Information and Society-E2

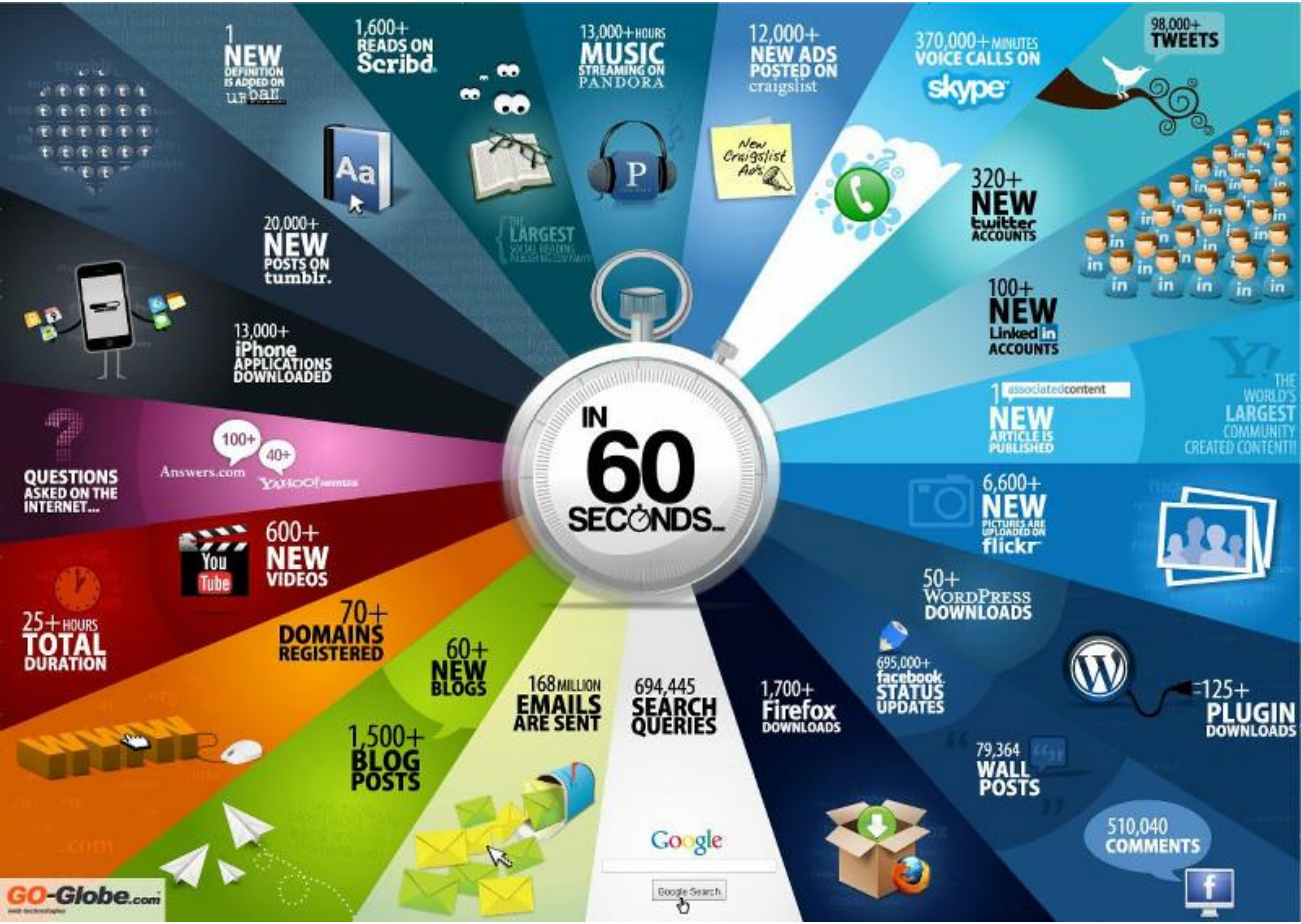
Social Media Analysis, Cloud Computing

Rafik Hadfi

Department of Social Informatics

Kyoto University

Email: rafik.hadfi@i.kyoto-u.ac.jp



IN 60 SECONDS...

1 NEW DEFINITION IS ADDED ON urban

1,600+ READS ON Scribd.

THE LARGEST SOCIAL READING PUBLISHING COMPANY

13,000+ HOURS MUSIC STREAMING ON PANDORA

12,000+ NEW ADS POSTED ON craigslist

370,000+ MINUTES VOICE CALLS ON skype

98,000+ TWEETS

320+ NEW twitter ACCOUNTS

100+ NEW LinkedIn ACCOUNTS

1 associatedcontent NEW ARTICLE IS PUBLISHED

THE WORLD'S LARGEST COMMUNITY CREATED CONTENT!!

6,600+ NEW PICTURES ARE UPLOADED ON flickr

50+ WORDPRESS DOWNLOADS

125+ PLUGIN DOWNLOADS

695,000+ facebook STATUS UPDATES

79,364 WALL POSTS

510,040 COMMENTS

1,700+ Firefox DOWNLOADS

694,445 SEARCH QUERIES

168 MILLION EMAILS ARE SENT

60+ NEW BLOGS

1,500+ BLOG POSTS

70+ DOMAINS REGISTERED

600+ NEW VIDEOS

25+ HOURS TOTAL DURATION

QUESTIONS ASKED ON THE INTERNET...

100+ Answers.com

40+ YAHOO! ANSWERS

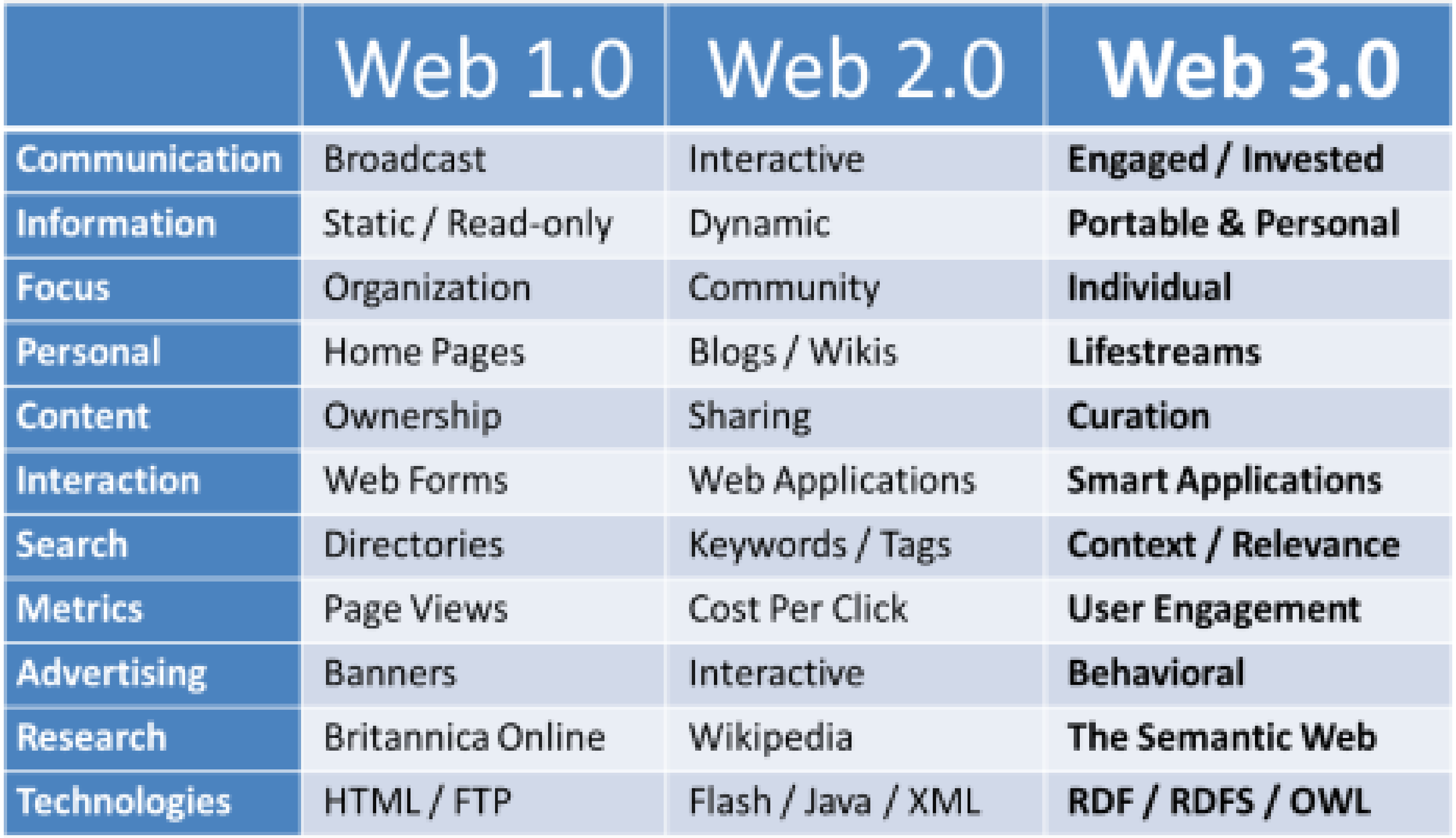
13,000+ iPhone APPLICATIONS DOWNLOADED

20,000+ NEW POSTS ON tumblr.

GO-Globe.com

What is Social Media?

Web 1.0: 1995 – 2005 (Homepage)
Web 2.0: 2005 – (Social network services, video sharing, photo sharing, …)
Web 3.0: ?



	Web 1.0	Web 2.0	Web 3.0
Communication	Broadcast	Interactive	Engaged / Invested
Information	Static / Read-only	Dynamic	Portable & Personal
Focus	Organization	Community	Individual
Personal	Home Pages	Blogs / Wikis	Lifestreams
Content	Ownership	Sharing	Curation
Interaction	Web Forms	Web Applications	Smart Applications
Search	Directories	Keywords / Tags	Context / Relevance
Metrics	Page Views	Cost Per Click	User Engagement
Advertising	Banners	Interactive	Behavioral
Research	Britannica Online	Wikipedia	The Semantic Web
Technologies	HTML / FTP	Flash / Java / XML	RDF / RDFS / OWL

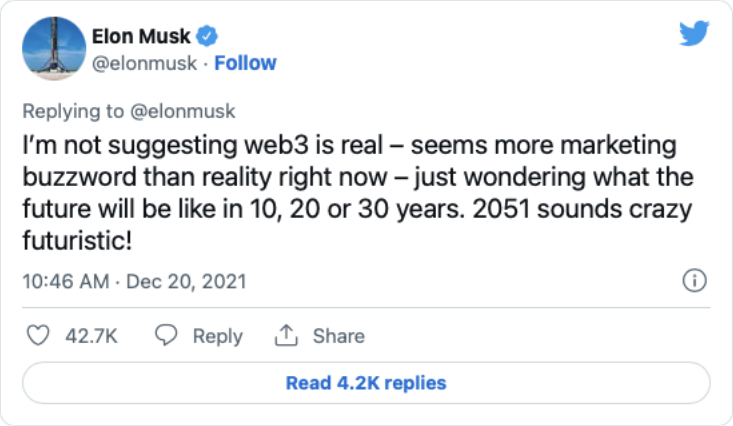
The Web 3.0: The Web Transition Is Coming
https://hackernoon.com/the-web-3-0-the-web-transition-is-coming-892108fd0d

Recent Version of Web 3.0

Web 3 (also known as **Web 3.0** and sometimes stylized as **web3**) is an idea for a new iteration of the World Wide Web which incorporates concepts such as decentralization, in the form distributed ledger such as blockchain technologies, and token-based economics.

