

Advanced Topics

Social Spam, Campaigns, Misinformation and Crowdturfing

(Not part of Exam)

Large-Scale Social Systems

Online Social Networking			
Social Media			
Information sharing communities			
Social Games			 
Location-based Services			 
Crowd-based services			 

Key Organizing Principles

- Openness:
 - Social systems are inherently **open to users** who generate, share and consume information
 - E.g., post a message, upload and watch a video
- Collaboration:
 - Many users **organically** participate in social systems to engage in **collaborative activities**
 - E.g., organize for political change, share disaster-related information
- Real-time information propagation:
 - Users, media and organization post information related to **hot events** in (near) **real-time**
 - E.g., emergency alerts, natural disaster news and sports games
- Crowdsourcing **tasks** or hiring **cheap workers** from all over the **world**:
 - People can hire workers from crowdsourcing sites with paying little money
 - E.g., workers from Amazon Mechanical Turk for labeling data, workers from Fiverr for editing a document



Challenges and Research Approach

These necessary positive aspects may also lead to **negative consequences**

- Spam of many flavors
 - Comment spam (~90% on websites = 46 billion)
 - Spam tweets (1% = 3 million/day) and Twitter spammers (5% = 25 million)
 - Spam videos (20%)
- Traditional Attacks
 - Phishing, malware and etc
- Campaigns
- Misinformation
- Crowdturfing
- Misuse
 - Crowdsourcing the wrong guy in the Boston bombings at Reddit
- ...



Fake Accounts

9% on Facebook = 87 million accounts in 2012 [Facebook]



Comment Spam

- 83 ~ 90% on websites = 46 billion comments [Akismet and Mollom. 2010, Kant et al. WSDM 2012]

Rosiane facebook.com/profile.php?id=100003406202721 x m.smealen@mail.ru 100.143.232.12	Submitted on 2012/07/02 at 09:27 you people may not belivee at all but i can and will tell you that between heaven and earth are things beyond the reach of ordinary man and women.you people do not know what knowledge is and you would not gain any knowledge if its not by some devine revelation.is this the book of the devil maybe but it sure as hell is not for ordinary folks like you people to read, you could not handle it any one of you, before you open the book of the devil you better make sure your in a right pad with GOD Jehova.
Urvi facebook.com/profile.php?id=100003406194827 x info@sms-vluchtelingen.nl 188.143.232.12	Submitted on 2012/07/02 at 02:20 I had a spambot at my potrey site post something regarding the size of her husband.All i can say is Mr. Jeremy must be glad he isn't married to her.Then there's the one with the guy wanting to sell his bridal dresses.
best affiliate website home-businessreviews.com/Turkey-Affiliate-Websit... x justin@111558@gmail.com 46.109.196.107	Submitted on 2012/06/29 at 04:34 Make \$1,000's Weekly with a Health Internet Business of Your Very Own Now get a complete fully-operational "Health eBiz" in a box! This amazing site: * Closes sales automatically for you! * Has a complete electronic sales manager that makes all upsells for you!



Spam Tweets and Twitter Spammers

1% Spam tweets and 5% Twitter spammers

- 3 million spam tweets/day and 25 million spam accounts

[Twitter and TwitSweeper, 2010]



[cphmlink](#): for all those wondering about the twitter trend i do believe **Apple Shampoo** refers to a Blink 182 song.

10 minutes ago from web · [Reply](#) · [View Tweet](#)



[TrevBusiness](#): [#apple](#) shampoo Getting strange calls? Reverse Phone Detective to find out who's bothering you! www.SmallBusinessSolved.com/r/

10 minutes ago from web · [Reply](#) · [View Tweet](#)



[TrevBusiness](#): [#apple](#) shampoo This website to give me great ideas to do on a daily to get the love flowing www.SmallBusinessSolved.com/m/

10 minutes ago from web · [Reply](#) · [View Tweet](#)



UNIVERSITY OF
TORONTO

Spam Videos

- **183** million U.S. Internet users watched more than **37 billion** online videos in Oct 2012. [comScore]
- **20%** of online videos are spam [VideoSurf]

Amanda Knox: Murder on Trial VIDEO

dowopor

12 videos

Subscribe



Like Add to Share

Uploaded by dowopor on Oct 3, 2011

Click on <http://hotznews.com/wow/> - FREE to watch Amanda Knox: Murder on



UNIVERSITY OF
TORONTO

Collective Attention Spam

- Target popular and trendy topics/items
- Feed spam contents once the topics/items become popular

Results for #DrakeCriesWhen

Tweets · All

Refine results



CherryIKornblut Cheryl Kornbluth

No way. She pulls this again! [bit.ly/oFalmo](#) #gritosmexicanos #DrakeCriesWhen #FastFoodAddiction Glen Rice

