Contribution to Knowledge Search in Video Content

Author: Francisco Javier Rodríguez González

Thesis Advisors: Cristina López Bravo, Enrique Costa Montenegro
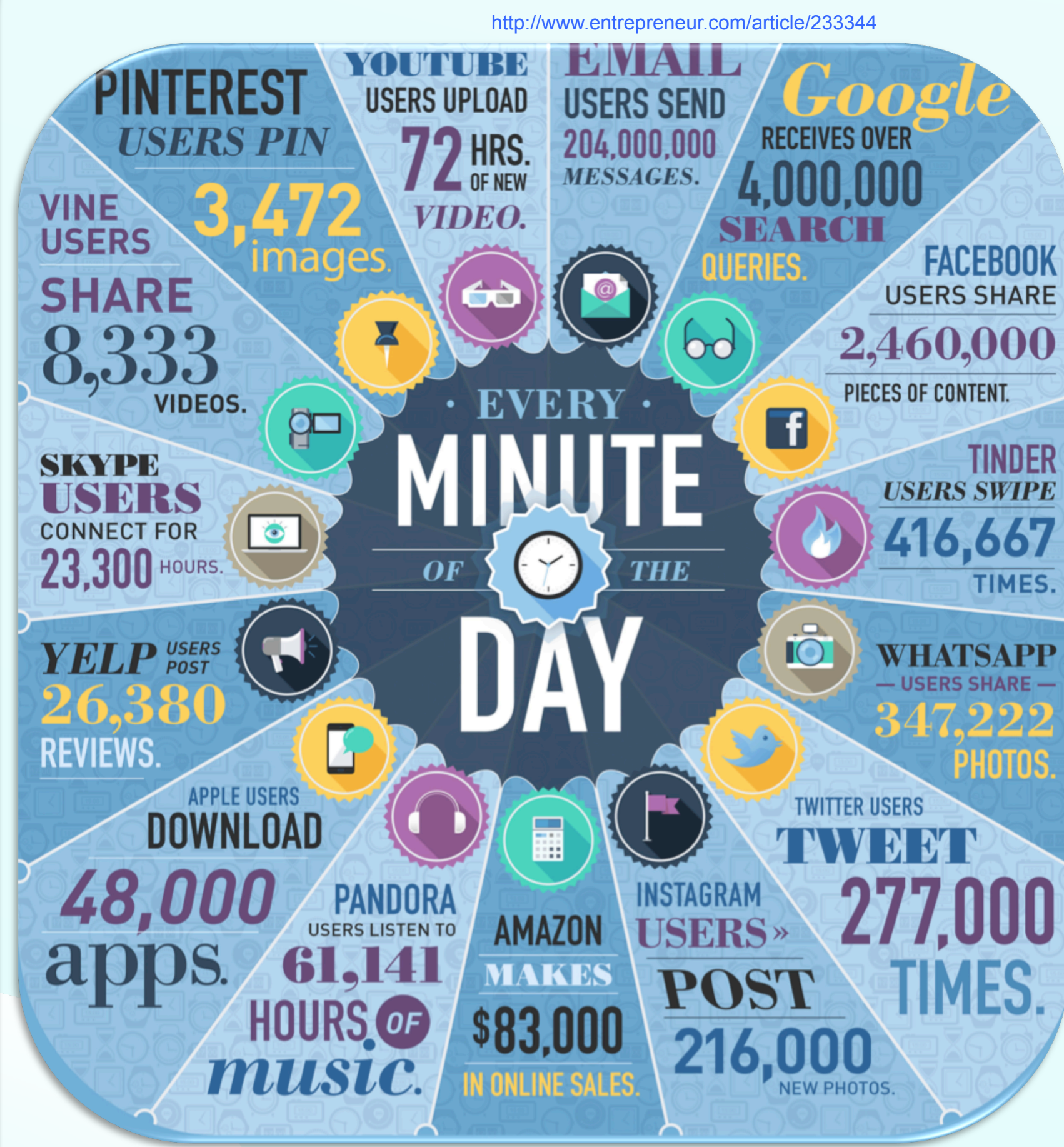