Some technologists and journalists have contrasted it with Web 2.0, wherein they say data and content are centralized in a small group of companies sometimes referred to as "Big Tech".

The term "Web3" was coined in 2014 by Ethereum co-founder Gavin Wood, and the idea gained interest in 2021 from cryptocurrency enthusiasts, large technology companies, and venture capital firms.



Elon Musk
@elonmusk · Follow

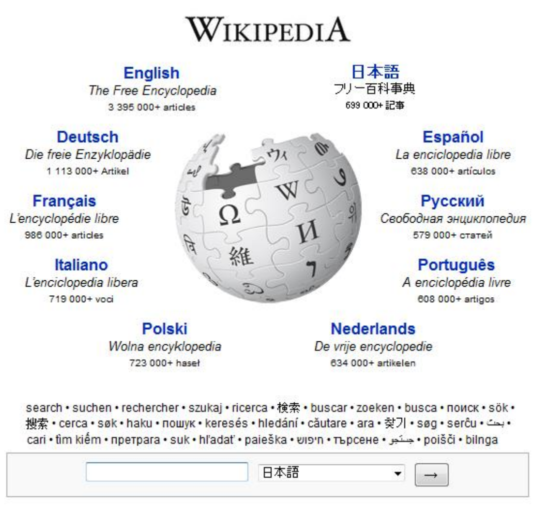
Replying to @elonmusk

I'm not suggesting web3 is real – seems more marketing buzzword than reality right now – just wondering what the future will be like in 10, 20 or 30 years. 2051 sounds crazy futuristic!

10:46 AM · Dec 20, 2021

https://en.wikipedia.org/wiki/Web3

What is Social Media?


WIKIPEDIA

English
The Free Encyclopedia
3 395 000+ articles

日本語
フリー百科事典
699 000+ 記事

Deutsch
Die freie Enzyklopädie
1 113 000+ Artikel

Español
La enciclopedia libre
638 000+ artículos

Français
L'encyclopédie libre
986 000+ articles

Русский
Свободная энциклопедия
579 000+ статей

Italiano
L'enciclopedia libera
719 000+ voci

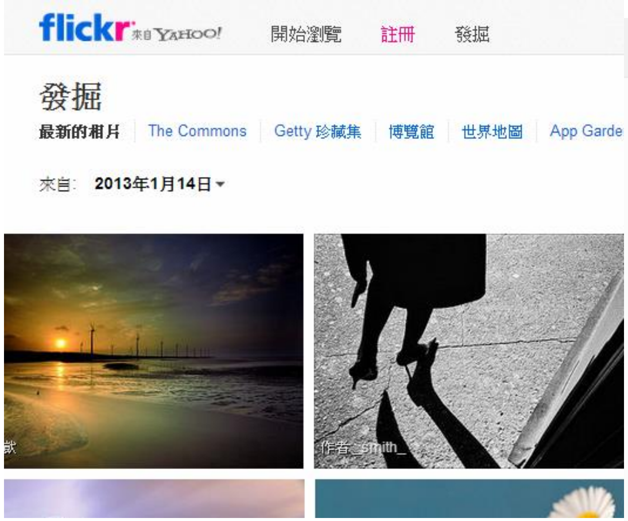
Português
A enciclopédia livre
608 000+ artigos

Polski
Wolna encyklopedia
723 000+ haseł

Nederlands
De vrije encyclopedie
634 000+ artikelen

search • suchen • rechercher • szukaj • ricerca • 検索 • buscar • zoeken • busca • поиск • sök • 搜索 • cerca • søk • haku • пошук • keresés • hledání • căutare • ara • 찾기 • søg • serĉu • بحث • cari • tìm kiếm • претрара • suk • hľadať • paieška • חיפוש • търсене • جستجو • poišči • bilnga

日本語 →

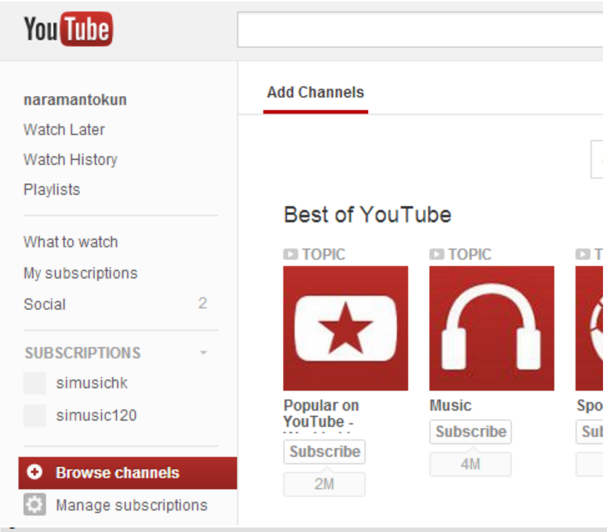

flickr 來自 YAHOO! 開始瀏覽 註冊 發掘

發掘

最新的相片 The Commons Getty 珍藏集 博覽館 世界地圖 App Garde

來自: 2013年1月14日

作者 _smith_


YouTube

naramantokun
Watch Later
Watch History
Playlists

What to watch
My subscriptions
Social 2

SUBSCRIPTIONS
simusichk
simusic120

Browse channels
Manage subscriptions

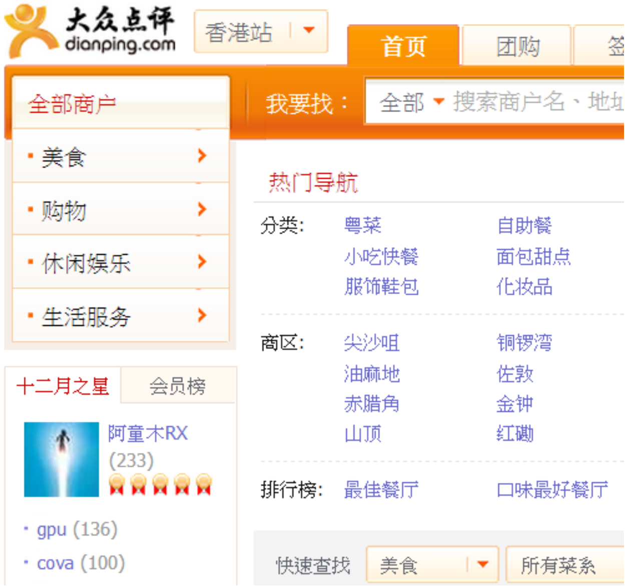
Add Channels

Best of YouTube

TOPIC
Popular on YouTube -
Subscribe
2M

TOPIC
Music
Subscribe
4M

T
Spo
Sul


大众点评 dianping.com 香港站

首页 团购 签

全部商户 我要找： 全部 搜索商户名、地址

- 美食
- 购物
- 休闲娱乐
- 生活服务

热门导航

分类： 粤菜 自助餐 小吃快餐 面包甜点 服饰鞋包 化妆品

商区： 尖沙咀 铜锣湾 油麻地 佐敦 赤腊角 金钟 山顶 红磡

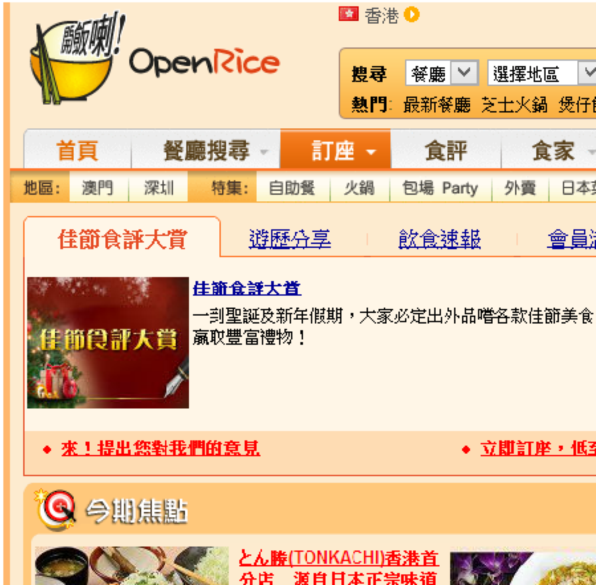
排行榜： 最佳餐厅 口味最好餐厅

十二月之星 会员榜

阿童木RX
(233)

- gpu (136)
- cova (100)

快速查找 美食 所有菜系


香港

開飯喇! OpenRice

搜尋 餐廳 選擇地區
熱門: 最新餐廳 芝士火鍋 煲仔

首頁 餐廳搜尋 訂座 食評 食家

地區: 澳門 深圳 特集: 自助餐 火鍋 包場 Party 外賣 日本

佳節食評大賞 遊歷分享 飲食速報 會員

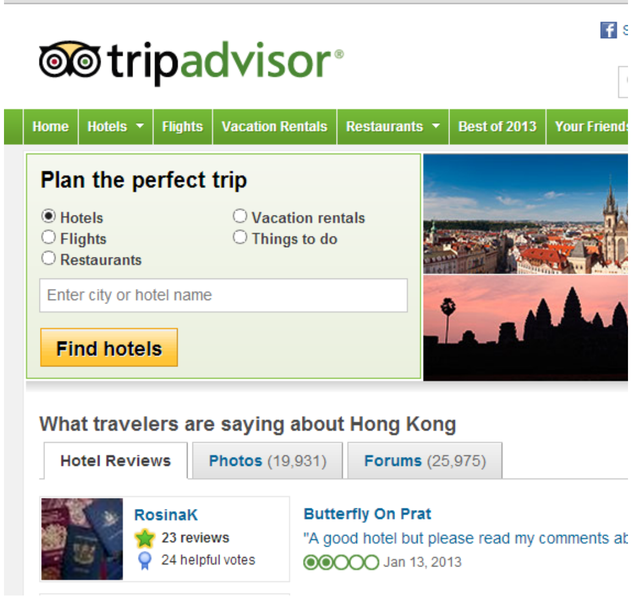
佳節食評大賞

一到聖誕及新年假期，大家必定出外品嚐各款佳節美食
贏取豐富禮物！

- 來！提出您對我們的意見
- 立即訂座，低至

今期焦點

とん勝(TONKACHI)香港首分店 源自日本正宗味道


tripadvisor

Home | Hotels | Flights | Vacation Rentals | Restaurants | Best of 2013 | Your Friends

Plan the perfect trip

● Hotels ○ Vacation rentals
○ Flights ○ Things to do
○ Restaurants

Enter city or hotel name

Find hotels

What travelers are saying about Hong Kong

Hotel Reviews | Photos (19,931) | Forums (25,975)

RosinaK
23 reviews
24 helpful votes

Butterfly On Prat
"A good hotel but please read my comments ab
Jan 13, 2013

What is Social Media?

