computational social media

lecture 6: watching

daniel gatica-perez

RESEARCH INSTITUTE

announcements

indicative feedback

project reminder

assignment #3 will be discussed today

reading #7 will be presented today:

M. Redi, N. O Hare, R. Schifanella, M. Trevisiol, A. Jaimes
6 Seconds of Sound and Vision: Creativity in Micro-Videos
CVPR 2014

indicative feedback: statistics

thank you for participating

11 responses (19 students):

Overall, I think this course is good.

Possible answers :	Nb. answers :	In % :
Strongly agree	6	55
Agree	4	36
Disagree	1	9
Strongly disagree	0	0
No opinion	0	0

indicative feedback: comments

Positive points

Liked course <u>content</u>; course is good; well structured,

Balanced <u>workload</u> (lectures, homework, papers); workload appropriate to learn skills

<u>Lectures</u> are interesting

<u>HWs</u> are well guided; HWs are clear and well thought through; useful HW

<u>Paper</u> discussion is favorite part; appreciate discussion-based nature of course

<u>Project</u> is useful for developing analysis skills; gave enough freedorm

<u>Instructor</u> is agreeable

Points to improve

<u>Projects</u>

Hard to have topic ideas at beginning of course; provide list of possible project topics

Project relevance is not clear; project seems a bit redundant given the HWs

<u>Homework</u>

Clarify that data in HW#2 will be reused in HW#3

<u>Lectures</u>

Need to be more stimulating

in-depth evaluation

all DH spring 2022 courses will have online in-depth evaluation

this will be launched next week, you will receive a link to the survey

reminder: project schedule

1. **team building: DONE**

   email the list of your team members on **Week 2: Fri 04.03.2022**

   each team will have a designated project mentor

2. **project pitch: DONE**

   5-minute presentation of your project on **Week 5: Fri 25.03.2022**

   structure: title, problem, goals, approach, evaluation

3. **project progress presentation: : DONE**

   5-minute presentation per team on **Week 10: Fri 29.04.2022**

4. **final project presentation** on **Fri 10.06.2022**

   talk: 25-minute presentation + 20-minute questions

   schedule: 09:00-15:30

5. **final project report** by **Fri 17.06.2022**

   ACM conference paper format (6 pages + references + appendix)

project presentation and report

presentation
- each team has 45-min slot: 25-min talk + 20 mins for questions
- all team members are expected to present
- each member's contribution needs to be explained (who did what)
- structure: intro, goals, data, methods, results & discussion, conclusion

report
- ACM conference paper format: 6 double-column pages + references (not counted in 6-pp.) + appendices (if needed, not counted in 6-pp.)
- latex template available (ask your project mentor in case of questions): https://www.acm.org/publications/proceedings-template
- structure: abstract, introduction, data, methods, results & discussion, conclusion, references
- introduction: include project goals and description of each team member's contribution (who did what and who wrote what)
- a collaborative tool like overleaf is recommended
- submit slides & report by **Fri 17.06.2022, 7pm**

final project presentation day (10.06.2022)

09:00-09:45	group 1
09:45-10:30	group 2
10:30-10:45	break
10:45-11:30	group 3
11:30-12:15	group 4
12:15-13:00	lunch break
13:00-13:45	group 5
13:45-14:30	group 6
14:30-14:45	break
14:45-15:30	group 7

+ everybody is invited to attend the full day
+ please reserve the slot for your team
+ room to be confirmed (most likely ELD020)

this lecture

1. the rise of online video

2. conceptualizing YouTube

     #1: YouTube is a strange place

     #2: YouTube is a popular place

     #3: YouTube is a place for pirates

     #4: YouTube is a marketplace

     #5: YouTube is a place for expression

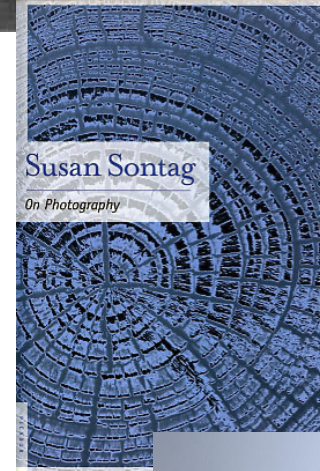   #6: YouTube is a data source

1.
the rise of online video

"Everything exists to end up in a book"