18 seconds ago



AprylHemmelgarn Apryl Hemmelgarn

I wonder if this really works [bit.ly/oFalmo](#) Glen Rice #UKnowUHungryWhen #gritosmexicanos #DrakeCriesWhen #FastFoodAddiction

18 seconds ago



UnSchwegel6358 Un Schwegel

Anybody know is this really works!?? [bit.ly/oFalmo](#) #gritosmexicanos Glen Rice #UKnowUHungryWhen #DrakeCriesWhen #FastFoodAddiction

18 seconds ago



DebraRozzi7831 Debra Rozzi

No way. She pulls this again! [bit.ly/oFalmo](#) #gritosmexicanos #DrakeCriesWhen #FastFoodAddiction Glen Rice

17 seconds ago



MaryaLicchetto9 Marya Licchetto

Omg...Is this real? [bit.ly/oFalmo](#) #gritosmexicanos #DrakeCriesWhen #FastFoodAddiction Glen Rice

16 seconds ago



How To Really Make Money Online, FAST!

Welcome Video

Training Video One

Training Video Two

Training Video Three

Join The Contest

Join The Contest Page



Video Training

Contest

Join The Contest

Contest Leader

First of all... Have you Checked Out The Amazing Free Fast Cash Commissions Training? It teaches a super simple yet effective way to make easy money online, it's a must see! It's Super Easy... All You Have To Do Is Sign Up & Spread Your Contest Link Online!



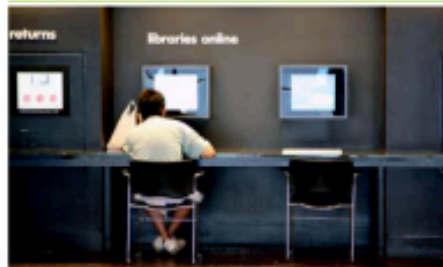
UNIVERSITY OF
TORONTO

Campaigns

Astroturfing

The need to protect the internet from 'astroturfing' grows ever more urgent

The tobacco industry does it, the US Air Force clearly wants to ... astroturfing – the use of sophisticated software to drown out real people on web forums – is on the rise. How do we stop it?



A real person using the internet. Unfortunately we can no longer assume what we are reading is written by one of these creatures. Photograph: Jeff Blackler/Fox Features

Fake review campaign

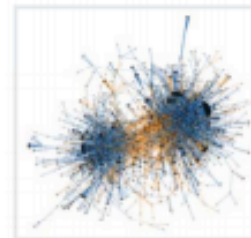
<p>1 of 1 people found the following review helpful: ★★★★★ Practically FREE music, December 4, 2004 <i>This review is from: Audio Xtract (CD-ROM)</i> I can't believe for \$10 (after rebate) I got a program that gets me free unlimited music. I was hoping it did half what was ...</p>	<p>2 of 2 people found the following review helpful: ★★★★★ Like a tape recorder..., December 8, 2004 <i>This review is from: Audio Xtract (CD-ROM)</i> This software really rocks. I can set the program to record music all day long and just let it go. I come home and my ...</p>	<p>★★★★★ Wow, internet music! ..., December 4, 2004 <i>This review is from: Audio Xtract (CD-ROM)</i> I looked forever for a way to record internet music. My way took a long time and many steps (frustrating). Then I found Audio Xtract. With more than 3,000 songs downloaded in ...</p>
<p>3 of 8 people found the following review helpful: ★★★★★ Yes – it really works, December 4, 2004 <i>This review is from: Audio Xtract Pro (CD-ROM)</i> See my review for Audio Xtract - this PRO is even better. This is the solution I've been looking for. After buying iTunes, ...</p>	<p>3 of 10 people found the following review helpful: ★★★★★ This is even better than..., December 8, 2004 <i>This review is from: Audio Xtract Pro (CD-ROM)</i> Let me tell you, this has to be one of the coolest products ever on the market. Record 8 internet radio stations at once, ...</p>	<p>2 of 9 people found the following review helpful: ★★★★★ Best music just got ..., December 4, 2004 <i>This review is from: Audio Xtract Pro (CD-ROM)</i> The other day I upgraded to this TOP NOTCH product. Everyone who loves music needs to get it from Internet ...</p>
<p>5 of 5 people found the following review helpful: ★★★★★ My kids love it, December 4, 2004 <i>This review is from: Pond Aquarium 3D Deluxe Edition</i> This was a bargain at \$20 - better than the other ones that have no above water scenes. My kids get a kick out of the ...</p>	<p>5 of 5 people found the following review helpful: ★★★★★ For the price you..., December 8, 2004 <i>This review is from: Pond Aquarium 3D Deluxe Edition</i> This is one of the coolest screensavers I have ever seen, the fish move realistically, the environments look real, and the ...</p>	<p>3 of 3 people found the following review helpful: ★★★★★ Cool, looks great..., December 4, 2004 <i>This review is from: Pond Aquarium 3D Deluxe Edition</i> We have this set up on the PC at home and it looks GREAT. The fish and the scenes are really neat. Friends and family ...</p>

Big John's Profile

Cletus' Profile

Jake's Profile

Political campaign



COMPUTING

Bogus Grass-Roots Politics on Twitter

Data-mining techniques reveal fake Twitter accounts that give the impression of a vast political movement.

