Information Retrieval

BITS Pilani
Pilani Campus

Abhishek
April 2020

innovate achieve lead

BITS Pilani

Pilani Campus

CS F469, Information Retrieval

Lecture topics: Multimedia IR

What is Multimedia?

- Basically our digital life
- Audio
  - Music
  - Speach
  - Sound
- Visuals
  - Images
  - Sketches
- Misc
  - Videos
  - 3D objects
  - …
- Any combination
  - PDF, web pages

Multimedia IR: Use cases

- Images:
  - Find look-alike pictures. (Google Reverse Image search)
    - Medical Images
  - recognize landmarks. (Google Lens)
  - people wearing white suit

Multimedia IR: Use cases

- Audio:
  - Finding similar sounds
    - Music
    - Animal sounds
    - Copyrights
  - Retrieve relevant speech or podcasts or music

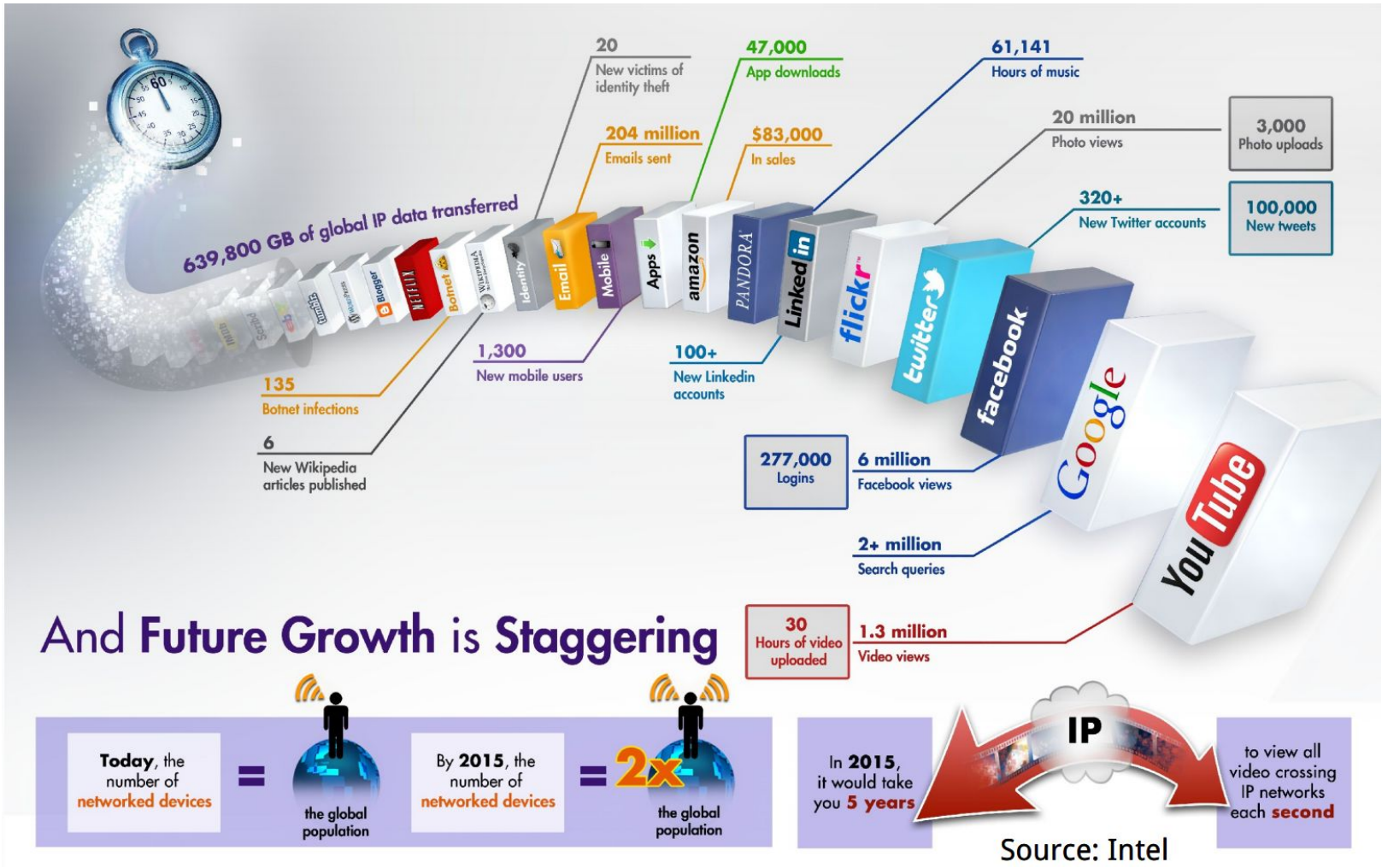
Shazam [http://www.shazam.com/] discover what song is playing
AudioID: https://www.mufin.com/products/
http://themefinder.org/

Multimedia IR: Use cases

- Videos:
  - Finding similar videos
    - Copyrights
  - Retrieve relevant news telecasts
  - Find specific actions: Kathak dance
  - Lecture

https://berify.com/

Volume of Multimedia Data



Basic IR Working Intuition

Piggyback Retrieval

- Convert everything to Text
- Use text based retrieval

Use metadata or loose annotations (automated or manual)

- **Image caption** (Text surrounding images on web page)
- Video to Text
- Music to Text
- Metadata of videos, music (title)

Eg: http://max-image-caption-generator-web-app.mybluemix.net/

Information Loss

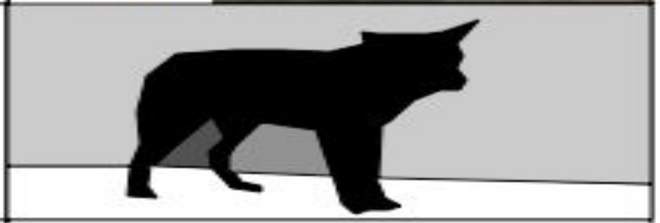
- Annotation:

"Wasp on a yellow Flower"

Will miss the query:

" insect feeding"

Semantic Gap

Semantics object relationships and more	Wolf **on** Road **with** Snow **on** **Roadside in Yosemite National Park, California on 24/1/2004 at 23:19:11GMT**
Object Labels symbolic names of objects	
Objects prototypical combinations of descriptors	
Descriptors feature-vectors	Segmented blobs, Salient regions, Pixel-level histograms, Fourier descriptors, etc...
Raw Media images	

Semantic Gap

Raw Media

This is what we have to work with

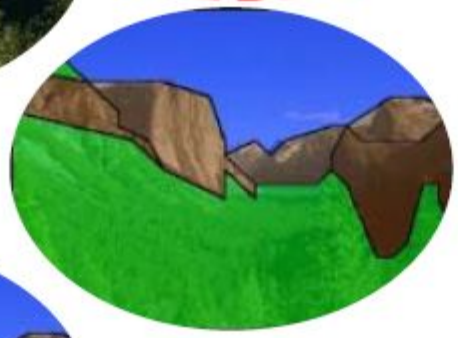


Image-level descriptors

Content descriptors

This is what we want

Semantic content

Photo of Yosemite valley showing El Capitan and Glacier Point with the Half Dome in the distance

Example - Sub-image matching

- Given a query image, find the parent image in the database with which it matches, either as a whole or as a part
- Give location information showing where in the parent the query is positioned
- The images may be very high resolution
- The query and target may be at different resolutions.

Example - Sub-image matching

Example - Sub-image matching

Best matching image with sub-image identified

Thank You!