Stéphane Mallarmé, 1842-1998

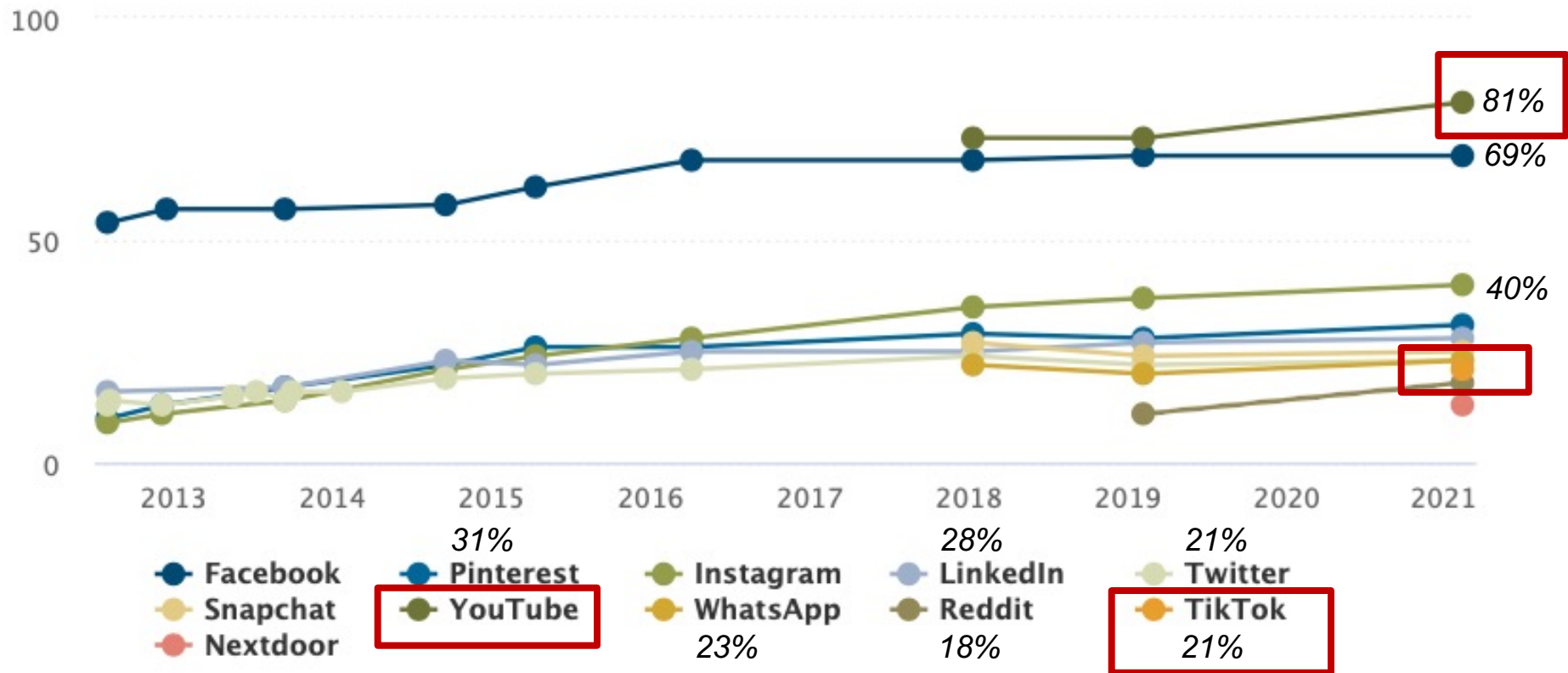"Everything exists to end up in a photograph"

Susan Sontag, *On Photography*, 1977

"Everything exists to end up in YouTube"

reminder: participation in social media

% of U.S. adults who say they ever use ...

81%
69%
40%
31%
28%
21%
23%
18%
21%

Facebook   Pinterest   Instagram   LinkedIn   Twitter
Snapchat   YouTube   WhatsApp   Reddit   TikTok
Nextdoor

https://www.pewresearch.org/internet/fact-sheet/social-media/

YouTube statistics

(last accessed may 2022, some numbers from earlier years)

founded in February 2005

#2 most popular website (after google)

2022 stats:

2B logged-in monthly users

500+ hours of content uploaded per minute
   (300h in 2015; 100h in 2014)

earlier stats:

1B hours of video watched every day (2020)

mobile: over 70% of global watchtime (2015)

more 18-34 & 18-49 year-olds than any cable network in US (2015)

https://www.youtube.com/about/press/

image (cc): alan klim @ flickr

Vintage TV designed by Alan Klim. Copyrighted material. Do not reproduce, modify or use without the author's permission

synchronous
few-to-many
passive
centralized

asynchronous
many-to-many
interactive
decentralized

this lecture

1. the rise of online video
2. conceptualizing YouTube
  #1: YouTube is a strange place
  #2: YouTube is a popular place
  #3: YouTube is a place for pirates
  #4: YouTube is a marketplace
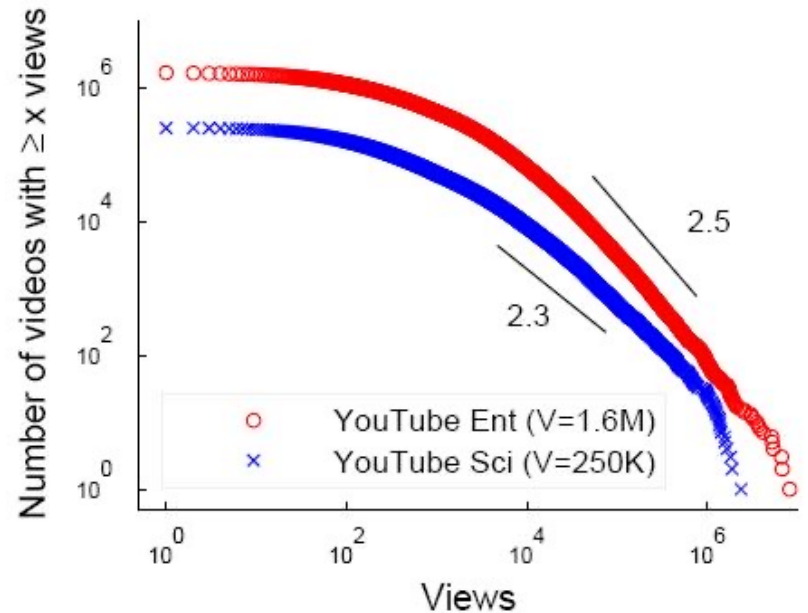 #5: YouTube is a place for expression
  #6: YouTube is a data source

View #1: YouTube is a strange place

YouTube and the | silliness / rudeness / absurdity | of crowds

J. Stossel, ABC 20/20:

"Do you like watching kids doing stupid and reckless things?

Beauty queens falling down?

Or a thousand prisoners dancing to the music of Thriller?