TUESDAY, NOVEMBER 2, 2010 | BY KURT KLEINER

1/21 PAGES

How true? This network graph shows the connections between 6,278 accounts that used the hashtag #gop in September and October 2010.
Indiana University

Researchers have found evidence that political campaigns and special-interest groups are using scores of fake Twitter accounts to create the impression of broad grass-roots political expression. A team at Indiana University used data-mining and network-analysis

techniques to detect the activity.

"We think this technique must be common," says **Filippo Menczer**, an associate professor at Indiana University and one of the principal investigators on the project. "Wherever there are lots of eyes looking at screens, spammers will be there; so why not with politics?"

Website	Campaigns	% Crowdturfing	Tasks	\$ per Subm.
Amazon Turk (US)	41K	12%	2.9M	\$0.092
ShortTask* (US)	30K	95%	527K	\$0.096
MinuteWorkers (US)	710	70%	10K	\$0.241
MyEasyTask (US)	166	83%	4K	\$0.149
Microworkers (US)	267	89%	84K	\$0.175

Wang et al. WWW 2012



UNIVERSITY OF
TORONTO

Adversarial Propaganda

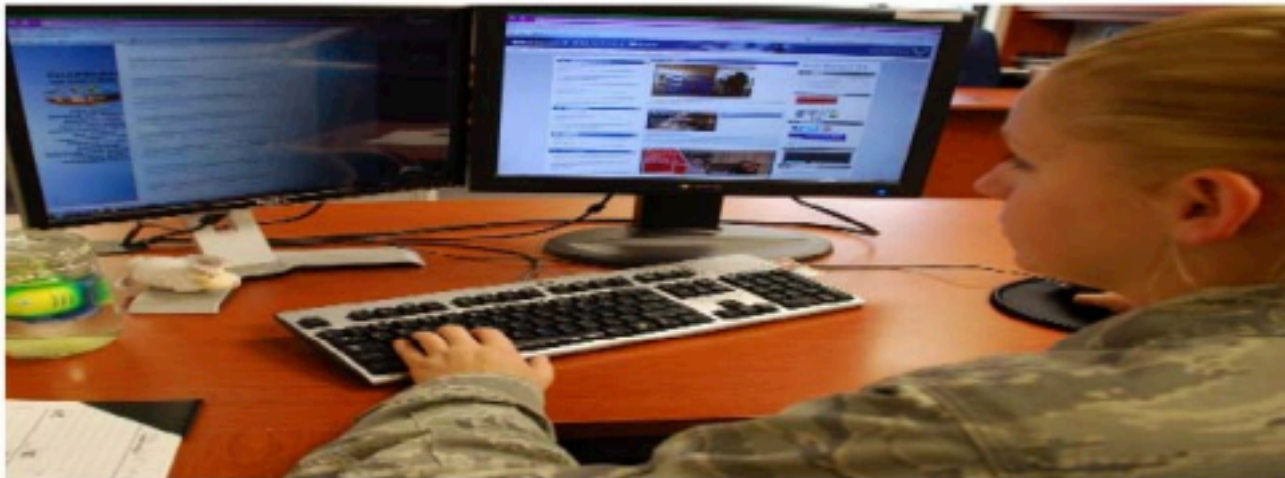
- Create and spread rumors and Misinformation
- Target a product/ government

Pentagon Wants a Social Media Propaganda Machine

BY ADAM RAWNSLEY 07.15.11 2:40 PM

[Follow @arawnsley](#)

[Like](#) 1.4k
[Tweet](#) 835
[+1](#) 113
[Share](#) 165



You don't need to have 5,000 friends of Facebook to know that social media can have a notorious mix of rumor, gossip and just plain disinformation. The Pentagon is looking to build a tool to sniff out social media propaganda campaigns and spit some counter-spin right back at it.