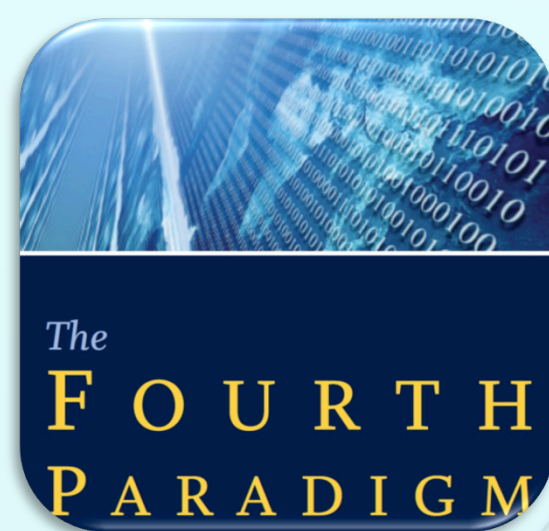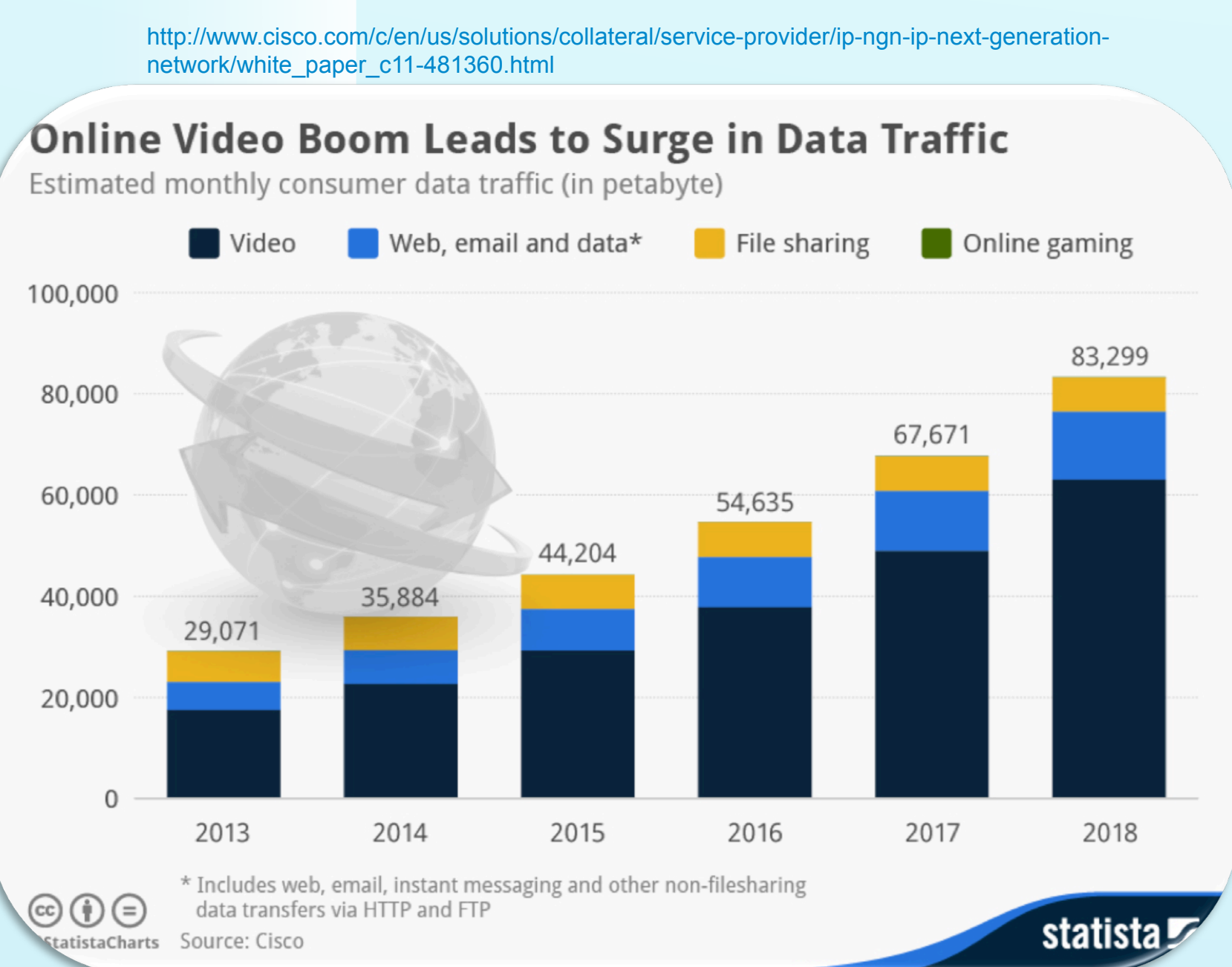**PhD Program:** Information and Communications Technology of the University of Vigo