It's all in YouTube"

Burgess & Green:

"Rather than video about nothing, this could be situated in the much longer history of *vernacular creativity – the wide range of everyday creative practices* ... practiced outside the cultural value systems of high culture or commercial practice"

J. Burgess and J. Green, *YouTube. Online Video and Participatory Culture*, Polity, 2009

Survey: YouTube Stars More Popular Than Mainstream Celebs Among U.S. Teens

EMAIL   *81*   450   56   7

PRINT   TALK   g+1   Tweet   f Share

AUGUST 5, 2014 | 09:00AM PT

Susanne Ault

YouTube cultural 'generativity'
(Burgess & Green, 2009)

«Multiple roles: a **broadcast platform**, a **media archive**, and a **social network**»

«Various forms of cultural, social, and economic values are collectively **produced by users,** via their consumption, evaluation, and entrepreneurial activities»

«Double function as both a **top-down** platform for the distribution of popular culture and a **bottom-up** platform for vernacular creativity»
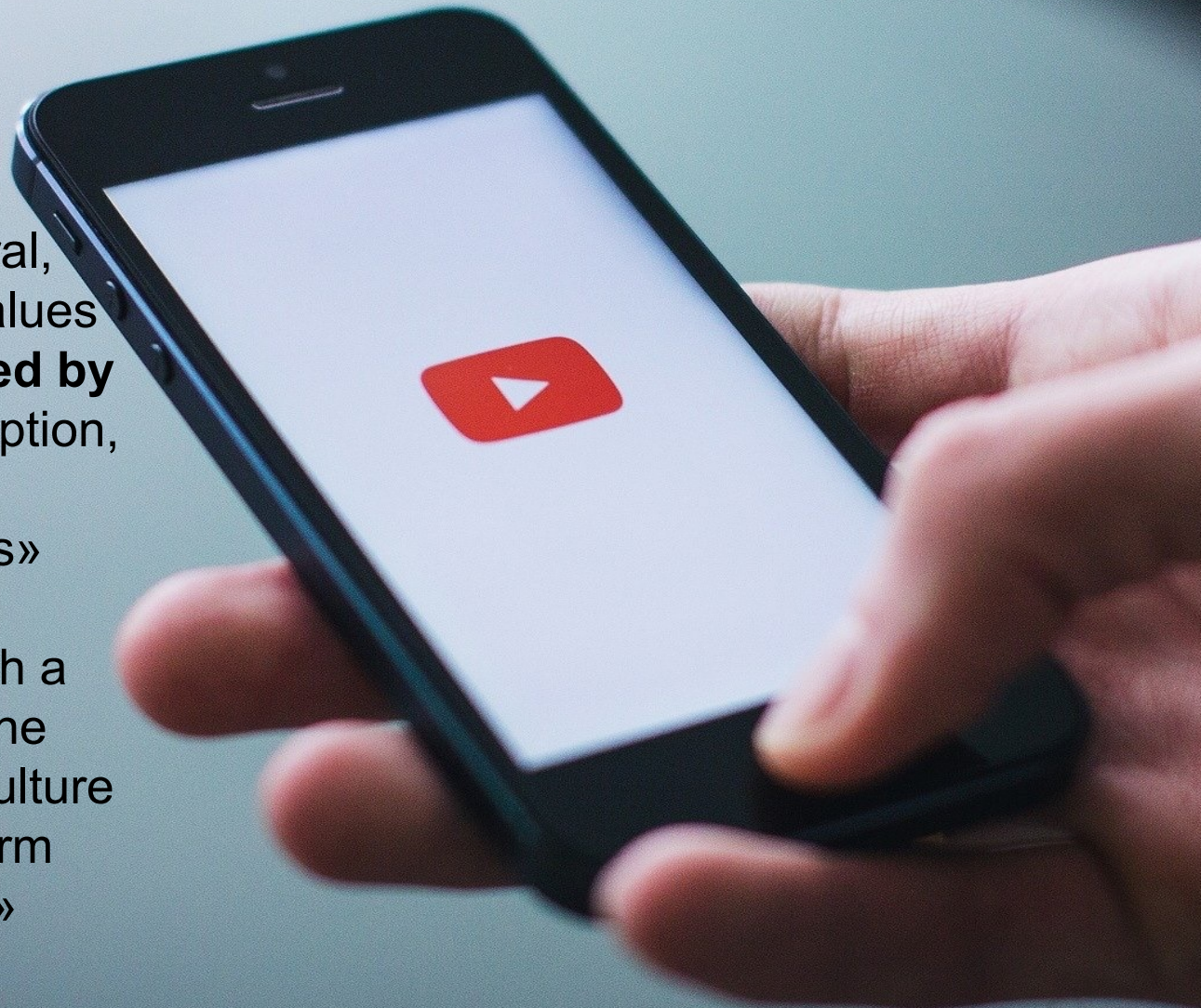
Image: StockSnap @Pixabay

this lecture

1. the rise of online video
2. conceptualizing YouTube

      #1: YouTube is a strange place

      #2: YouTube is a popular place

      #3: YouTube is a place for pirates

      #4: YouTube is a marketplace

      #5: YouTube is a place for expression

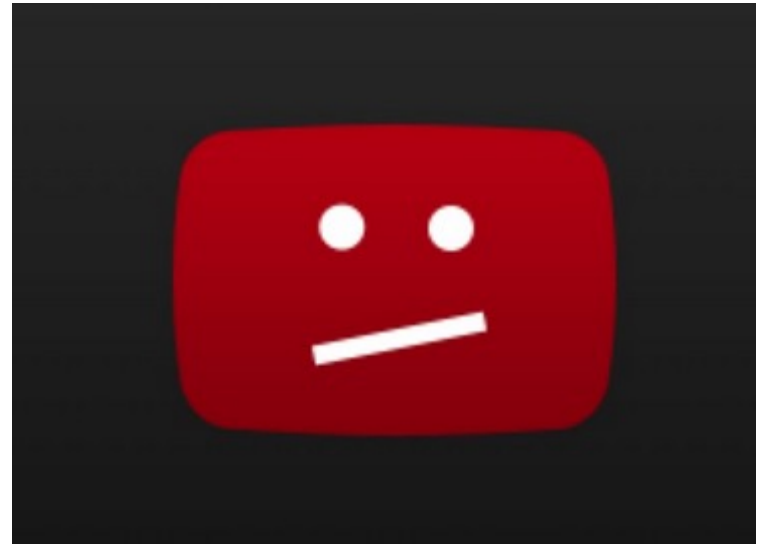   #6: YouTube is a data source

View #2: YouTube is a popular place

YouTube and popular videos

video popularity distributions (power-law with truncated tails)

large-scale analysis (10^6 videos)

no content was analyzed

video popularity distribution of YouTube: power-law in waist, sharp decay in tail