On Thursday, Defense Department extreme technology arm Darpa unveiled its [Social Media in](#) [Wired]



UNIVERSITY OF
TORONTO

Misinformation (Fake)



McDonalds in Virginia Beach flooded.

pic.twitter.com/FZBoCyDM

Reply Retweet Favorite



Fake
Images



I TOLD Y'ALL! Shark on the highway in New Jer
@maxthewanted would appreciate this. #Hurric
pic.twitter.com/kaYMjWzT

1:09 AM - 30 Oct 2012



UNIVERSITY OF
TORONTO

Crowdturfing (Crowdsourcing + Astroturfing)

- A **Multimillion-dollar industry** in Chinese crowdsourcing sites
 - **90%** crowdturfing tasks [MIT Technology Review]
- **70~95%** crowdturfing tasks at several U.S. crowdsourcing sites [Wang et al., WWW 2012]

Twitter Post: CPP Scam

Work done: 222/250

Employer: [Member_968289](#) 

You will earn **\$0.60**

This task takes less than **30** min to finish

Job ID: 364488d297e8

What is expected from Workers?

You must have 50 Twitter followers. Make sure you are logged into your Twitter account

1. Open your browser and search on Google "college pro painters success"
2. Click on any search result that starts with `collegepropaintersscam.com`
3. Go to Home Page of the website
4. Retweet any article



Examples of Crowdturfing

- Vietnamese propaganda spread by 1,000 crowdturfers

Vietnam admits deploying bloggers to support government

By Nga Pham
BBC News, Hanoi

Vietnamese propaganda officials have admitted deploying people to engage in online discussions and post comments supporting the Communist Party's policies.

The party has also confirmed that it operates a network of nearly 1,000 "public opinion shapers".

They are assigned with the task of spreading the party line.

The tactic is similar to China's model of internet moderators who aim to control news and manipulate opinion.

'Political opportunists'

Hanoi Propaganda and Education Department head Ho Quang Loi said that the authorities had hired hundreds of so-called "internet polemist" in the fight against "online hostile forces".



The bloggers have been hailed for stopping negative online rumours



Social Spam

- Fake accounts (5 ~ 6 % on Facebook = 42 million)
 - [Facebook. 2012]
- Comment spam (83 ~ 90% on websites = 46 billion)
 - [Akismet and Mollom. 2010, Kant et al. WSDM 2012]
- Spam Tweets (1% = 3 million/day) and Twitter Spammers (5% = 25 million)
 - [Twitter. 2010, TwitSweeper. 2010, Lee et al. SIGIR 2010, Lee et. al ICWSM 2011, Yang et al. WWW 2012]
- Tag spam
 - [Koutrika et al. TWEB 2008, Krause et al. AIRWEB 2008 , Neubauer et al. AIRWEB 2009]
- Spam videos
 - [Benevenuto et al. AIRWeb 2008, Benevenuto et al. SIGIR 2009]
- Fake Reviews
 - [Jindal and Bing ICDM 2007, Lim et al. CIKM 2010, Wang et al. TIST 2011, Mukherjee et al. WWW 2012]
- Voting spam
 - [Bian et al. AIRWEB 2008, Tran et al. NSDI 2009]
- Wikipedia vandalism
 - [Potthast et al. ECIR 2008, Chin et al. WICOM 2010, Adler et al. CICLing 2011]
- ...



How to detect these?

Blacklisting URLs

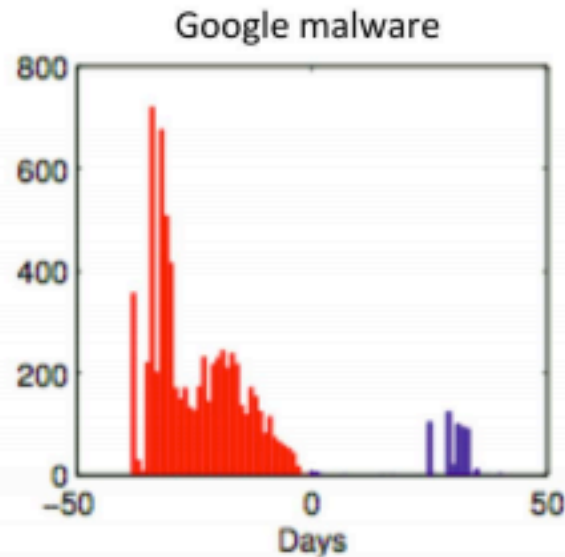
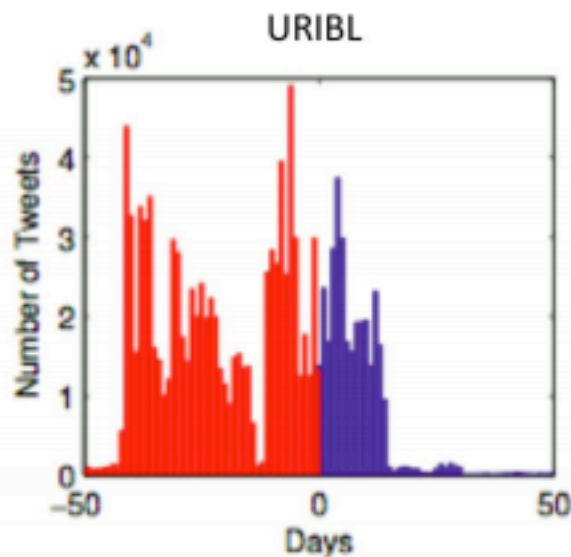
- Crawled URLs from Twitter
 - 25 million URLs crawled
 - 8% of them link to spam pages
- Over 80% of spam URLs were shortened
 - Mask landing site
 - <http://bit.ly/aLEmck> -> <http://i-drugspedia.com/pill/Viagra...>
 - Defeat blacklist filtering
 - bit.ly -> short.to -> malware landing page