Financial Time’s Definition

“Social media refers to the internet and mobile technology based channels of communication in which people share content with each other. Examples are social networking sites such as Facebook and Twitter.

In contrast to “social” media, earlier media channels made a clear distinction between a producer and a consumer of content. Since, for example, social media easily allows for a viewer of a video to share the same content with others, the boundary between consumers and producers is blurred.”

Social Media Sites/Services

- **Social Networking & Microblogs**
  Facebook, MySpace, Mixi, Twitter, Weibo

- **Video Sharing**
  Youtube, Niconico, Youku

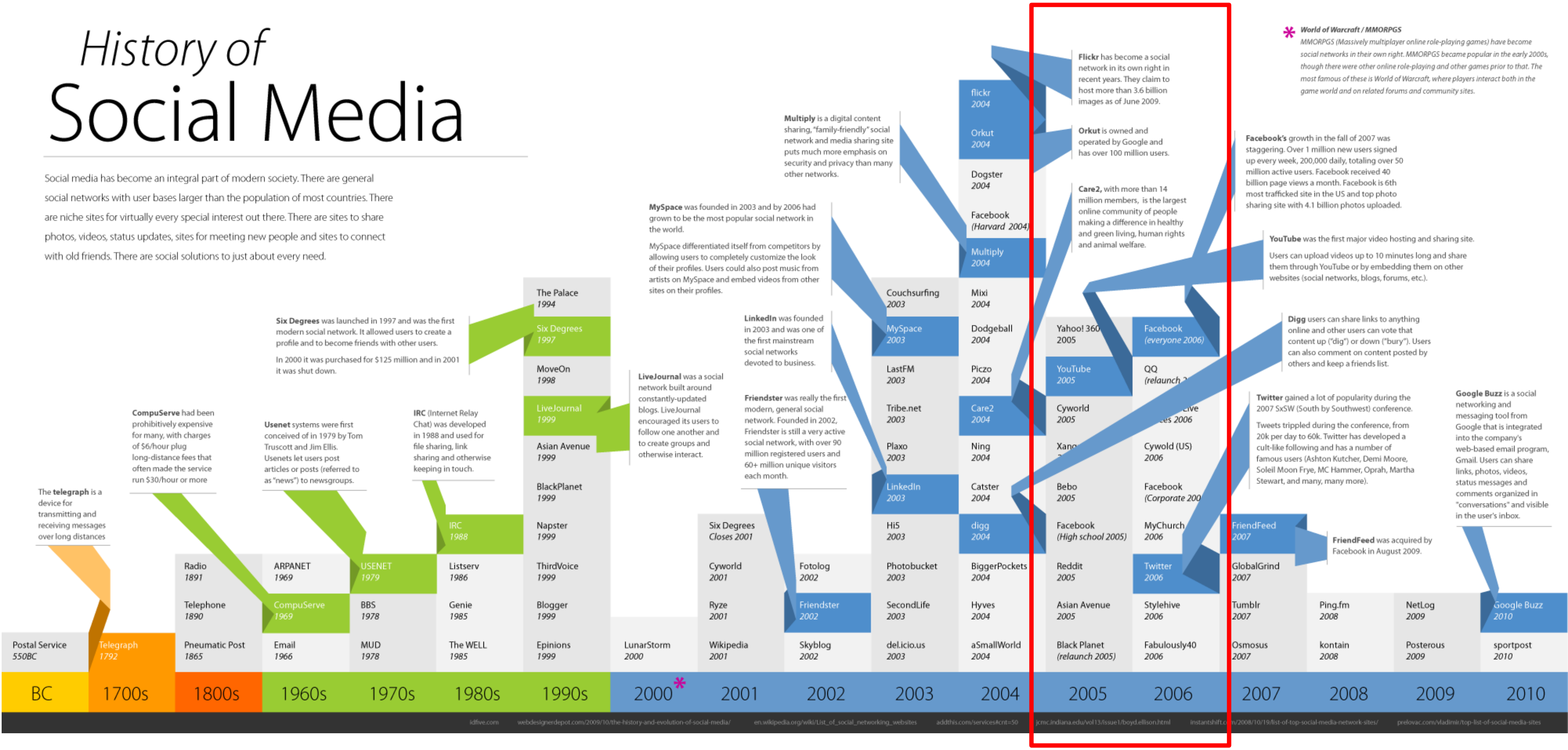
- **Photo Sharing**
  Pinterest, Flickr, Instagram

- **Consumer Reviews**
  Amazon, Rakuten, Taobao

- **Others**
  Wikipedia, Foursquare, …

Types of Social Media

- The variety of social media makes it hard to categorize them
- Marketing and social media experts broadly agree on 13 types
  1. blogs,
  2. business networks,
  3. collaborative projects,
  4. enterprise social networks,
  5. forums,
  6. microblogs,
  7. photo sharing,
  8. products/services review,
  9. social bookmarking,
  10. social gaming,
  11. social networks,
  12. video sharing,
  13. virtual worlds



History of Social Media

Social media has become an integral part of modern society. There are general social networks with user bases larger than the population of most countries. There are niche sites for virtually every special interest out there. There are sites to share photos, videos, status updates, sites for meeting new people and sites to connect with old friends. There are social solutions to just about every need.

The **telegraph** is a device for transmitting and receiving messages over long distances

CompuServe had been prohibitively expensive for many, with charges of $6/hour plug long-distance fees that often made the service run $30/hour or more

Usenet systems were first conceived of in 1979 by Tom Truscott and Jim Ellis. Usenets let users post articles or posts (referred to as "news") to newsgroups.

IRC (Internet Relay Chat) was developed in 1988 and used for file sharing, link sharing and otherwise keeping in touch.

Six Degrees was launched in 1997 and was the first modern social network. It allowed users to create a profile and to become friends with other users.

In 2000 it was purchased for $125 million and in 2001 it was shut down.

LiveJournal was a social network built around constantly-updated blogs. LiveJournal encouraged its users to follow one another and to create groups and otherwise interact.

MySpace was founded in 2003 and by 2006 had grown to be the most popular social network in the world.

MySpace differentiated itself from competitors by allowing users to completely customize the look of their profiles. Users could also post music from artists on MySpace and embed videos from other sites on their profiles.

LinkedIn was founded in 2003 and was one of the first mainstream social networks devoted to business.

Friendster was really the first modern, general social network. Founded in 2002, Friendster is still a very active social network, with over 90 million registered users and 60+ million unique visitors each month.

Multiply is a digital content sharing, "family-friendly" social network and media sharing site puts much more emphasis on security and privacy than many other networks.

Flickr has become a social network in its own right in recent years. They claim to host more than 3.6 billion images as of June 2009.

Orkut is owned and operated by Google and has over 100 million users.

Care2, with more than 14 million members, is the largest online community of people making a difference in healthy and green living, human rights and animal welfare.

* ***World of Warcraft / MMORPGS***

MMORPGS (Massively multiplayer online role-playing games) have become social networks in their own right. MMORPGS became popular in the early 2000s, though there were other online role-playing and other games prior to that. The most famous of these is World of Warcraft, where players interact both in the game world and on related forums and community sites.

Facebook's growth in the fall of 2007 was staggering. Over 1 million new users signed up every week, 200,000 daily, totaling over 50 million active users. Facebook received 40 billion page views a month. Facebook is 6th most trafficked site in the US and top photo sharing site with 4.1 billion photos uploaded.

YouTube was the first major video hosting and sharing site.

Users can upload videos up to 10 minutes long and share them through YouTube or by embedding them on other websites (social networks, blogs, forums, etc.).

Digg users can share links to anything online and other users can vote that content up ("dig") or down ("bury"). Users can also comment on content posted by others and keep a friends list.

Twitter gained a lot of popularity during the 2007 SxSW (South by Southwest) conference.

Tweets trippled during the conference, from 20k per day to 60k. Twitter has developed a cult-like following and has a number of famous users (Ashton Kutcher, Demi Moore, Soleil Moon Frye, MC Hammer, Oprah, Martha Stewart, and many, many more).

Google Buzz is a social networking and messaging tool from Google that is integrated into the company's web-based email program, Gmail. Users can share links, photos, videos, status messages and comments organized in "conversations" and visible in the user's inbox.

FriendFeed was acquired by Facebook in August 2009.

idfive.com webdesignerdepot.com/2009/10/the-history-and-evolution-of-social-media/ en.wikipedia.org/wiki/List_of_social_networking_websites addthis.com/services#cnt=50 jcmc.indiana.edu/vol13/issue1/boyd.ellison.html instantshift.com/2008/10/19/list-of-top-social-media-network-sites/ prelovac.com/vladimir/top-list-of-social-media-sites

http://www.flickr.com/photos/pictopedia/5200988483/
http://www.ritholtz.com/blog/wp-content/uploads/2010/12/socialMediaTL_05.png

Social Media vs. Media

- Larger variance in quality (professional vs. amateurs)
- Cover more niche areas (different social media publishers may have different interests)
- Immediacy – (e.g. contributed by local people, or any one who happens to be at the scene)
- Others: reach, frequency, accessibility, permanence, etc.

Ref:

- http://en.wikipedia.org/wiki/Social_media
- Nigel Morgan; Graham Jones; Ant Hodges. "Social Media". The Complete Guide to Social Media From The Social Media Guys.

Why is Social Media Important?

- **A global phenomenon**: Internet and mobile users spend more and more time on social media sites and services

- Social media sites are platforms where users
  1. **Interact** and **share** information with each other
  2. Form **groups** and **communities**, in which members share some characteristics (political views, hobbies, religion, opinions, interests, etc.)
  3. Express **opinions** and discuss them
  4. Seek **advice** and **recommendations** on various decision making tasks
  5. Etc.

What is Social Media Analysis?