Motivation

- We live in the **Information Era**

http://www.entrepreneur.com/article/233344

- **Big Data (BD)** market: $125 Billion

- Data-driven Science

  LHC 50-100 PB/year

http://research.microsoft.com/en-us/collaboration/fourthparadigm/

- Video traffic is growing

http://www.cisco.com/c/en/us/solutions/collateral/service-provider/ip-ngn-ip-next-generation-network/white_paper_c11-481360.html

Online Video Boom Leads to Surge in Data Traffic
Estimated monthly consumer data traffic (in petabyte)

* Includes web, email, instant messaging and other non-filesharing data transfers via HTTP and FTP
Source: Cisco

- Studying using Internet resources, internal training, is common

- **Video** is usually **indexed upon** the metadata, **not** on the content itself

Thesis Objectives

Define the architecture of a software system that is able to **search relevant content inside video**, correlating all its information sources, voice, image, scenery, metadata, author, context

Requirements

- **Web Scalable** and Highly Available
- State of the art **Open Source projects**
- "Correlated" **layered architecture**
  - ◇ OCR and voice transcripts
  - ◇ Semantic dimension
  - ◇ Image and scene recognition
- Cloud model, real time
- **Right click to Search** point in time

Example: Find AWS growth in Keynote

Figure Source: AWS re:Invent 2014 | Day 1 Keynote with Andy Jassy
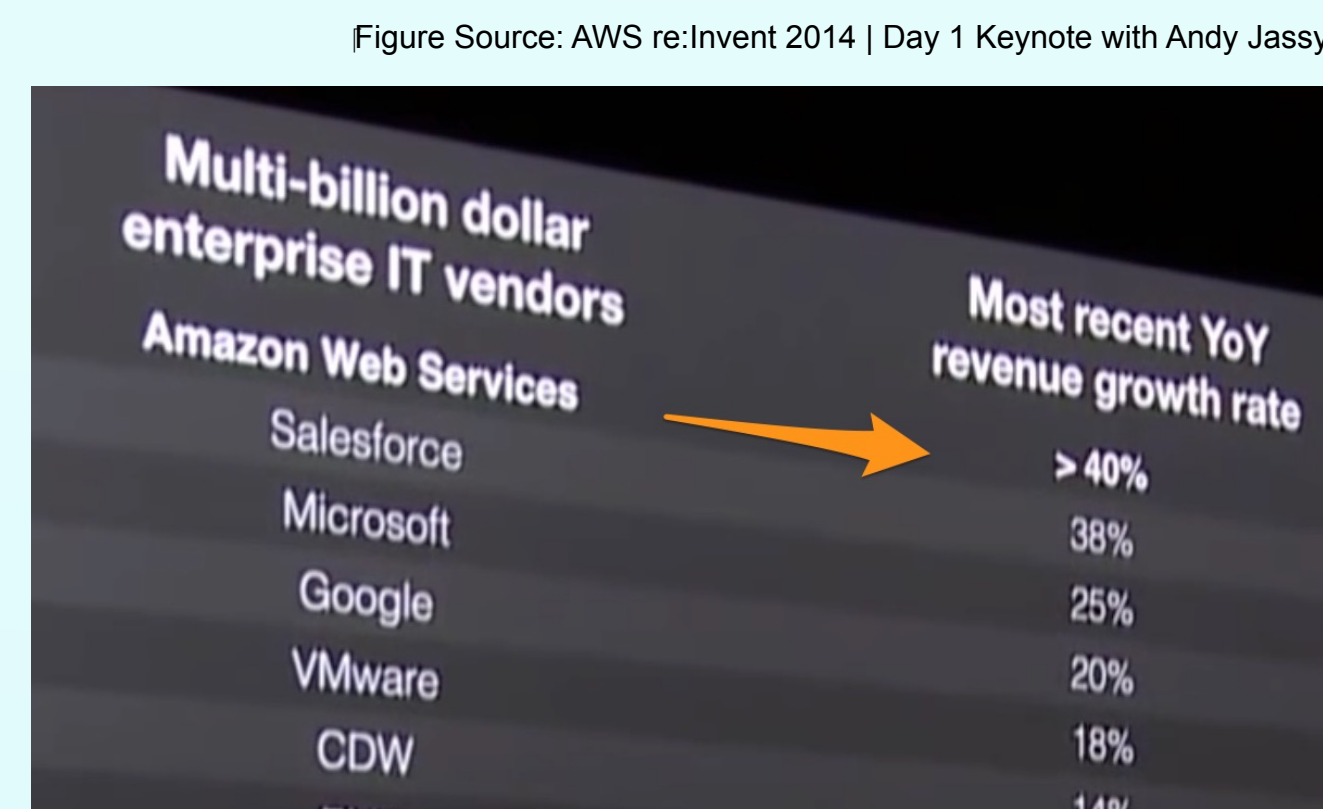
Other Applications:
   Keynotes, Transcripts, Accelerated viewing
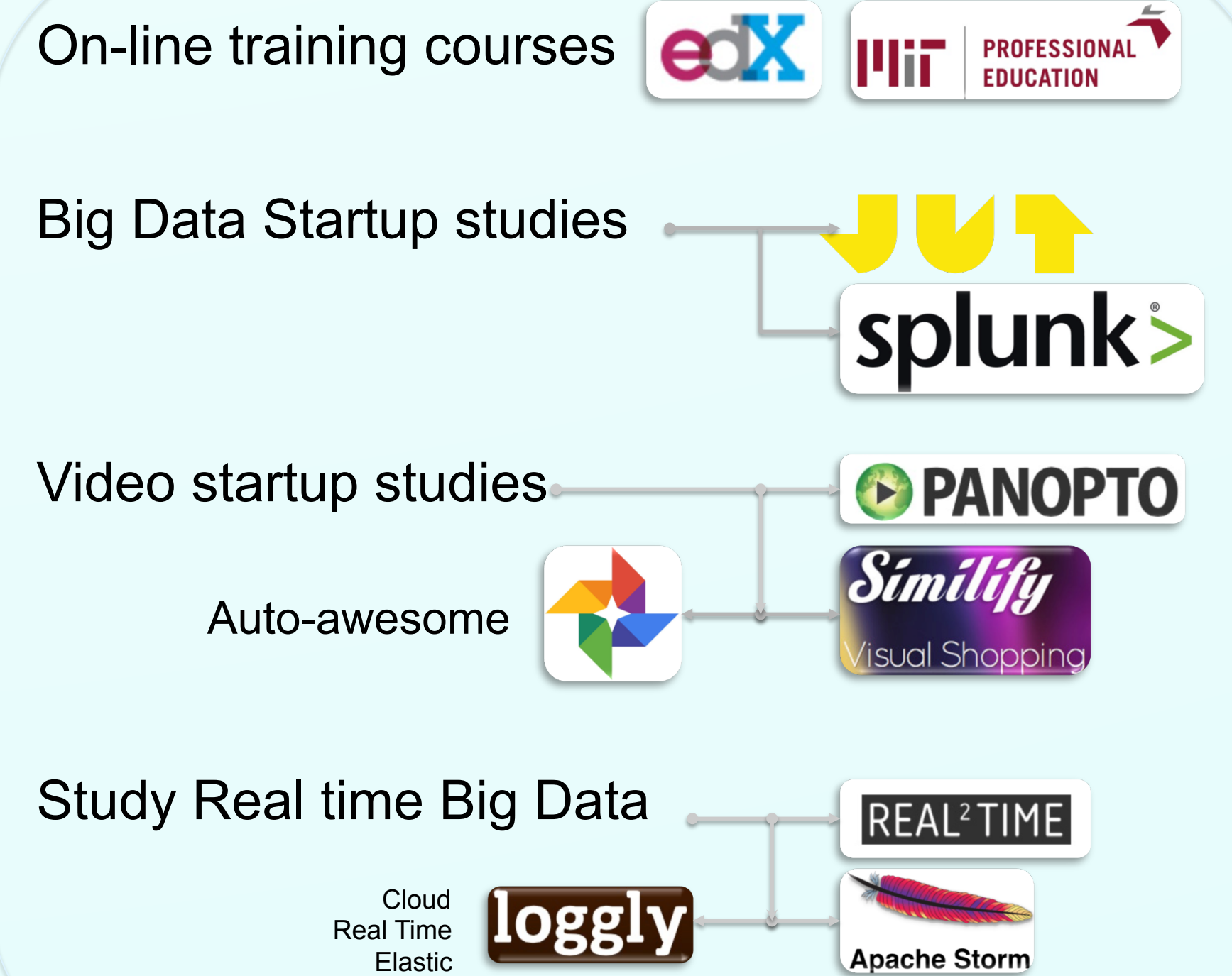
Research Plan

