Improve retrieval and recommendation by using the given tags

Running Example

http://www.npr.org/sections/parallels/2016/03/22/471401729/in-vulnerable-europe-a-third-major-terrorist-attack-in-a-year

Tags from Clarifai API

clarifai ABOUT PRICING DEVELOPER BLOG CONTACT US SIGN UP NOW

Predicted Tags

bridge people travel city architecture building vehicle
man tourism landscape outdoors tourist road adult
sky group transportation system

https://www.clarifai.com/#demo

Tags from Microsoft Vision API

Features:

Feature Name	Value
Description	{ "type": 0, "captions": [{ "text": "a man standing on top of a bridge", "confidence": 0.13269974921728017 }] }
Tags	[{ "name": "sky", "confidence": 0.9977843165397644 }, { "name": "outdoor", "confidence": 0.9974787831306458 }, { "name": "building", "confidence": 0.996910035610199 }, { "name": "arch", "confidence": 0.9431114792823792 }, { "name": "person", "confidence": 0.9322513937950134 }, { "name": "bridge", "confidence": 0.9236949682235718 }]
Image Format	jpg
Image Dimensions	900 x 599
Clip Art Type	0 Non-clipart
Line Drawing Type	0 Non-LineDrawing
Black & White Image	False

https://www.microsoft.com/cognitive-services/en-us/computer-vision-api

Tags derived by our pipeline

Prior Art (Academic)

Text mining for automatic image tagging

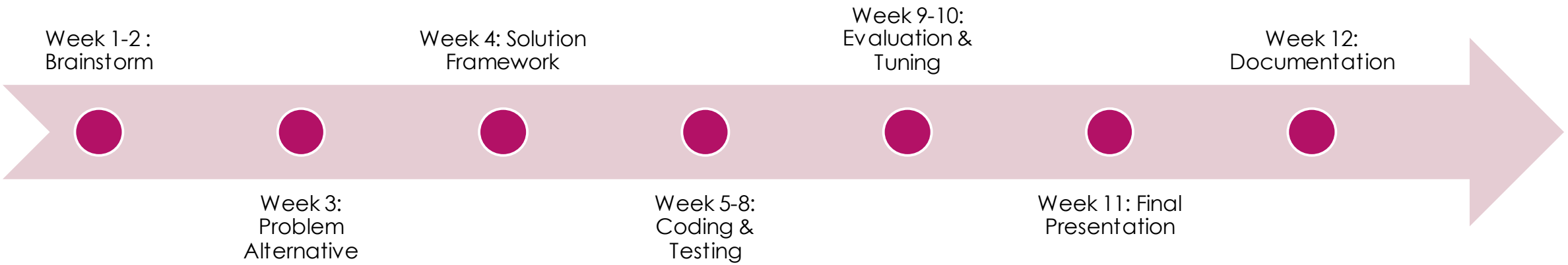
- Leong et al., *Proceedings of the 23rd International Conference on Computational Linguistics: Posters*. Association for Computational Linguistics, 2010.

A review on automatic image annotation techniques

- Zhang, Dengsheng, Md Monirul Islam, and Guojun Lu. *Pattern Recognition* 45.1 (2012): 346-362.

Prior Art (Industry)