Grier, C., Thomas, K., Paxson, V., and Zhang, M. @spam: the underground on 140 characters or less. In CCS, 2010.



Blacklist Performance

- Blacklists are slow to list spam domains
 - 80% of clicks are seen in first day
- Retroactively blacklist



Red = Lag

Blue = Lead



Comparison to Email Click Through

- Spam Email clickthrough: .003-.006%
 - From Spamalytics, Kanich et al. CCS 2008
- Twitter clickthrough: .13%
 - Collected 245,000 spam URLs
 - Define clickthrough as clicks / reach
 - Reach defined as *tweets * followers*



Social Spam Detection Approaches

- Supervised spam detection approach
 - The most popular approach
 - Require labeled data for training purpose
- Ranking users based on their social graph
- Use crowd wisdom (humans) to identify fake accounts



Open Research Challenges

- Need for large, accurate, up-to-date data sets
 - APIs
 - Hard crawling
 - Shared datasets
 - Purchasing data (e.g., Gnip)
 - Data grant or know an insider
- Labeling
 - Manual labeling
 - Use crowd wisdom
 - Get labeled data from a social media site
 - Blacklist



Open Research Challenges

- Integration of multiple techniques for data processing and modeling
 - Big data analysis, machine learning (data mining), information retrieval, visualization, etc
- Interdisciplinary research for analysis
 - computer science, social science, psychology, etc
- Arms race (endless battle)
 - Spammers and malicious users change their behaviors or use new techniques to avoid existing detection approaches
 - Spammers and malicious users move to another site



Useful Tools

- Machine learning
 - Weka: <http://www.cs.waikato.ac.nz/ml/weka/>
 - scikit-learn: <http://scikit-learn.org/stable/>
 - LingPipe (linguistic analysis): <http://alias-i.com/lingpipe/>
- Visualization
 - Matplotlib: <http://matplotlib.org/>
 - Gephi: <https://gephi.org/>
 - Graphviz: <http://www.graphviz.org/>



Useful Tools

- Big data analysis and visualization
 - Hadoop (MapReduce): <http://hadoop.apache.org/>
 - Pig: <https://pig.apache.org/>
 - Hive: <https://hive.apache.org/>
 - Cascalog: <http://cascalog.org/>
 - Giraph: <https://giraph.apache.org/>
- Scalable machine learning
 - Mahout: <https://mahout.apache.org/>
- Large scale stream processing
 - Storm: <http://storm.incubator.apache.org/>
 - Summingbird: <https://github.com/twitter/summingbird>