1. Study BD State of the Art (SOTA) ☑
   i. Academia research perspective ☑
   ii. Start-ups perspective ☑
2. Study Video Search SOTA ☑
   i. Machine learning ☑
   ii. Start-ups in the realm ☑
3. Study Apache Storm, Spark ☑
4. Build a Storm Sandbox ☑
   I. Cloud ingest optimization(RT) ☑
   II. Scalability, HA experiments ☑
5. System Architecture ☑
   I. Definition ☑
   II. Open Source modules to integrate ☑
   III. Develop Scene and semantic modules
   IV. System output and search interface
   V. Quality, Information Index definition ☑
   VI. Develop functional prototype

Results and Discussions

On-line training courses

Big Data Startup studies

Video startup studies

Auto-awesome

Study Real time Big Data

Cloud
Real Time
Elastic

Storm Ingest Optimization and Cloud model

- Connectors: Twitter, Instagram, Youtube, Alchemy
- Build parallelism: study Twitter Rate limits
- Latency optimization: 33 % faster

Definition of the System Architecture

1. Real Time scalability: Storm, Spark, Cassandra, ES
2. Best of breed Open Source + scene innovation
3. Cloud model for data ingest, output. REST API

Next Year Planning

- Work on Paper: Storm ingest optimization
- Study Machine Learning: **Caffe**, Scikit-learn
- Implement Data Layer
- Work on a Connector to Youtube per frame
- Implement Caffe and OCR modules
- Define Correlation (OCR, Image, t)

References

[1] Billions Views/day. 300 h upload/min
http://www.youtube.com/yt/press/statistics.html

[2] 6 Predictions for the $125 Billion BD market
http://www.forbes.com/sites/gilpress/2014/12/11/6-predictions-for-the-125-billion-big-data-analytics-market-in-2015/2/