M. Cha, H. Kwak, P. Rodriguez, Y.-Y. Ahn and S. Moon, "I Tube, You Tube, Everybody Tubes: Analyzing the World's Largest User Generated Content Video System," ACM Internet Measurement Conference, Oct. 2007

social network & content-diffusion platform

YouTube as a **content-driven social network**: social network where interaction is mediated by content (user-content-user)

explicit (subscriptions) & **implicit** (comments) graphs

low **reciprocity:** only 25% of users have one or more reciprocate links on the subscription graph

poor **assortative linking** («preference for a network's nodes to attach to others that have similar degree»): users subscribe to much more popular users

low **homophily** (26%) measured by the predominant upload category that users upload to

similar to Twitter regarding these features
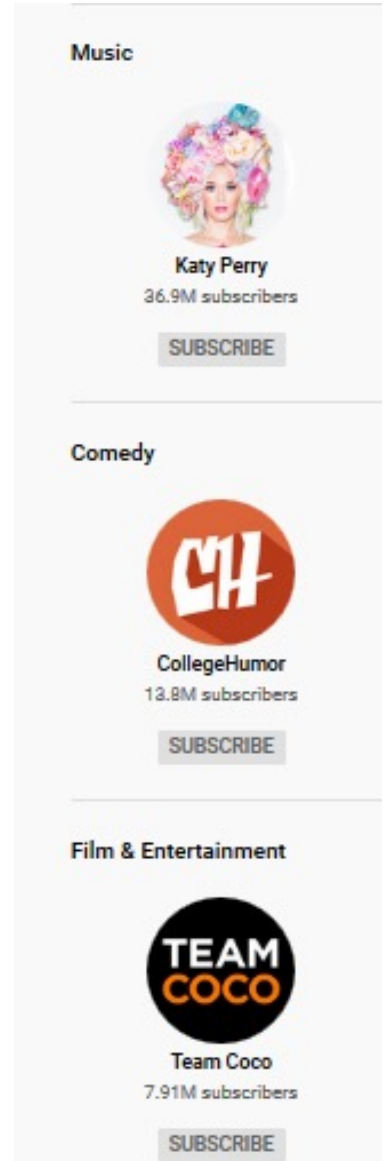
Music

Katy Perry
36.9M subscribers

SUBSCRIBE

Comedy

CollegeHumor
13.8M subscribers

SUBSCRIBE

Film & Entertainment

TEAM COCO

Team Coco
7.91M subscribers

SUBSCRIBE

M. Wattenhofer, R. Wattenhofer, Z. Zhu, The YouTube Social Network, AAAI Int. Conf. on Web and Social Media (ICWSM), 2012

this lecture

1. the rise of online video
2. conceptualizing YouTube
      #1: YouTube is a strange place
      #2: YouTube is a popular place
      #3: YouTube is a place for pirates
      #4: YouTube is a marketplace
      #5: YouTube is a place for expression
      #6: YouTube is a data source

View #3: YouTube is a place for pirates

YouTube and copyright infringement

near-duplicate (ND) detection: YouTube key business feature

YouTube partners with media companies to detect unauthorized copies

NDs are seen as redundant and often deemed to be eliminated

Yet, NDs are not always uploaded with bad intent

M. Cherubini, R. Oliveira, N. Oliver, "Understanding Near-Duplicate Videos: A User-Centric Approach", ACM Multimedia 2009

YouTube Content ID (launched 2007)

Copyright owners can use a system called Content ID to easily identify and manage their content on YouTube.

May 2020: "9,000+ partners use Content ID, including major network broadcasters, movie studios and record labels. They have claimed over 800 million videos, helping them control their content and make money on videos containing copyrighted material. More than 75 million active reference files in Content ID database."

https://support.google.com/youtube/answer/2797370?hl=en
https://www.youtube.com//about/press/

Content ID official video:

https://www.youtube.com/watch?v=9g2U12SsRns
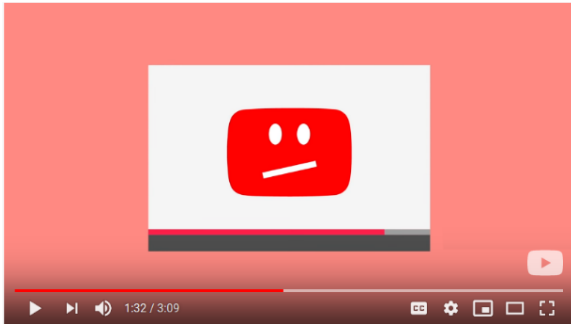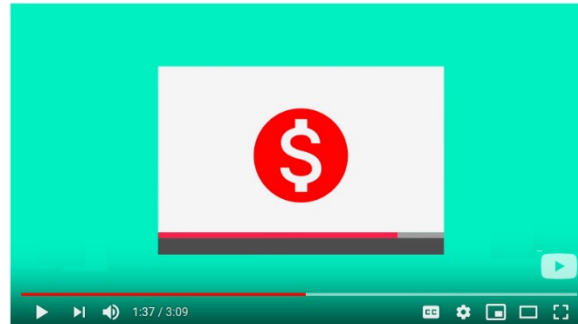
block content · monetize video · track viewer data to get detailed analytics

https://www.youtube.com/watch?v=9g2U12SsRns

this lecture

1. the rise of online video
2. conceptualizing YouTube
      #1: YouTube is a strange place
      #2: YouTube is a popular place
      #3: YouTube is a place for pirates
      #4: YouTube is a marketplace
      #5: YouTube is a place for expression
      #6: YouTube is a data source

View #4: YouTube is a marketplace

LSTN Headphones advertised their video on YouTube to share their story.