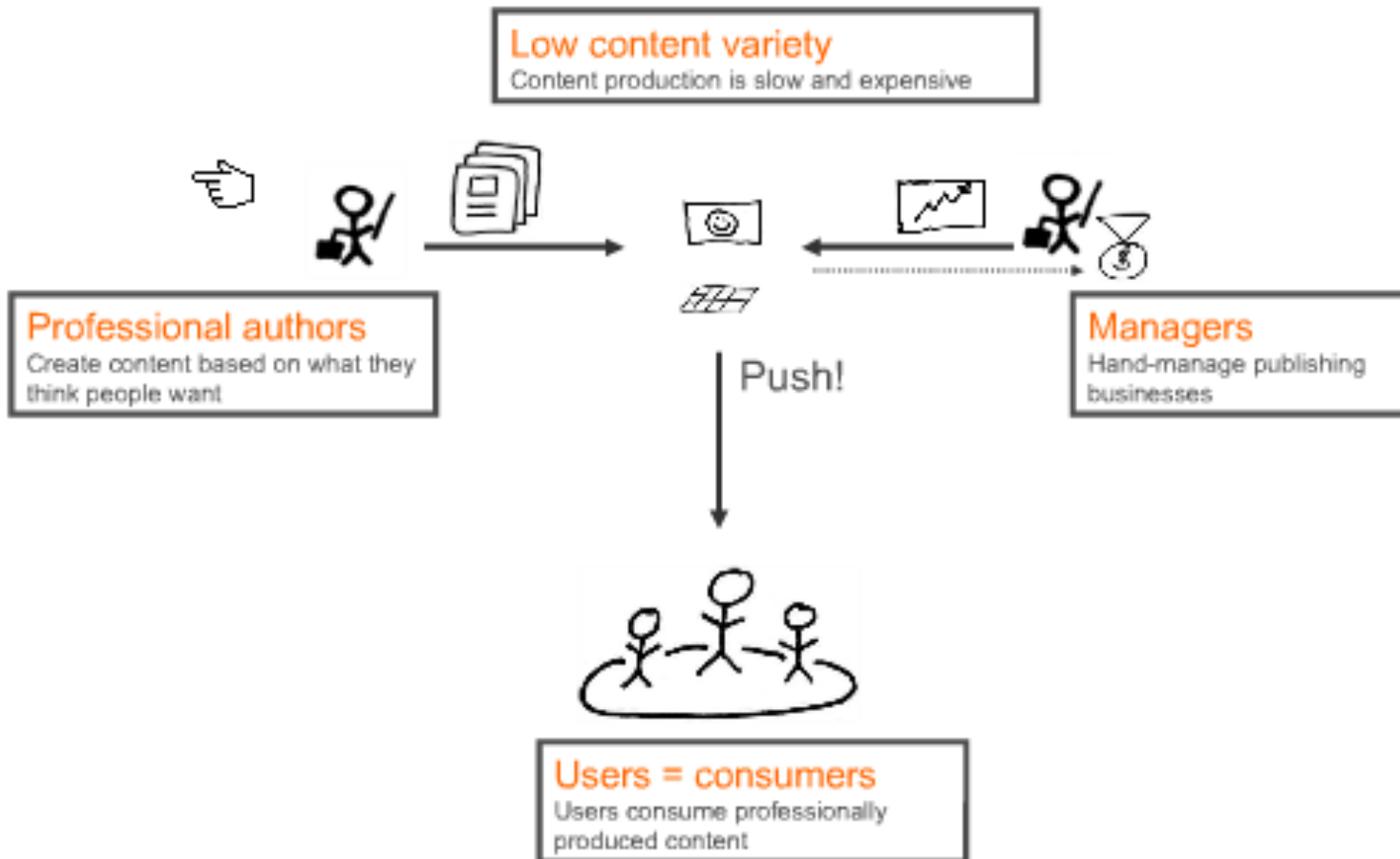
Web 3.0

Web 1.0



web1.0

Media companies put content in the web and push it to users.



UNIVERSITY OF
TORONTO

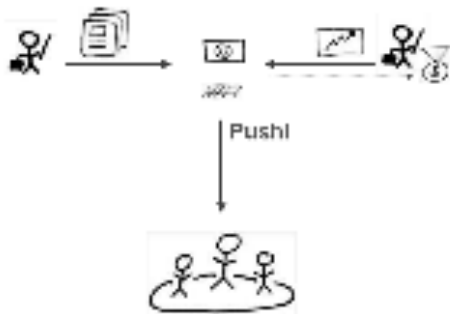
Web 1.0



web1.0 problem

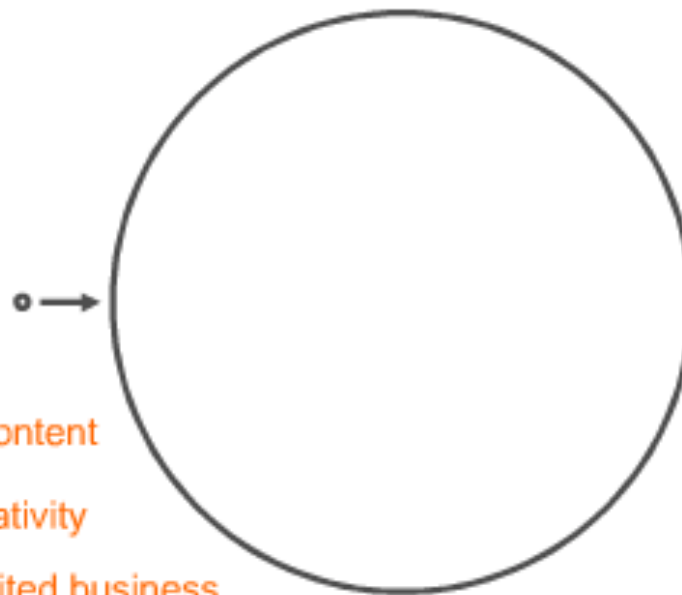
This old media model ignored the new possibilities of the internet.

web1.0



Problems

10,000 editors serve 500,000,000 internet users



Limited content

Limited creativity

Limited business



UNIVERSITY OF
TORONTO

Web 2.0



web2.0

New platforms allow users to generate content themselves

User generated content

Blogs, videos, photos, music etc.