[3] Apache Storm, Spark, Real2Time

[4] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar, L.Fei-Fei, "Large-scale Video Classification with Convolutional Neural Networks", Conference on Computer Vision and Pattern Recognition (CVPR), 2014 IEEE
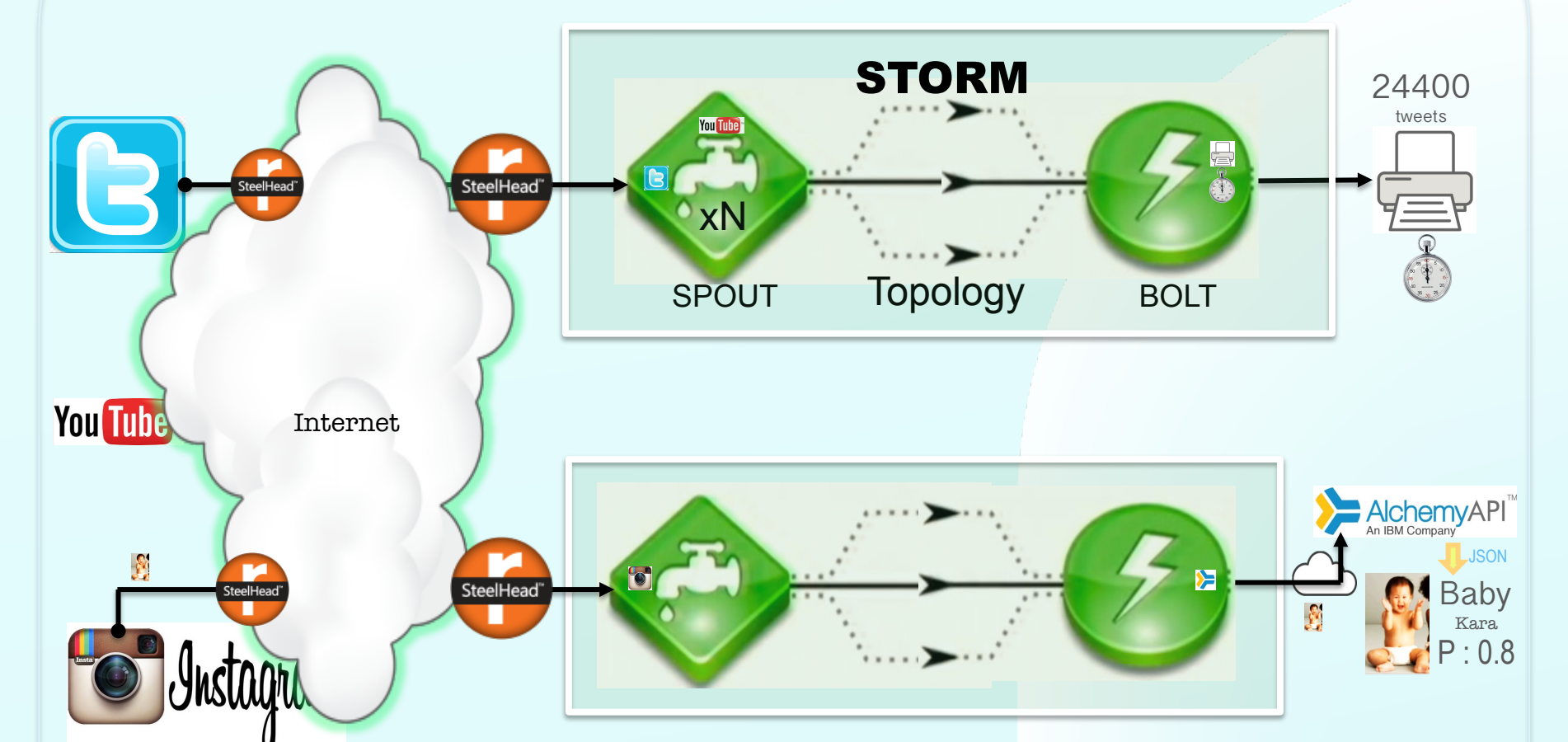
[5] D. Kinoshenko, S. Mashtalir, A. Stephan, V Vinarski, "Neural Network Segmentation of Video via time series analyses", Int. Journal, Information Theories and Applications", V. 18, # 3, 2011