21,452 views • Feb 22, 2016

👍 153   👎 11   ➡ SHARE   ☰+ SAVE   •••

YouTube Advertisers ✓
342K subscribers

SUBSCRIBE

https://www.youtube.com/watch?v=A0QQWfJsixA&t=22s

Learn how Majestic Heli Ski uses YouTube and Video Ads to grow their business.

11,851 views • Aug 23, 2017

👍 41   👎 10   ➜ SHARE   ☰+ SAVE   •••

YouTube Advertisers ✔
342K subscribers

SUBSCRIBE

https://www.youtube.com/watch?v=wf-UMK7YCc4&t=102s

https://www.youtube.com/ads/making-a-video-ad/

YouTube Ads
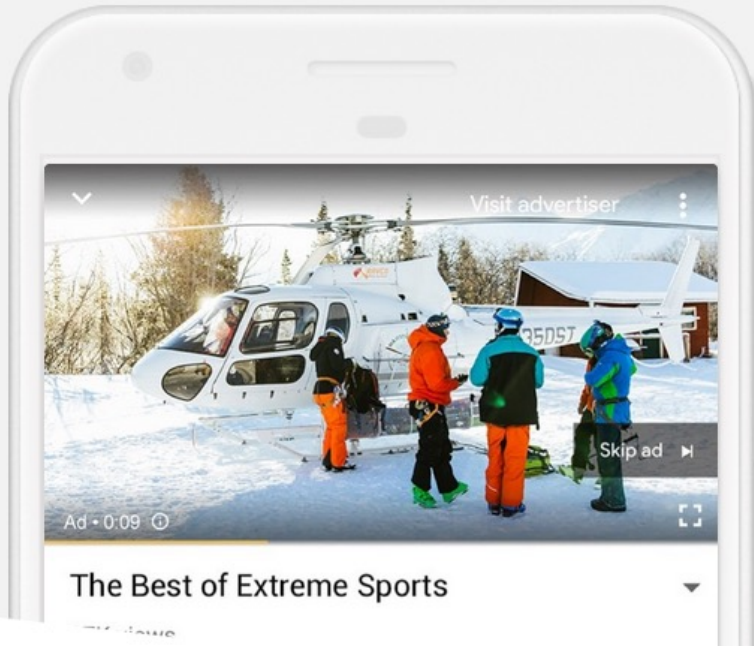
Only pay when they watch your video ad

You only pay when someone chooses to watch at least 30 seconds or clicks on your TrueView ad.

GET THE MOST FROM YOUR BUDGET →

Visit advertiser

Skip ad ▶|

Ad • 0:09 ⓘ

The Best of Extreme Sports ▾

You'll create your video ad with Google Ads

In a few easy steps your video ad will be up and running. Although your ad will appear on YouTube, you'll manage your campaign using Google Ads, an advertising platform used by businesses running ads on Google and its advertising network – which includes YouTube.

You'll now be redirected to sign in to or create a Google Ads account.

GET STARTED

HOW YOU'LL SET UP YOUR AD

1. Select your video from YouTube ^

If the video you want to use for your ad isn't already on YouTube, upload it now. Don't have a video? Get help here.

2. Choose who should see your ad ^

Select the locations where you want to show your ad and define the types of people you want to reach.

3. Choose how much you want to spend ^

Set the budget for your video ad campaign.

https://www.youtube.com/ads/signup/?sf=video_youtube&sourceid=awo&hl=en&subid=ww-ww-et-v_ads_yt_1index-new

this lecture

1. the rise of online video
2. conceptualizing YouTube
      #1: YouTube is a strange place
      #2: YouTube is a popular place
      #3: YouTube is a place for pirates
      #4: YouTube is a marketplace
      #5: YouTube is a place for expression
      #6: YouTube is a data source

View #5: YouTube is a place for expression

Source: Jawed Karim (Apr 2005)

Source: Joe Biden (Apr 2020)

nonverbal communication: what do YouTubers express?

gaze

proxemics

gestures

prosody

body posture

one world (MadV, Nov. 2006)