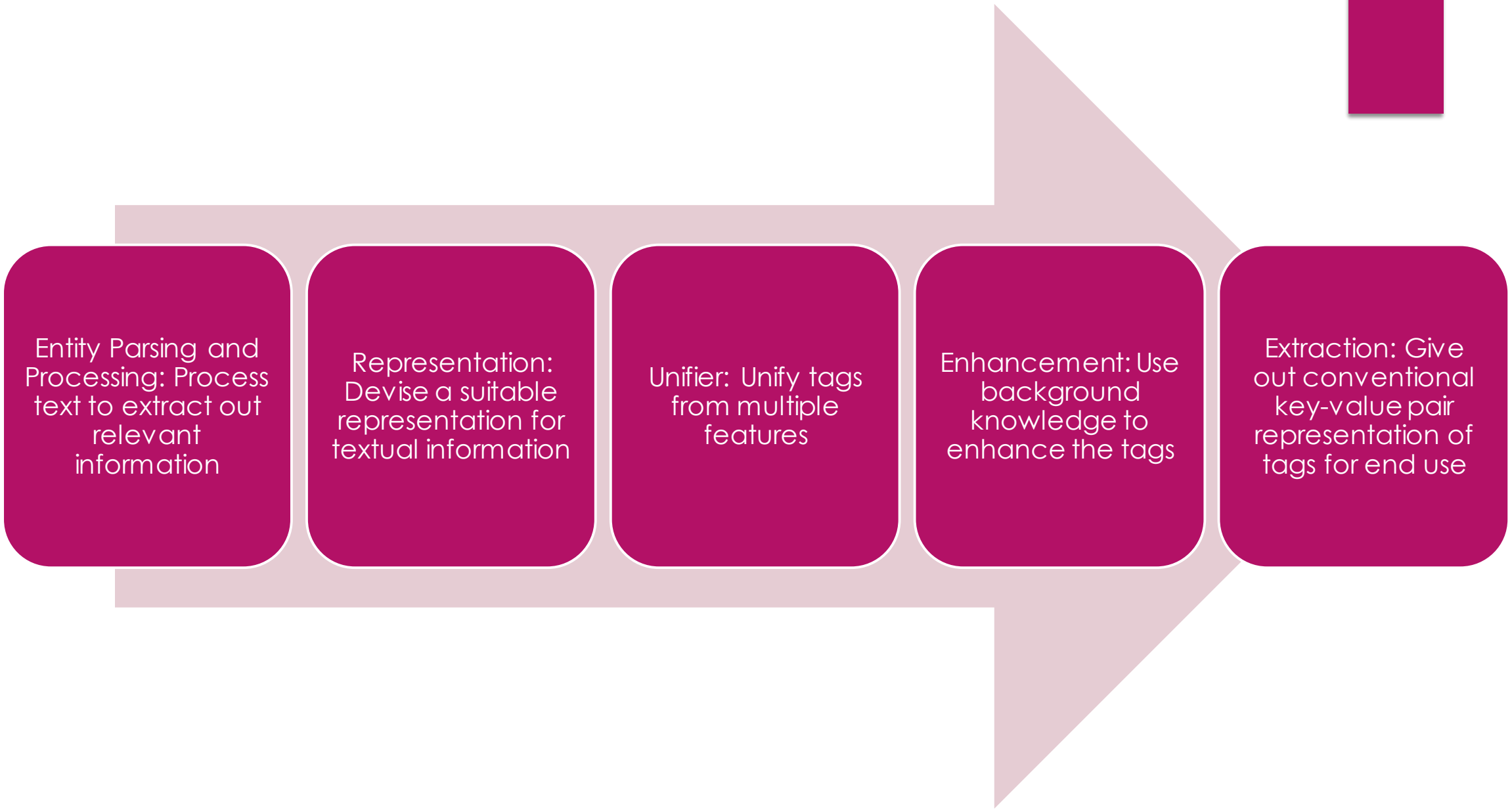
Workplan


Week 1-2 : Brainstorm
Week 3: Problem Alternative
Week 4: Solution Framework
Week 5-8: Coding & Testing
Week 9-10: Evaluation & Tuning
Week 11: Final Presentation
Week 12: Documentation

Objectives

- Combine visual and textual features for tagging of images
- Enhance the given set of tags using background knowledge from large knowledge bases
- Use the final tags to improve retrieval and recommendation engines
- Develop a cross-platform and re-entrant pipeline

Solution Framework


Entity Parsing and Processing: Process text to extract out relevant information
Representation: Devise a suitable representation for textual information
Unifier: Unify tags from multiple features
Enhancement: Use background knowledge to enhance the tags
Extraction: Give out conventional key-value pair representation of tags for end use

Evaluation & Results

Metrics

Experimental Setup

Baselines

Imagga

- https://imagga.com/auto-tagging-demo

Microsoft Vision

- https://www.microsoft.com/cognitive-services/en-us/computer-vision-api

Text mining for automatic image tagging

- Leong et al. "Text mining for automatic image tagging." *Proceedings of the 23rd International Conference on Computational Linguistics: Posters*. ACL, 2010.

Gold-standard Human Annotations and Visual autotagging engine

- http://lit.csci.unt.edu/index.php/Downloads

Results

Future Work

- Extend this work for multimedia content using the same principal pipeline
- Capture multiple word senses in text
- Parallelize and optimize the application for low latency solution
- Offer solution for rarely used images
- Measure the performance of each component and tune them

Intellectual Property

Paper

- We are pushing for a publication in WWW 2017 where submission is due on October 24 with the title "Usage based Tag enhancement of images"

Invention Disclosure

- We have drafted an invention disclosure which is pending approval.

Fun & Bragging

https://blogs.adobe.com/adobelife/2016/08/23/dear-future-interns/