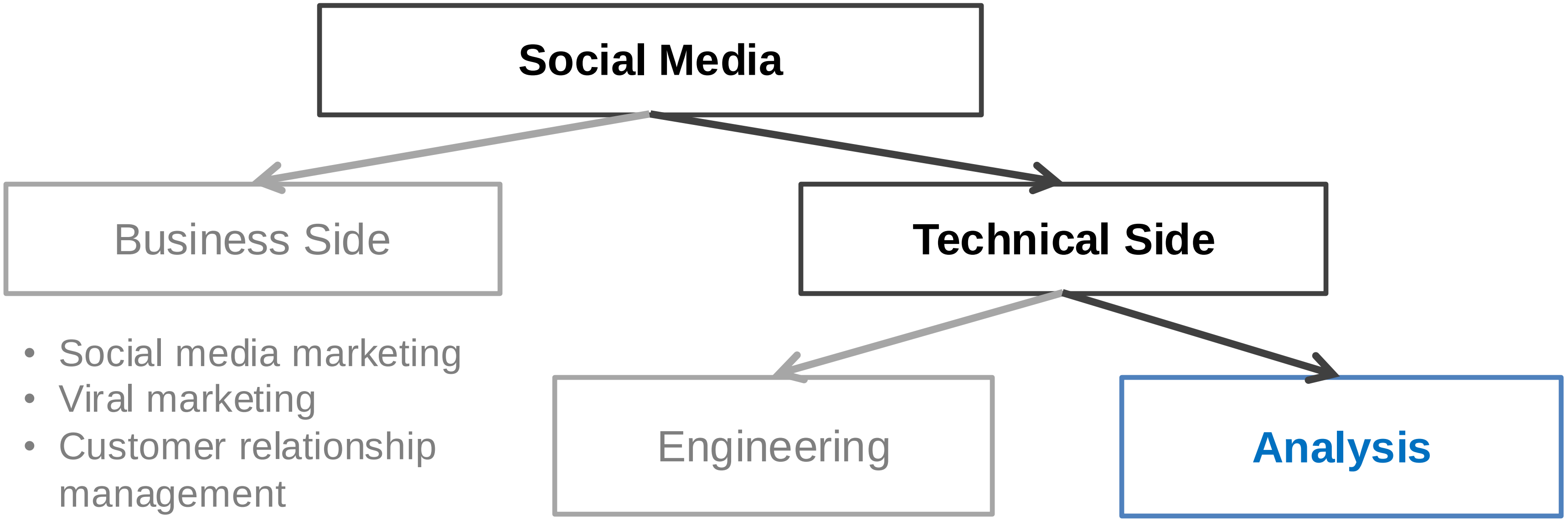
- Social media involve current affairs, people's opinions and feelings, reviews of consumer products and services, etc.
- A new channel for us to understand human behaviour, user preferences and reactions, trends and problems, etc.

- **Social Media Analysis**
  - To summarize and extract information from a large amount of data collected in a social media service, using statistical and mathematical techniques and algorithms

Why Social Media Analysis?

- Understand individual and social behaviour
- Understand user preferences and interests
- Learn about the opinions and sentiment of the people
- Understand how people talk about a person, a company or a brand on the Internet
- Understand how ideas and information propagates in a social network
- Build better social media services to support user interactions and information sharing
- Make predictions of various behaviour and events
- …

Studying Social Media



Social Media

Business Side

Technical Side

Engineering

Analysis

Business Side:

- Social media marketing
- Viral marketing
- Customer relationship management
- Customer services
- …

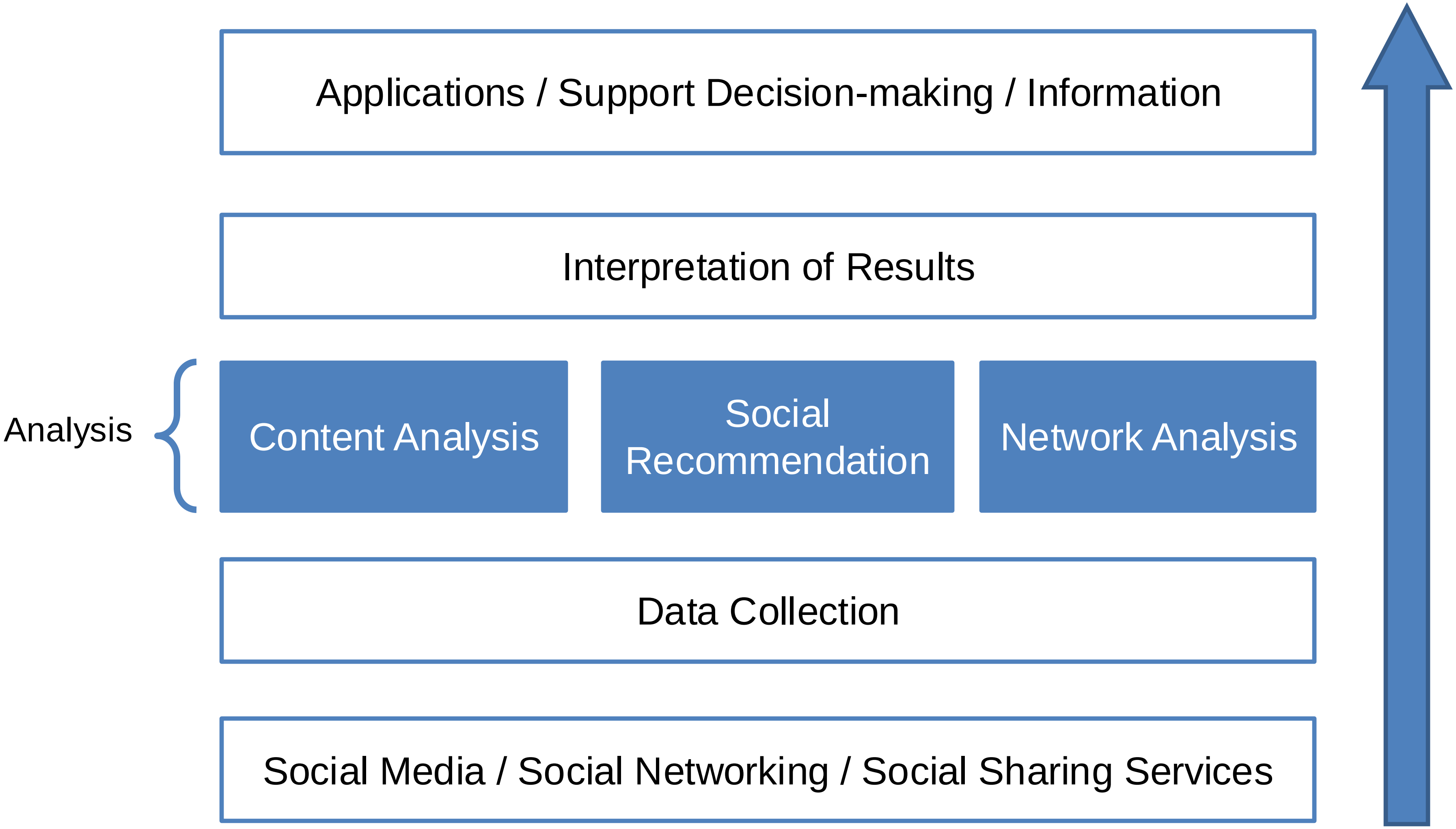
Engineering:

- Cloud computing
- Data centre operation
- Big data management
- Load balancing
- Content delivery network
- Mobile apps
- …

Analysis:

- Natural language Processing (NLP)
- Sentiment analysis & opinion mining algorithms
- Social recommendation systems
- Information propagation and influence
- …

Workflow of Social Media Analysis



Applications / Support Decision-making / Information

Interpretation of Results

Analysis

Content Analysis

Social Recommendation

Network Analysis

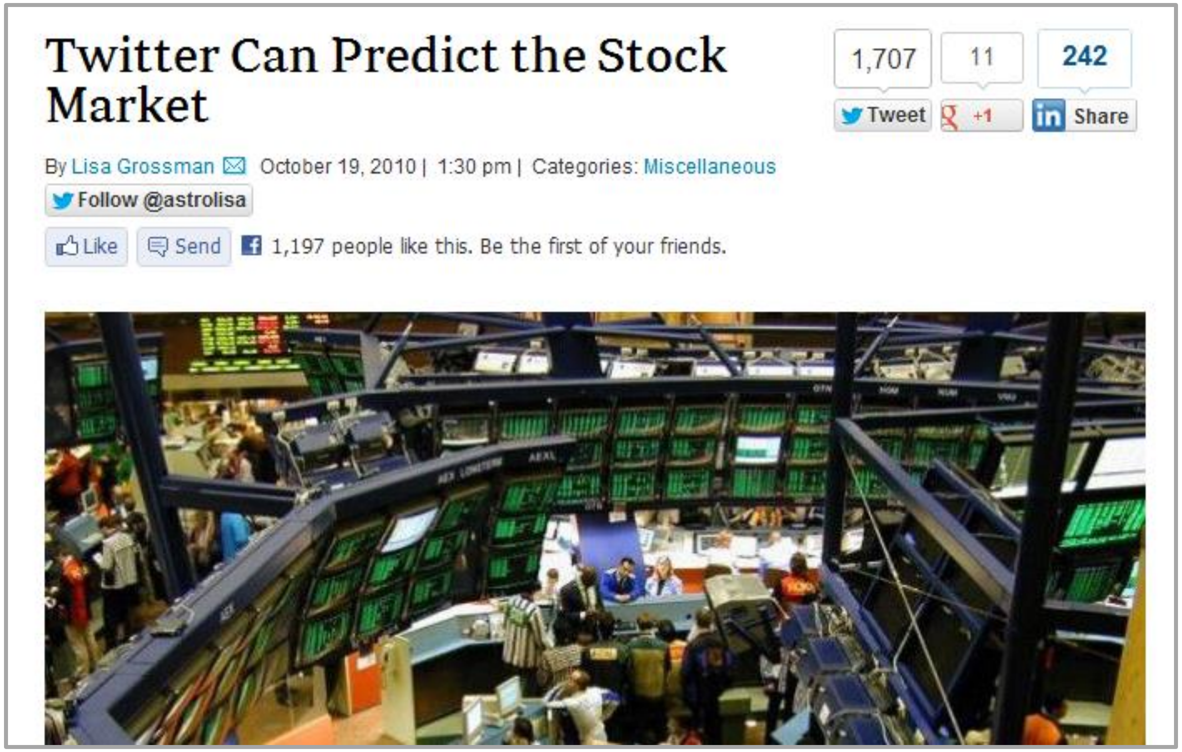
Data Collection

Social Media / Social Networking / Social Sharing Services

Some Research Questions in Social Media Analysis

Sentiment Analysis / Opinion Mining

- Understand what the crowd is thinking or feeling
- Understand what opinions are expressed towards products, companies, individuals, etc.