High content variety

dramatically increases, while technical quality goes down.

Authoring platforms

enable everyone to publish!

Web2.0 managers

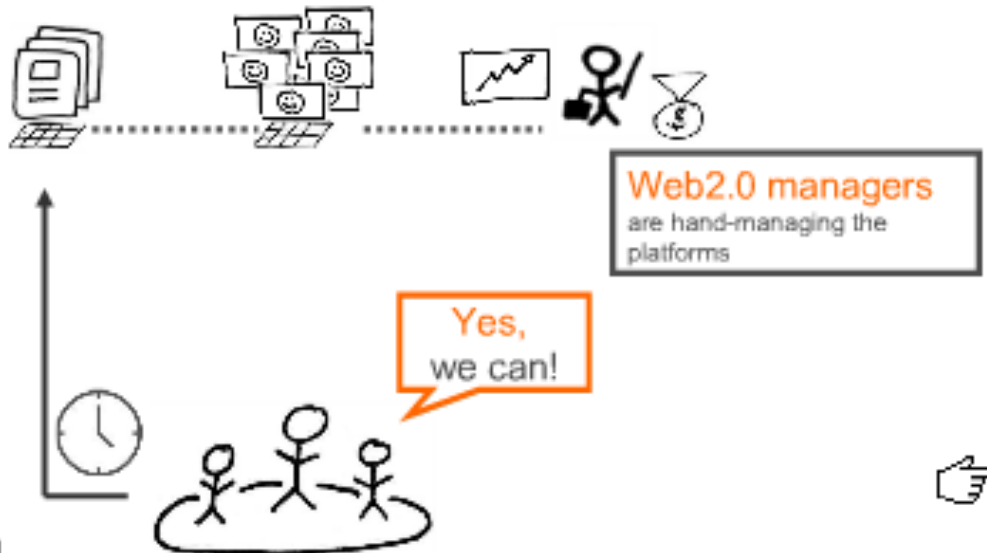
are hand-managing the platforms

Professional authors

can still publish, but are competing with – everyone!

Users = authors

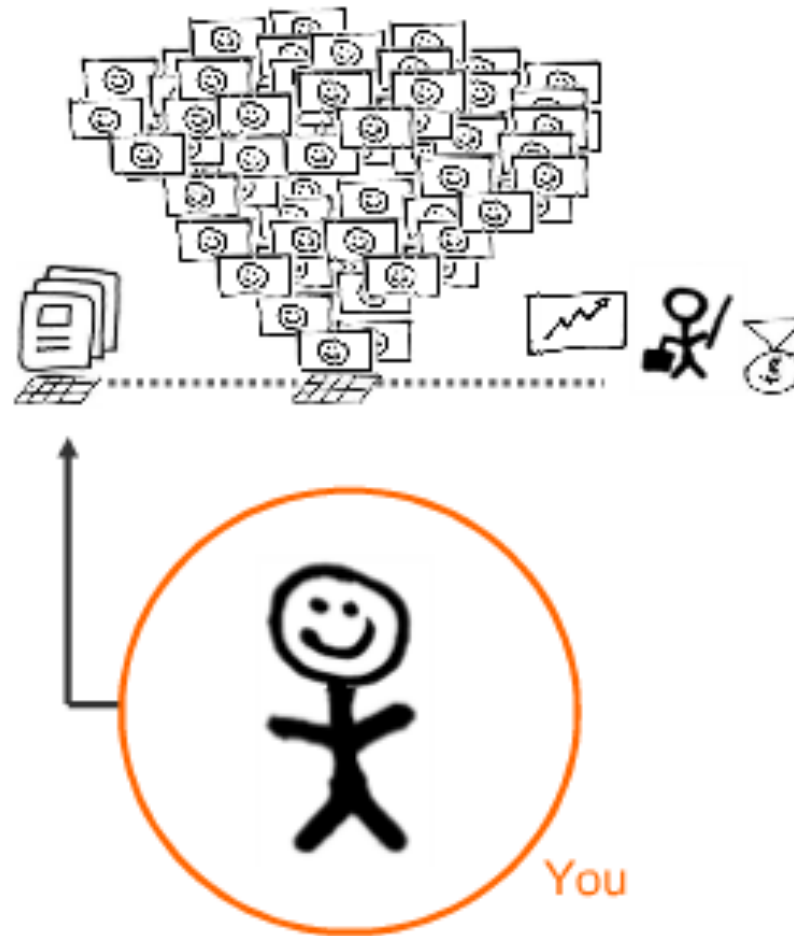
users generate content



UNIVERSITY OF
TORONTO

Web 2.0

YOU CAN PUBLISH!