This is an invitation

to make a stand

Post your response now

ONE WORLD

https://www.youtube.com/watch?v=FHc0xjRGI60
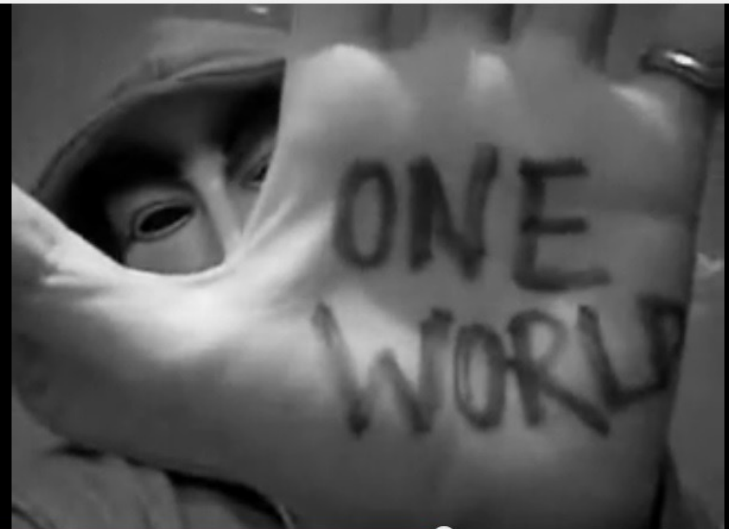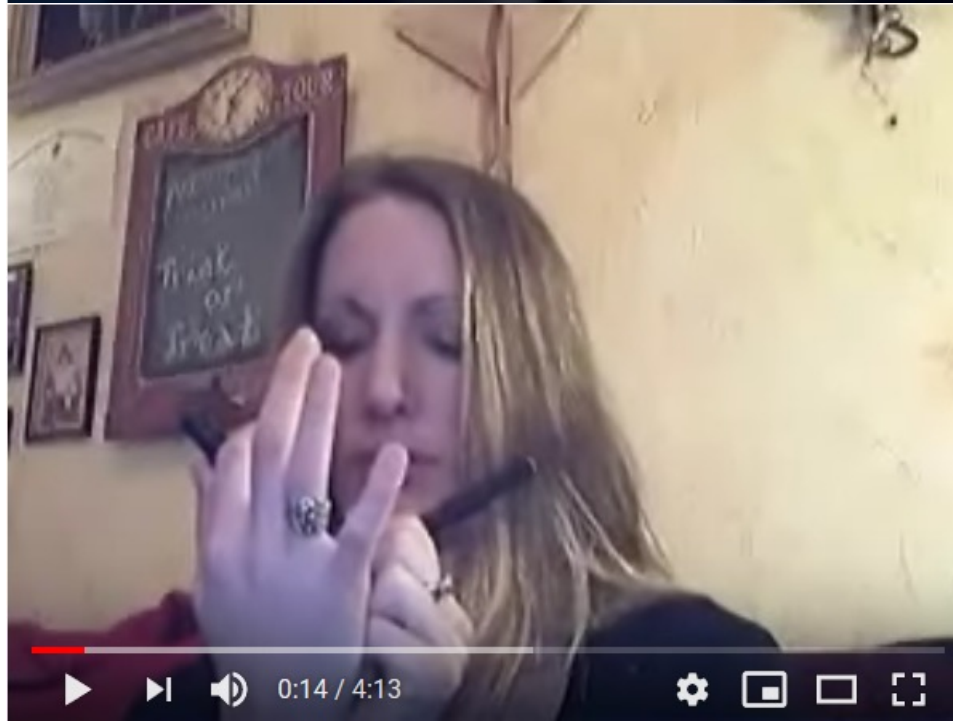
2000 video replies
https://www.youtube.com/watch?v=Z-BzXpOch-E
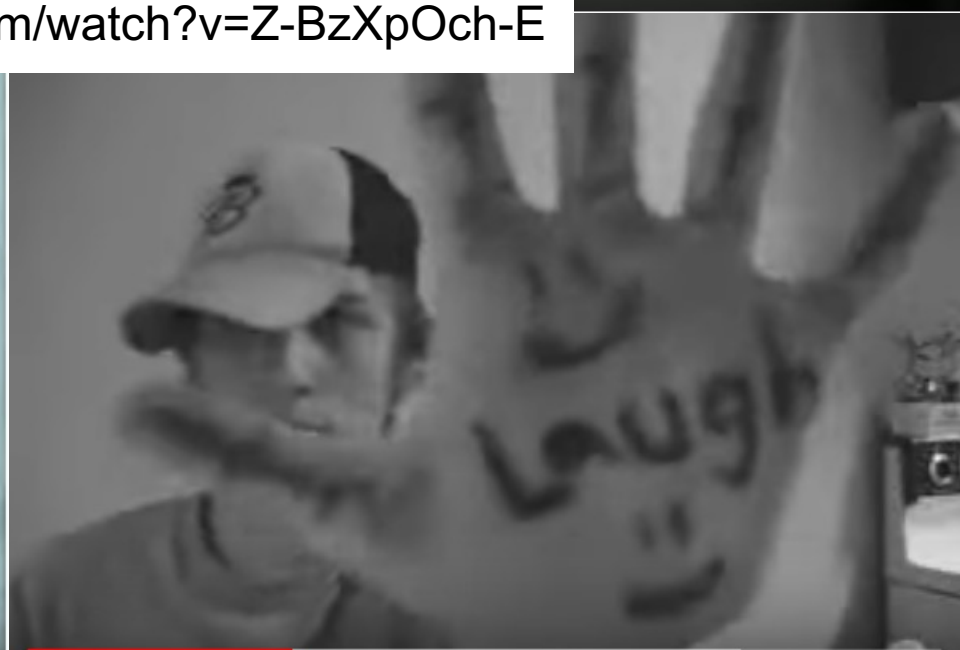
Eric Whitacre's Virtual Choir - 'Lux Aurumque'

Scroll for details

3:28 / 6:20

https://www.youtube.com/watch?v=D7o7BrIbaDs

Apr 16, 2020

https://www.youtube.com/watch?v=8c9QzDfFjxY

Apr 29, 2020

https://www.youtube.com/watch?v=vuQ0TvyRtsY

4000 submissions

learning media skills: making YouTube videos

PATRICIA G. LANGE

TECHNICAL IDENTITIES
AND
DIGITAL LITERACIES

KIDS ON YOUTUBE

2014

Ethnographic analysis of 200 videos &150 interviews of YouTube video makers (kids, youth, parents, and other adults)

"**Personally expressive media**: enables the creator to communicate aspects of the self [...] a wish to communicate personal ideas, feelings or experiences through video"