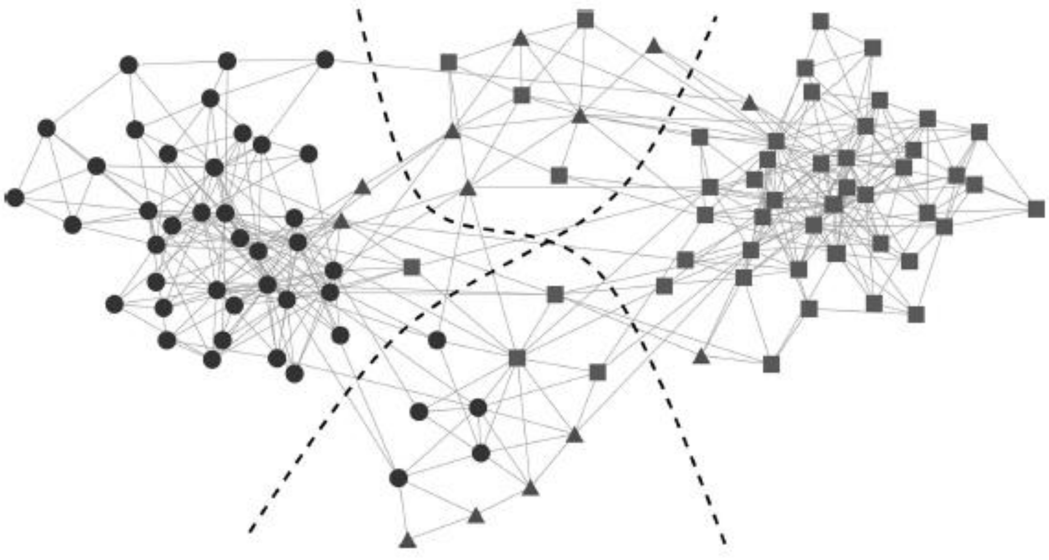
Twitter Can Predict the Stock Market

By Lisa Grossman ✉ October 19, 2010 | 1:30 pm | Categories: Miscellaneous

1,197 people like this. Be the first of your friends.

Social Network & Network Science

- How can we understand the roles of different persons given a social network?
- How to discover communities of users? Can one user belong to multiple communities/groups?

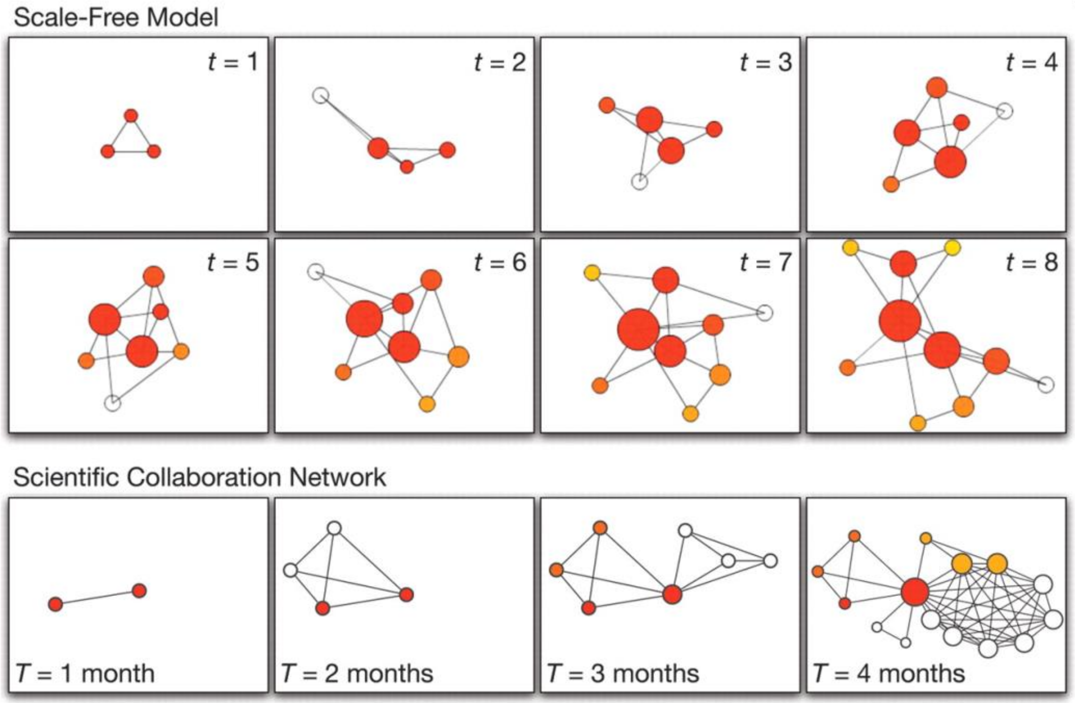


Kreb's network of books on American politics. Taken from "Modularity and community structure in networks" (Newman, 2006).

Vertices represent books and edges represent the fact that the two books were read by the same reader.

Social Network & Network Science

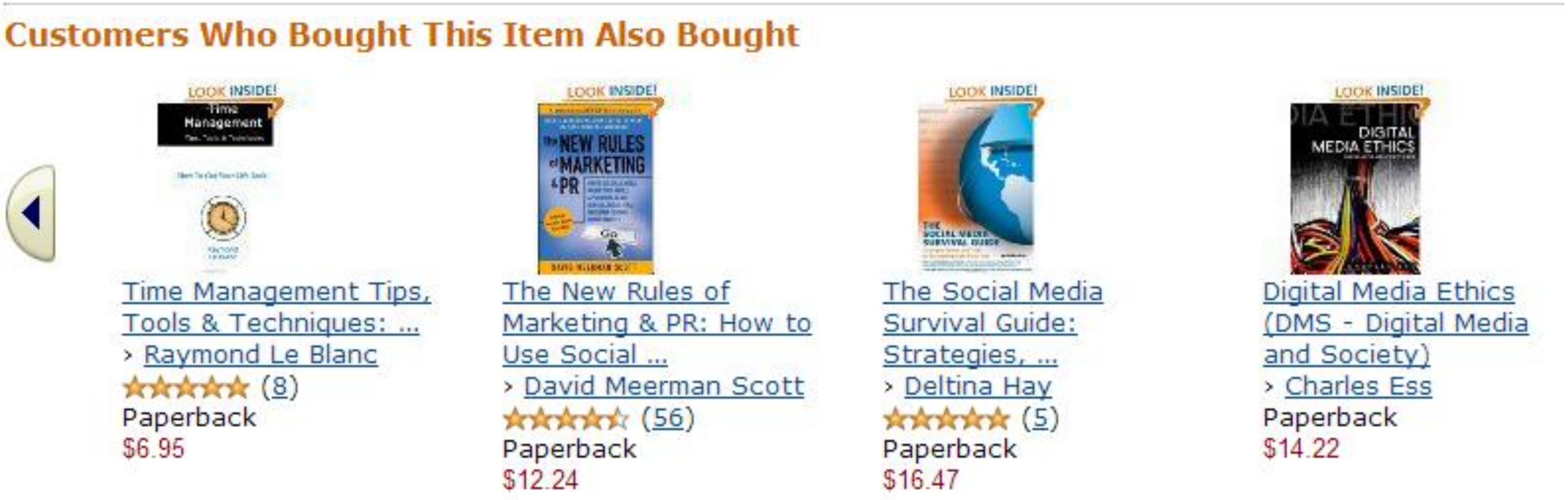
- How does information propagate in a network? Who is responsible of spreading the news or rumours?
- How can we predict or model the growth of a network? (small-world and scale-free networks)



Growth of a scale-free network. From "Scale-Free Networks: A Decade and Beyond" (Barabási 2009)

Search, Navigate, Explore & Recommend

- How can we improve searching by leveraging user-generated content?
- How can we assist navigation by analysing usage patterns? E.g. People read this page also read that page



Customers Who Bought This Item Also Bought

Time Management Tips, Tools & Techniques: ...
› Raymond Le Blanc
★★★★★ (8)
Paperback
$6.95

The New Rules of Marketing & PR: How to Use Social ...
› David Meerman Scott
★★★★½ (56)
Paperback
$12.24

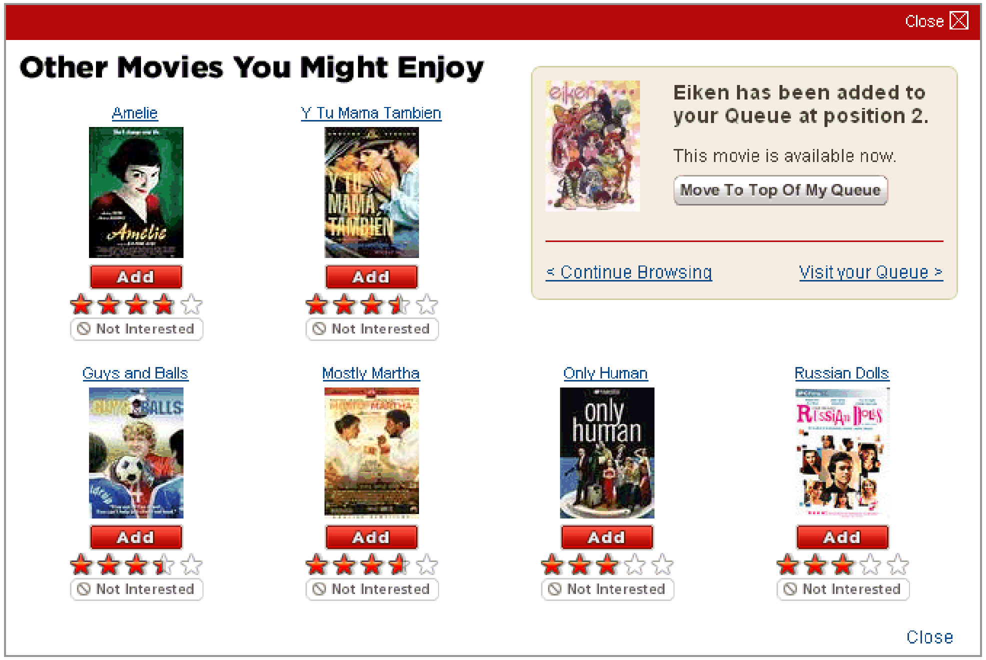
The Social Media Survival Guide: Strategies, ...
› Deltina Hay
★★★★★ (5)
Paperback
$16.47

Digital Media Ethics (DMS - Digital Media and Society)
› Charles Ess
Paperback
$14.22

Related books on Amazon (http://www.amazon.com/)

Making Recommendations

- How can we make recommendations?
  - ➢ Based on user profiles, history of user activities
  - ➢ Based on item attributes
  - ➢ Based on social networks