Web 2.0



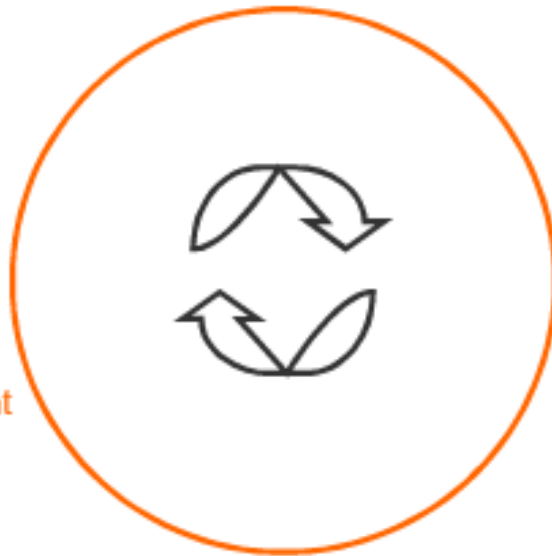
web2.0

web1.0

User generated content

We all serve each other

unlimited content



UNIVERSITY OF
TORONTO

Basically, web 2.0 is a
social change. The
technical part of the web
hasn't changed very much.



But, web 3.0 will be
driven by
technological changes.



Web 3.0 - the semantic
web - is about the
meaning of data.



Suppose, I am a
stamp collector...



Over the years I've
collected a lot of stamps.



Web 3.0

About every stamp, I made a document



Web 3.0

That's a lot of documents



How will I find that
specific stamp?



?



Web 3.0

Ah, of course !



Web 3.0

This is the web we have today: a huge collection of documents



Web 3.0

The words of all those documents are indexed. We can search for keywords.



Web 3.0

Now, suppose I search
for all red stamps...



What do I get ?



Red stamps
Stamps from Cambodia (Khmer Rouge)
Stamps from the Red Sea
Stamps from the 140 th anniversary of the Red Cross
Stamps with red dragons

Not very intelligent,
but how can a
computer know what I
mean?



Web 3.0

Answer: when we **structurally**
describe that a stamp is a
stamp and red is a color.

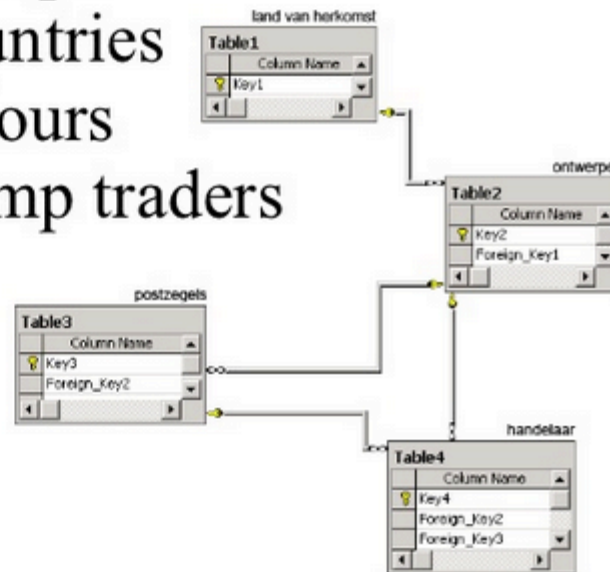


Describing data in a
structured way can
best be done in a
database.



Web 3.0

- A database with stamps
- A database with countries
- A database with colours
- A database with stamp traders



Web 3.0

Web 3.0 creates a big
collection of databases
which can be connected on
demand.



Web 3.0

Agreements are made on the structure of data and the way data is described. Where the data is located is irrelevant. Linking data is the power of web 3.0.



Web 3.0

So, “I want all the red stamps, designed in Europe, but used in the U.S.A., between 1980 and 1990” is a question that will get a **better answer** with web 3.0.



Web 3.0

This definition of web 3.0 is a ‘**narrow**’ definition. Like web 2.0, web 3.0 stands for a range of developments.



Web 3.0

A broader definition of web 3.0:

A fast broadband connection to the internet, always and everywhere .

Open source techniques and free data (Data as a Service)

Open identities

Software as a Service (e.g. Google docs)



References

1. <http://www2014.kr/asset/slide/Social+Spam,+Campaigns,+Misinformation+and+Crowdturfing-www2014-tutorial.pdf>
4. http://www.slideshare.net/freekbijl/web-30-explained-with-a-stamp?from=ss_embed