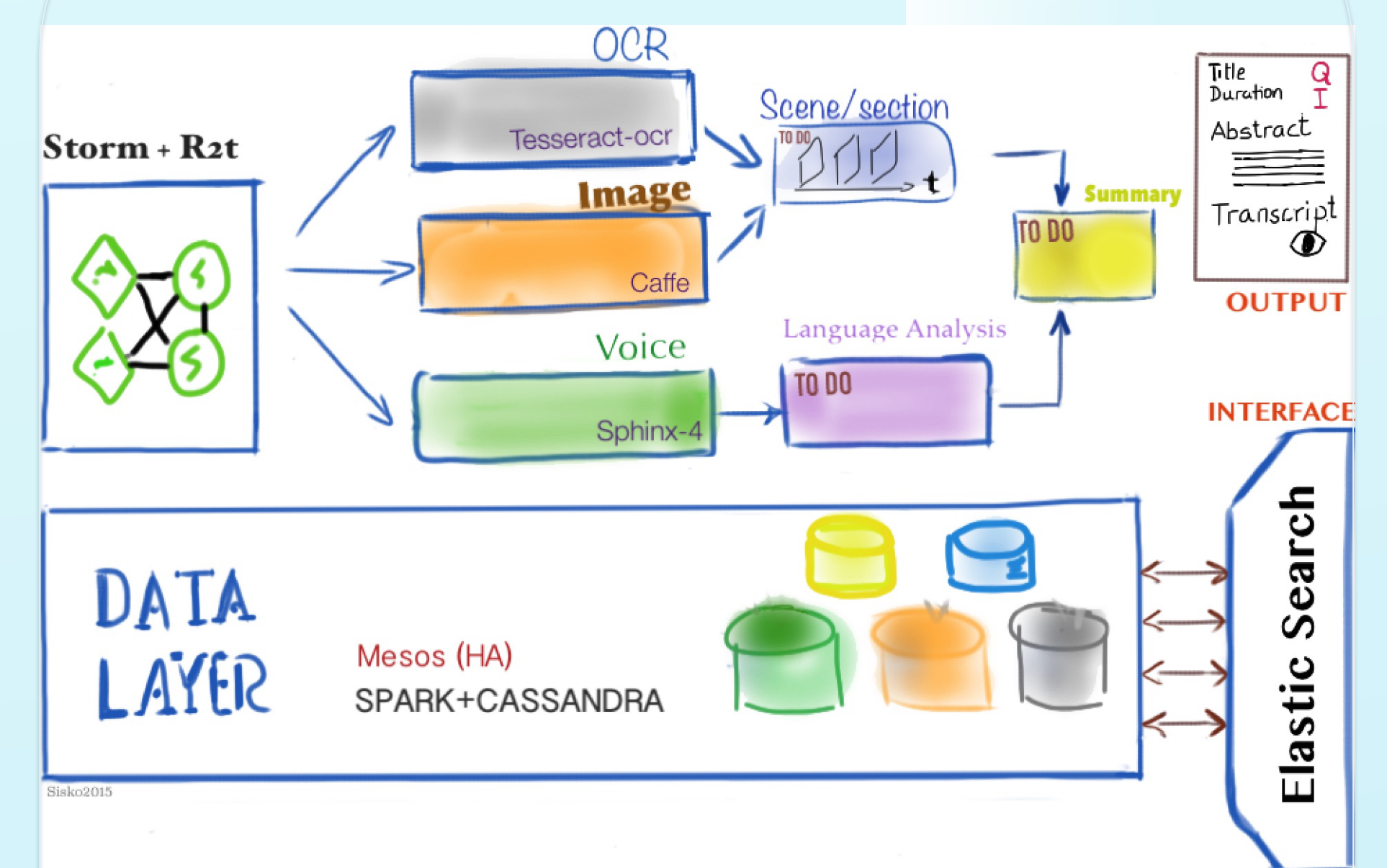[6] A. Karpathy, Li Fei-Fei, "Deep Visual-Semantic Alignments for Generating Image Descriptions", *arXiv preprint arXiv:1412.2306* (2014)