Other Movies You Might Enjoy

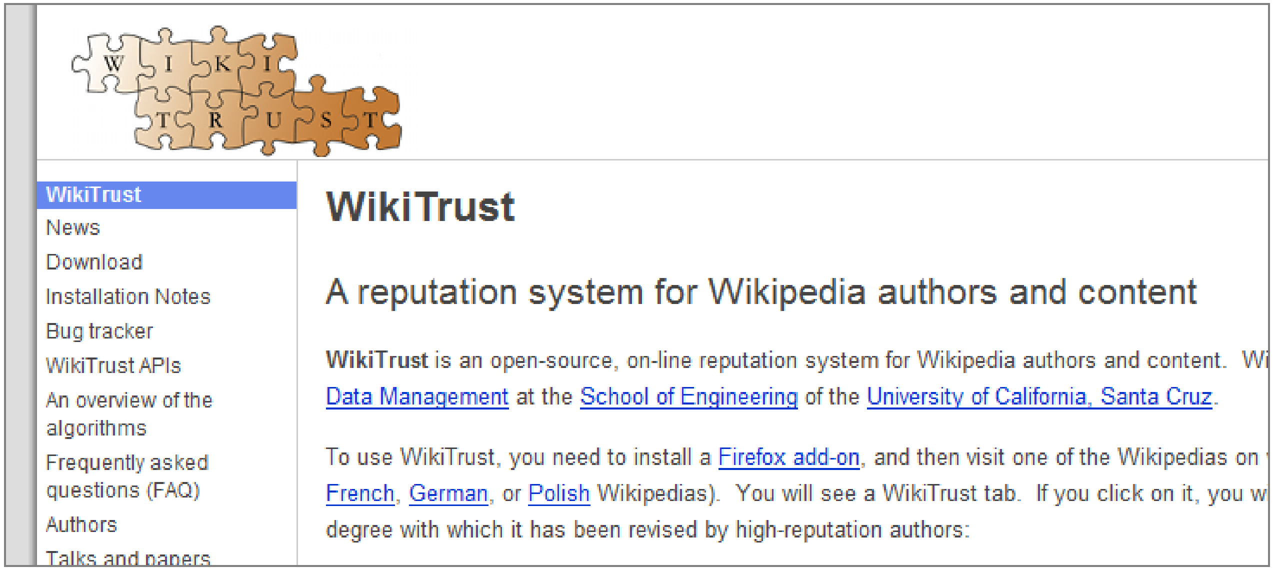
Eiken has been added to your Queue at position 2.

This movie is available now.

Movie recommendations on
http://www.netflix.com/

Credibility, Trust, Expertise

- Not all users are credible, how do we choose what to trust, who to trust?
- Shall we trust Wikipedia, given that it is written collaboratively by ordinary users?



WikiTrust

A reputation system for Wikipedia authors and content

WikiTrust is an open-source, on-line reputation system for Wikipedia authors and content. Wi
Data Management at the School of Engineering of the University of California, Santa Cruz.

To use WikiTrust, you need to install a Firefox add-on, and then visit one of the Wikipedias on
French, German, or Polish Wikipedias). You will see a WikiTrust tab. If you click on it, you w
degree with which it has been revised by high-reputation authors:

A reputation system for Wikipedia
(http://wikitrust.soe.ucsc.edu/home)

Quality of Shared Content

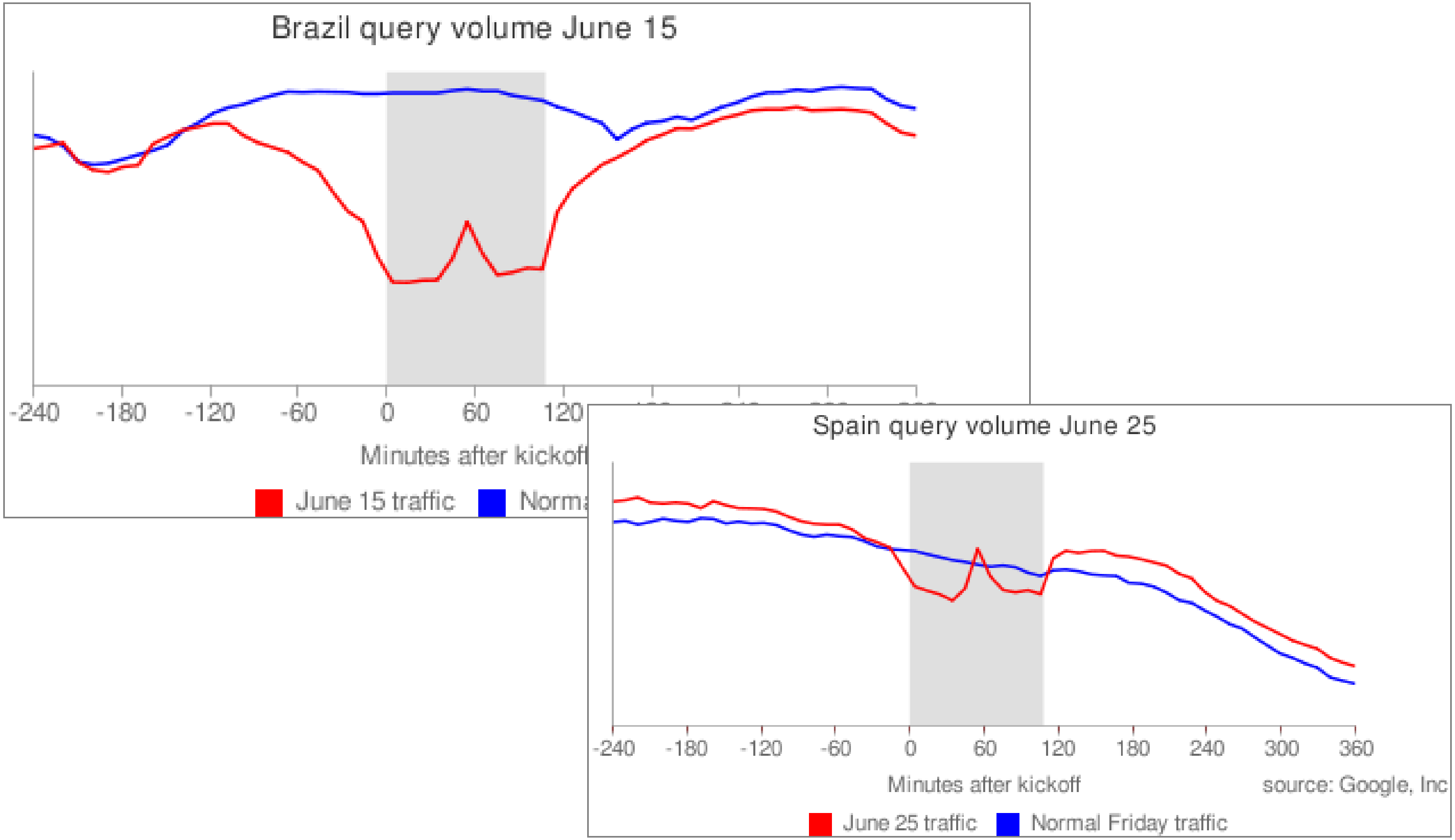
- How can we assess the quality of shared content on the Web?
- What is the relationship between quality of content vs. credibility of users?
- How can we avoid the ‘Tyranny of the Majority’?

Social Dynamics, Group Evolution

- How do users interact with each other? What are the drives behind the behaviour of the users?
- What motivates people to contribute to a Web site?
- How can we model the dynamics or evolution/growth of a group/community? What are the factors?
- What kinds of environment facilitate users to collaborate with each other on the Web?