Media skills are "not developed overnight .. [but] … built up through the micro-interactions that individuals have when creating personally interesting media"

how-to videos: YouTube as a knowledge accelerator

documenting and sharing experiences over time

Health Vlogs as Social Support for Chronic Illness Management

JINA HUH, Michigan State University
LESLIE S. LIU and TINA NEOGI, University of Washington
KORI INKPEN, Microsoft Research
WANDA PRATT, University of Washington

YouTube users posting vlogs about experience with chronic illness

The first and last vlog of each user was coded w.r.t content to understand how life experience changed over time

Vlogs are used as a way of building community and provide support to individuals who experience illness

this lecture

1. the rise of online video
2. conceptualizing YouTube

      #1: YouTube is a strange place

      #2: YouTube is a popular place

      #3: YouTube is a place for pirates
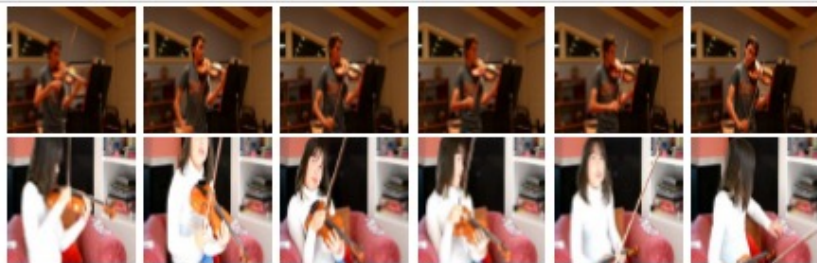
      #4: YouTube is a marketplace

      #5: YouTube is a place for expression

      #6: YouTube is a data source

View #6: YouTube is a data source

Kinetics human action video dataset:
400 actions, min. 400 YouTube videos each (306,000 videos)

playing violin
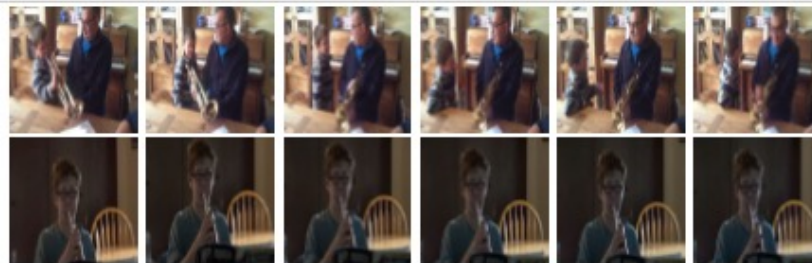
playing trumpet

braiding hair

brushing hair

dribbling basketball

dunking basketball

W. Kay, J. Carreira, K. Simonyan, B. Zhang, C. Hillier, S. Vijayanarasimhan, F. Viola, T. Green, T. Back, P. Natsev, M. Suleyman, and A. Zisserman. The kinetics human action video dataset. arXiv preprint arXiv:1705.06950, 2017

performance on Kinetics dataset

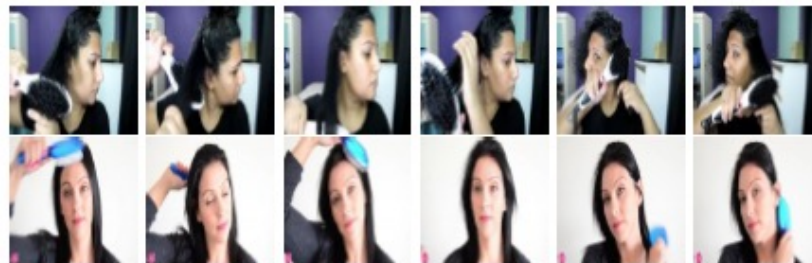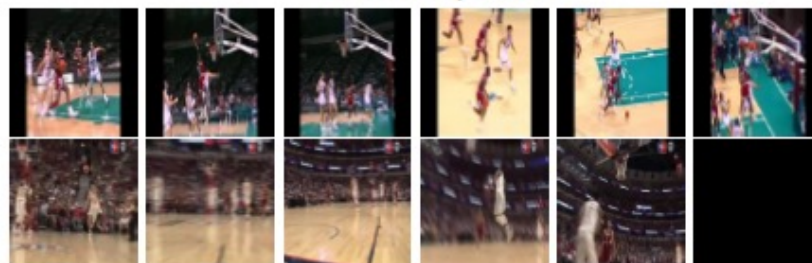| Architecture	Kinetics			ImageNet then Kinetics		
	RGB	Flow	RGB + Flow	RGB	Flow	RGB + Flow
(a) LSTM	53.9	–	–	63.3	–	–
(b) 3D-ConvNet	56.1	–	–	–	–	–
(c) Two-Stream	57.9	49.6	62.8	62.2	52.4	65.6
(d) 3D-Fused	–	–	62.7	–	–	67.2
(e) Two-Stream I3D	**68.4** (88.0)	**61.5** (83.4)	**71.6** (90.0)	**71.1** (89.3)	**63.4** (84.9)	**74.2** (91.3)

Table 3. Performance training and testing on Kinetics with and without ImageNet pretraining. Numbers in brackets () are the Top-5 accuracy, all others are Top-1.

J. Carreira and A. Zisserman. Quo Vadis, Action Recognition? A New Model and the Kinetics Dataset, CVPR 2017

YouTube-8M dataset

YouTube-8M is a large-scale labeled video dataset that consists of millions of YouTube video IDs, with high-quality machine-generated annotations from a diverse vocabulary of 3,800+ visual entities. It comes with precomputed audio-visual features from billions of frames and audio segments, designed to fit on a single hard disk. This makes it possible to train a strong baseline model on this dataset in less than a day on a single GPU! At the same time, the dataset's scale and diversity can enable deep exploration of complex audio-visual models that can take weeks to train even in a distributed fashion.