Let’s look at some examples of social media analysis

Search Log Analysis at Google

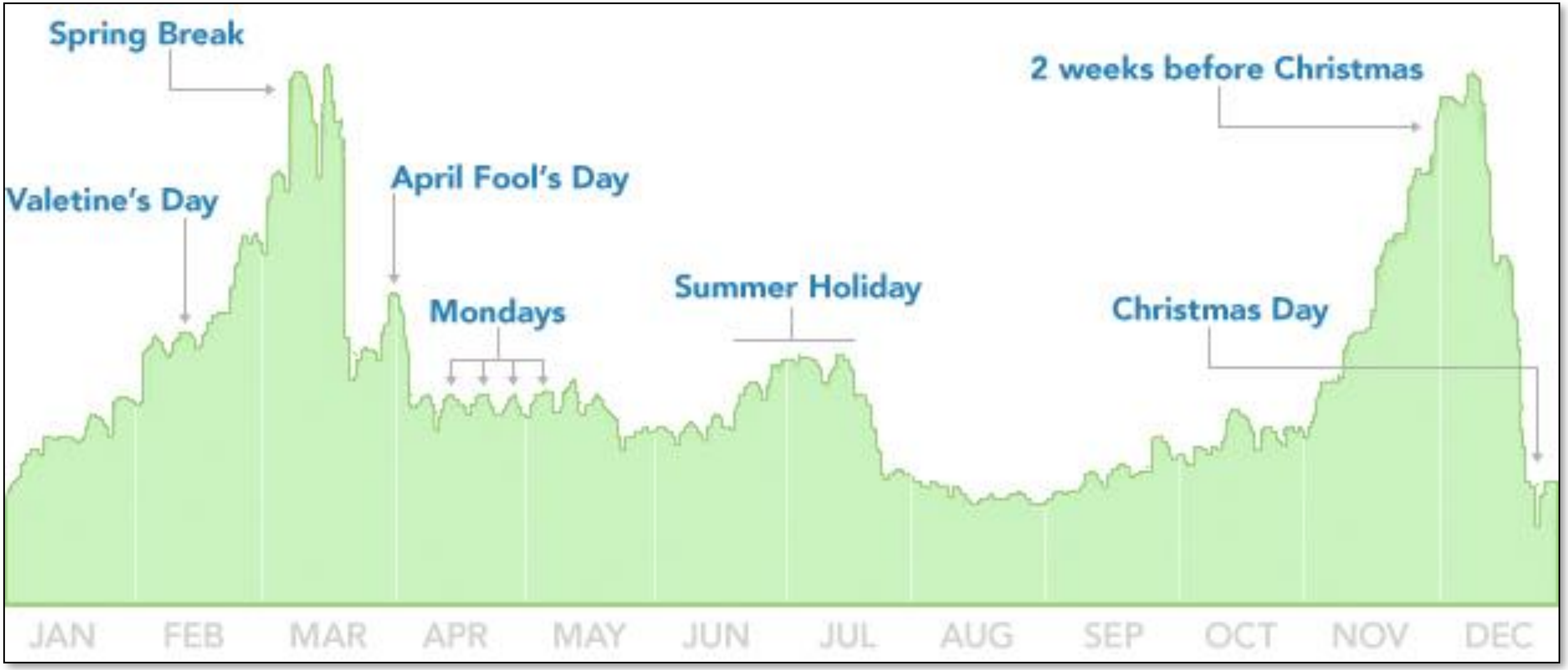


Brazil query volume June 15

Spain query volume June 25

source: Google, Inc

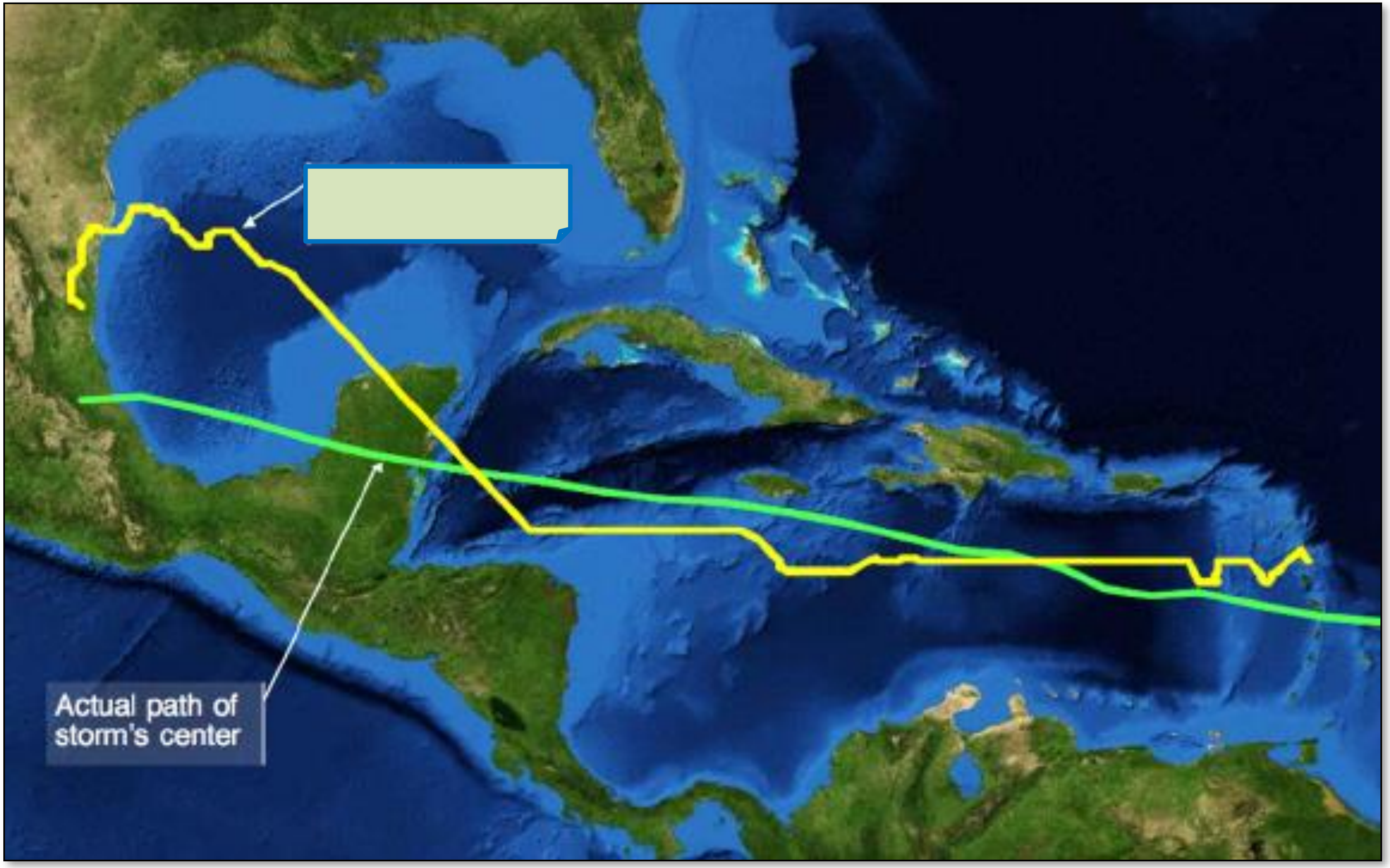
Visualization of Status Changes



David McCandless

http://www.thehighdefinite.com/2010/11

Movement



Lars Backstorm et al: Spatial Variation in Search Engine Queries, WWW 2008

Cloud Computing

Jimmy Lin and Dhris Dyer. Data-Intensive Text Processing with MapReduce.
Morgan & Claypool Synthesis Lectures on Human Language Technologies. 2010

Cloud Computing - Definition

- Cloud computing - culmination of many technologies such as grid computing, utility computing, SOA, Web 2.0, etc.
- Cloud computing
  - *"a computing paradigm, where a large pool of systems are connected in private or public networks, to provide dynamically scalable infrastructure for application, data and file storage. With the advent of this technology, the cost of computation, application hosting, content storage and delivery is reduced significantly"* [1]
- Key characteristics
  - ability to scale and provision computing power dynamically in a cost efficient way
  - ability of the consumer (e.g., end user, organization or IT staff) to make the most of that power without having to manage the underlying complexity of technology

[1] Torry Harris, Introduction to Cloud Computing

Cloud Computing Background

- Data in size of Terabytes and Petabytes become relatively common
  - E.g. Google was processing 20PB a day with MapReduce in 2008, eBay has more than 2PB of user data, FB over 2.5PB, Large Hadron Collider over 15PB data per year, Sloan Digital Sky Survey project with telescope producing 0.5PB archive monthly
- Large computational problems resulting from accumulation of huge datasets, e.g. friend suggest at Facebook or ad placement in Gmail
- Larger datasets or larger corpora lead to higher precision and performance of different algorithms/features choice converges along with data amount increase [1] (growing evidence that data may be more important than algorithms)
- Moore Law no longer applies - causes fundamental shift in software creation
- Our ability of storing data is overwhelming our ability to process it

[1] Michele Banko and Eric Brill. Scaling to very very large corpora for natural language disambiguation. In Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics (ACL 2001), pages 26-33, Toulouse, France, 2001

SaaS, PaaS, IaaS

- SaaS – Software as a Service
  - Complete application offered to customer, as a service on demand
  - A single instance of the service runs on the cloud & multiple end users are serviced. Applications are delivered through browser, and multiple customers can access them from various locations. It has become the most common form of cloud computing
  - Example: Gmail, GoogleDocs, Salesforce
- PaaS – Platform as a Service
  - Layer of software or development environment is encapsulated and offered as a service, upon which other higher levels of service can be built. Customer is free to build his own applications that will run on the providers infrastructure
  - Example: Google App Engine (maintains infrastructure such as storage layer and programming environment including backup, upgrade, patch, etc.), LAMP platform (Linux, Apache, MySql and PHP),
- IaaS – Infrastructure as a Service
  - Basic storage and computing capabilities offered as standardized services over the network. Servers, storage systems, networking equipment, data centre space etc. are pooled and made available to handle workloads. Computational resources are essentially rented - turning what was previously a need to purchase products (hardware, software and network bandwidth) into a service
  - Examples: Amazon’s EC2, Rackspace
- XaaS – Everything as a Service
- HaaS – Human as a Service (Crowd computing) → Next week

Characteristics of Cloud Computing (1/2)

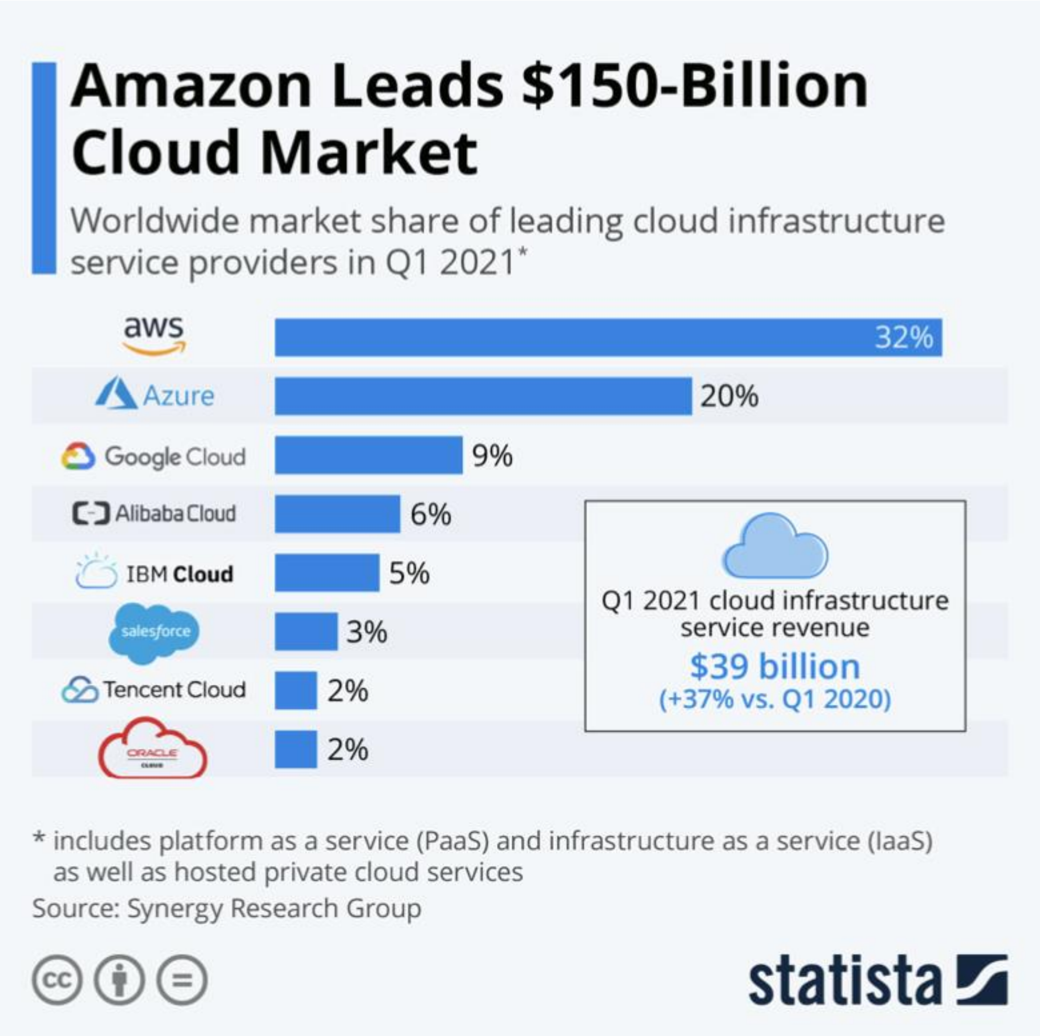
- Scale "out" not "up"
  - Large number of low-end servers preferred over small number of high-end ones
  - Cost of machines does not scale linearly – e.g., a machine with twice as many processors is often significantly more than twice expensive (*amount of data grows faster than the fall of memory/processor price)*
  - Non-linearity between load and power draw - lightly-loaded server is much less efficient than a heavily-loaded one (e.g., a server at 10% utilization may draw significantly more than half as much power as a server at 100% utilization level)
  - Communication overhead is small for large cluster of low end servers when compared to small cluster of high end mainframes
- Assumption of common failures
  - Failures are not only inevitable but common, and server mail fail anytime
  - With reliable servers a 10,000-server cluster would experience 10 failures a day (3 years of MTBF)
  - Infrastructure should be organized in a way to counteract this problem
    - E.g. slow decrease of computing power in case of many failures, seamless rejoining the cloud by repaired servers
- Move processing to data
  - In high performance computing (e.g., climate simulations) it is common to have processing and storage nodes
  - Data-intensive workloads are not very processor-demanding so rather than moving data around it is more efficient to "move processes around" (processors and storage are assumed to be co-located)

Characteristics of Cloud Computing (2/2)

- Hide system-level details from application developer
  - Normally, distributed programming involves keeping track of many details (i.e. details across several threads, processes and machines, etc.)
  - **MapReduce** addresses the challenges of distributed programming by providing abstraction to isolate developer from system-level details (e.g. locking of data structures)
  - User will see only simple, well-defined interfaces between small number of components (separation of “what” needs to be done and “how”)
- Process data sequentially and avoid random access
  - Seek times for random disk access are limited by mechanical nature of devices, e.g., 1 TB DB with 10^{10} 100-byte records needs months for access and mutation of records on a single machine, while only 1 day if done in sequential fashion
  - Desired to avoid random data access

Algorithm designers are faced with problems of diminishing returns as the increase of the degree of parallelization increases communication (e.g., mythical man-month)

Cloud computing service markets



Amazon Leads $150-Billion Cloud Market

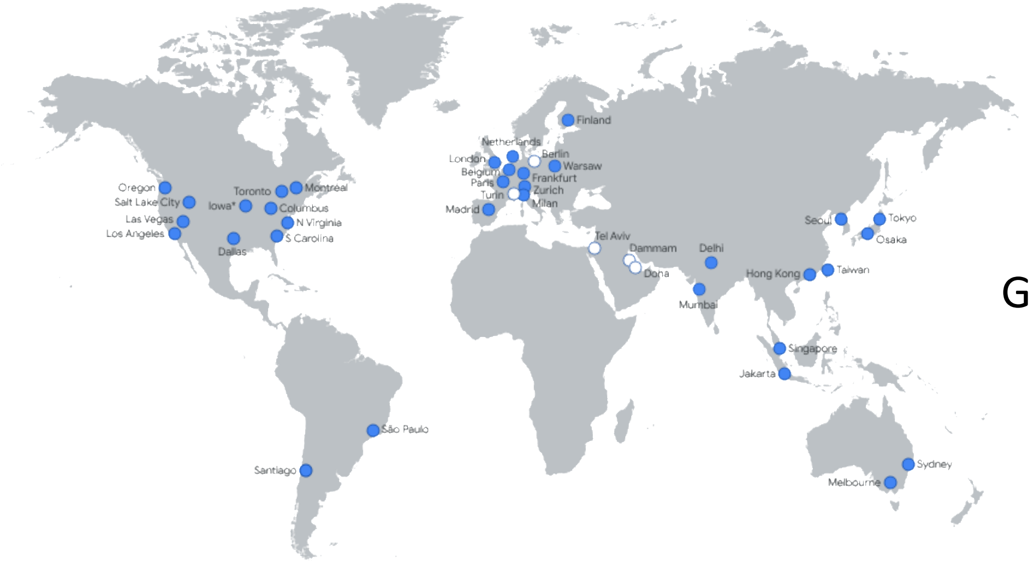
Worldwide market share of leading cloud infrastructure service providers in Q1 2021*

Q1 2021 cloud infrastructure service revenue
$39 billion
(+37% vs. Q1 2020)

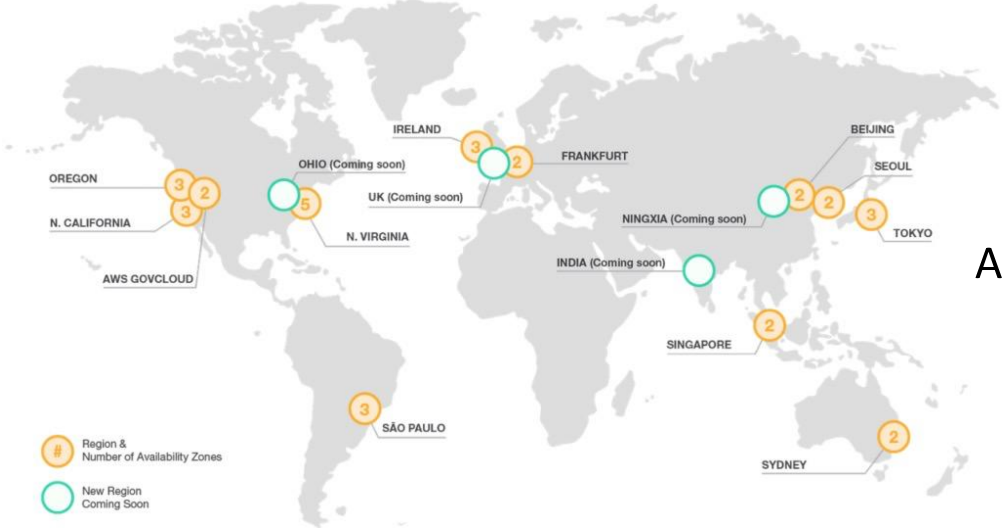
* includes platform as a service (PaaS) and infrastructure as a service (IaaS) as well as hosted private cloud services

Source: Synergy Research Group

Cloud Server Location



Google Cloud



Amazon AWS

Google Data Center Security

Microsoft Underwater Data Center

NY Times Example: TimesMachine Project

- New York Times wanted to make its historical archive available online
- The company needed to process 11M articles and turn them into pdf files
- Initial estimation showed that hundreds of servers and large storage were needed, resulting in huge costs and significant delay before the service could be deployed
- 100 EC2 instances and 4 terabytes of S3 storage allowed to finish the job in few days for a cost of $240

The New York Times

Netflix Example

All services of NETFLIX are migrated to cloud servers

Completing the Netflix Cloud Migration

[source: https://about.netflix.com/]

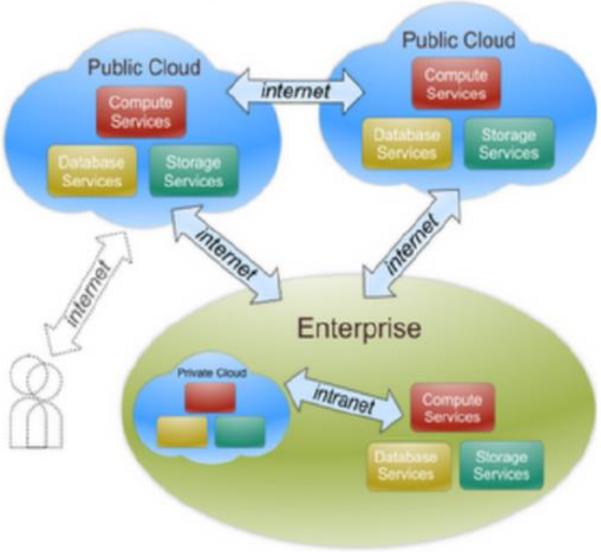
Netflix Example

Why Netflix migrates to cloud servers?

- It began in August of 2008, when they experienced a major database corruption and for three days could not ship DVDs to our members.
- That is when they realized that we had to move away from vertically scaled single points of failure, like relational databases in our datacenter, towards highly reliable, horizontally scalable, distributed systems in the cloud.
- They chose Amazon Web Services (AWS) as their cloud provider because it provided with the greatest scale and the broadest set of services and features.

Private, Public and Hybrid Clouds

- **Public Cloud -** Owned and operated by third parties. All customers share the same infrastructure pool with limited configuration, security protections, and availability variances managed and supported by cloud provider. Public cloud may be larger than an enterprises cloud
- **Private Cloud -** Built exclusively for a single enterprise. Aim to address concerns on data security; offers greater control, which is often lacking in a public cloud
  - On-premise Private Cloud: internal clouds hosted within ones own data center
  - Externally hosted Private Cloud: Hosted externally with a cloud provider, however the provider guarantees an exclusive cloud environment with full privacy. No sharing of physical resources
- **Hybrid Cloud -** Combination of public and private cloud models
  - Service providers can utilize 3rd party cloud providers in a full or partial manner thus increasing the flexibility of computing. Often used for providing on-demand, externally provisioned scale. Private cloud can be augmented with the resources of a public cloud in case of unexpected surges in workload



Cloud Computing Concerns (1/2)

- Security of confidential data
  - In cloud computing, it is common to store data of multiple customers at one common location. Shared infrastructure increases the potential for unauthorized access and exposure
  - Care must be taken to ensure that one customer’s data does not affect another customer’s data. In addition, cloud computing providers must be equipped with proper disaster recovery policies to deal with any unfortunate event
  - Important to be aware of data administrators and their extent of data access rights

Cloud Computing Concerns (2/2)

- Regulatory compliance policies
  - In some European countries, Government regulations do not allow customer's personal information and other sensitive information to be physically located outside the state or country. In order to meet such requirements, cloud providers need to setup a data center or a storage site exclusively within the country to comply with regulations. Having such an infrastructure may not always be feasible and is a challenge for cloud providers
  - In many instances, the actual storage location is not disclosed, adding onto the security concerns of enterprises
- Consistency around authentication, identity management, compliance
  - To reassure their customers, cloud providers must offer a high degree of transparency into their operations

Examples: Amazon Web Services

- Large collection of toolkits
  - Virtual machine hosting – EC2 (Elastic Compute cloud)
  - Storage - S3
  - Database – SimpleDB
  - Content delivery – CloudFront
  - Queue service – SQS
  - Etc.
- 99.95% uptime for EC2 guaranteed

Availability	Downtime per year	Downtime per month	Downtime per week
90%	36.5 days	72 hours	16.7 hours
95%	18.25 days	36 hours	8.4 hours
99%	3.65 days	7.20 hours	1.68 hours
99.9%	8.76 hours	43.2 min	10.1 min
99.95%	4.38 hours	21.56 min	5.04 min
99.99%	52.6 min	4.32 min	1.01 min
99.9999%	31.5 sec	2.59 sec	0.605 sec

Examples: Google App Engine

- Over 1 million customers. Google itself runs on this platform
- Provides backend datastore and API for anyone to build highly-scalable web applications
- Google maintains the whole infrastructure freeing the user from having to backup, upgrade or patch basic services
- 99.95% uptime SLA
- Carrot approach
  - Free usage for anyone up to certain limit and then billing over the limit usage
  - Rich set of available tools
- Proprietary implementation of MapReduce in C++
  - Bindings in Java, Python