Our goal is to accelerate research on large-scale video understanding, representation learning, noisy data modeling, transfer learning, and domain adaptation approaches for video. More details about the dataset and initial experiments can be found in our technical report and in previous workshop pages (2018, 2017). Some statistics from the latest version of the dataset are included below.
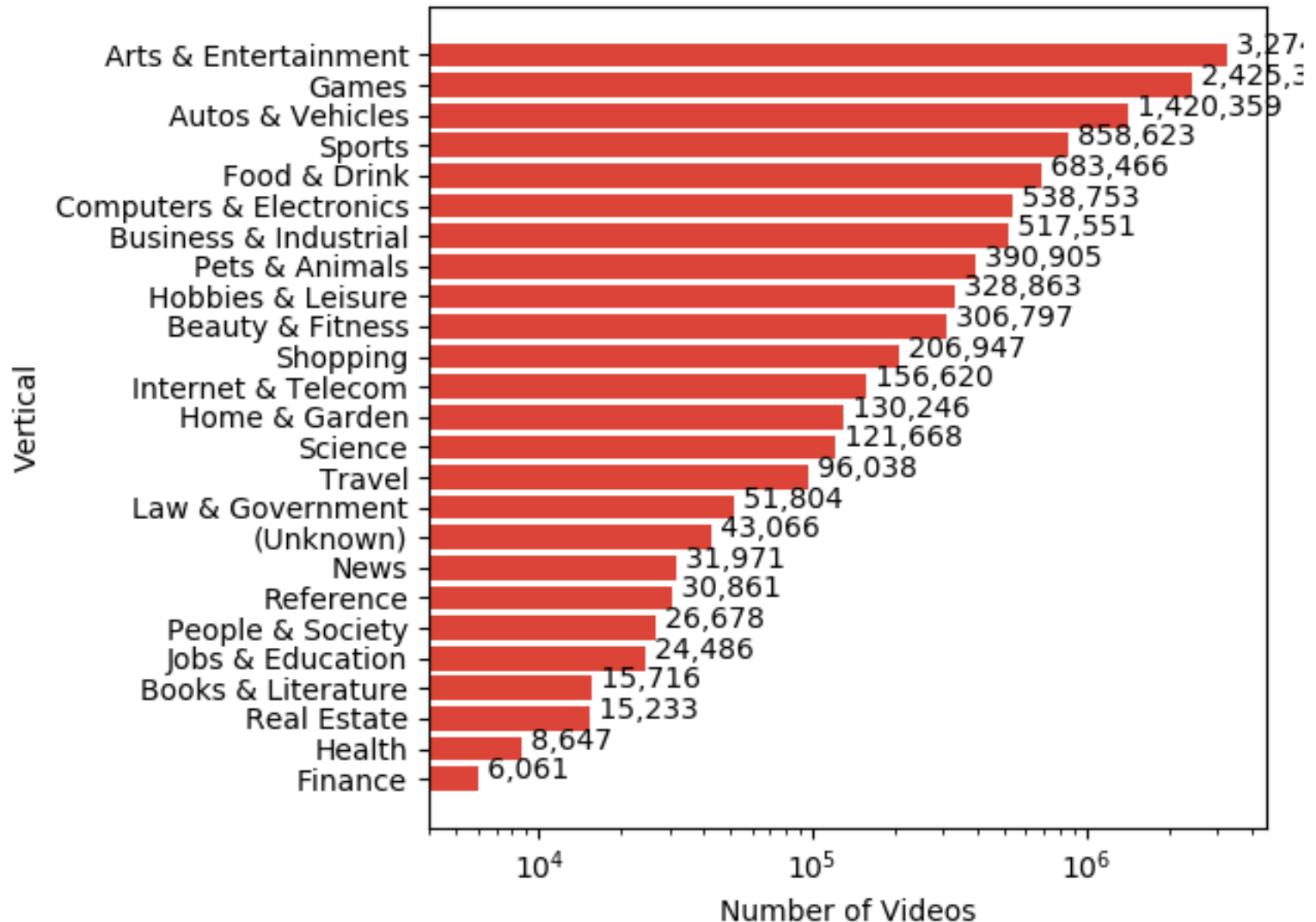
6.1 Million Video IDs	350,000 Hours of Video	2.6 Billion Audio/Visual Features	3862 Classes	3.0 Avg. Labels / Video

https://research.google.com/youtube8m/

number of videos per YouTube category

Bar chart showing Number of Videos per Vertical (log scale):

- Arts & Entertainment — 3,27...
- Games — 2,425,3...
- Autos & Vehicles — 1,420,359
- Sports — 858,623
- Food & Drink — 683,466
- Computers & Electronics — 538,753
- Business & Industrial — 517,551
- Pets & Animals — 390,905
- Hobbies & Leisure — 328,863
- Beauty & Fitness — 306,797
- Shopping — 206,947
- Internet & Telecom — 156,620
- Home & Garden — 130,246
- Science — 121,668
- Travel — 96,038
- Law & Government — 51,804
- (Unknown) — 43,066
- News — 31,971
- Reference — 30,861
- People & Society — 26,678
- Jobs & Education — 24,486
- Books & Literature — 15,716
- Real Estate — 15,233
- Health — 8,647
- Finance — 6,061

https://research.google.com/youtube8m/

what to remember

online video   is different than text & still images

continues to rise

YouTube is: multifaceted: broadcast platform,
media archive, social network

a top-down and bottom-up ecosystem

top-down: content providers &
advertisers looking for monetization

bottom-up: appropriated by users in
myriad ways, from creative expression
to misinformation

a data source for research (visual recognition)

questions?

daniel.gatica-perez@epfl.ch