Scalable Web Programming

CS193S - Jan Jannink - 2/16/10
Administrative Stuff

• Submit a running website on Thursday
  
  • some functionality can still be simple
  
  • some placeholders acceptable

• Some tests required
  
  • complete coverage not necessary

• critical internal API
Weekly Syllabus

1. Scalability: *(Jan.)*
2. Agile Practices
3. Ecology/Mashups
4. Browser/Client
5. Data/Server: *(Feb.)*
6. Security/Privacy

7. Analytics*
8. Cloud/Map-Reduce
9. Publish APIs: *(Mar.)**
10. Future

* assignment due
Privacy Review - Google Buzz, One Week Later

• High profile launch

• High traffic
  • millions of users have tried it

• Privacy leaks (guaranteed to happen but difficult to predict)
  • woman’s ex-husband & new boyfriend

• Changes
  • many features are now opt in instead of opt out
Internet Analytics (continued)

- US still represents ~50% of allocated IP addresses (1.5 billion)

- 87 billion Google searches in December (2.8 billion / day, ~1 trillion / year)

- If you don’t do anything a site might get found twice a day

- How do you raise awareness in a saturated environment?
  - Have what people want, make sure crawlers index it
  - Pay for raised awareness, make sure audience sees it
  - Find those who connect with your content
Web Analytics

• Key Metrics (low to high value)
  
  • page views
  
  • unique users
  
  • session (bounce rate)
  
  • sign ups
  
  • repeat visits
Comparative Tools

• Alexa
  • toolbar based

• Compete
  • network traffic over carrier networks

• Quantcast
  • ~50% of top 1000 sites submit their traffic
Google Analytics

• Track
  • referrals
  • page events
  • ad campaigns

• Benefits
  • accurate site crawl
  • more detailed ranking of pages
Organic Traffic vs. Episodic Traffic

- Organic traffic growth process
  - run experiments (with real users)
    - bring users to site
    - get feedback
  - repeat as necessary
  - use results to derive a site specific growth formula
- Transform episodic traffic into organic traffic
Traffic Referral

• Relevant content + good site structure + traffic = better Google ranking

• Many content aggregators
  • slashdot effect
  • digg, reddit, mixx

• Social media
  • buzz, FriendFeed, StumbleUpon, Delicious

• Traditional media, advertising
imeem Example

• One year in stealth / alpha (20k accounts, hundreds of simultaneous users)

• Eight months in public beta before takeoff (May ’06)

• First sustained marketing in January ’06 (Sundance festival)

• Tried film, music, fashion, while developing features all along
  • partner requested features helped polish platform

• Viral content player + playlist feature proved the winner
  • embedding into myspace helped drive a lot of traffic
Lessons Learned

• If possible, build a platform not a product

• Iterate quickly on a feedback loop around new users
  • increased satisfaction brings more users back

• Use marketing events to grow user base, develop partnerships & feature set
  • partnerships should always include highly engaged counterparts

• Choose activities that will put you in touch with power users
  • power users make the service live and sell the product organically
Power Law of Participation

Partners

Power Users

Members

Visitors

Collective Intelligence

Collaborative Intelligence

Lead

Moderate

Collaborate

Write

Network

Share

Refactor

Core

Periphery

Low Threshold with Tool

(cc) Ross Mayfield 2006
The Platform Lesson

• Allow feature shift without having to start from scratch

• Build towards user wants in small steps learning from each

• Focus on every success, deprioritize but don’t eliminate failures

• General models of user behavior are powerful
  • Sharpen them based on data not idealized models

• Make the platform about activities as opposed to location
  • Define features around doing, not static models
Role of Analytics

- It is rare to succeed at growth without working at it
- Compare reported user behavior to actual behavior
- Profile your user categories numerically
- Understand relationships between user categories
  - producer consumer dynamic among users
  - develop feature set to enable users to increase engagement
- Discover which efforts are working
Worth Checking Out

• Google analytics
  • http://www.google.com/analytics/

• Analytics blog
  • http://www.kaushik.net/avinash/

• Logging howto
  • http://www.campin.net/newlogcheck.html

• Nagios
  • http://www.nagios.org/
Q & A Topics

- Project questions
- Developing a platform
- Making marketing & customer service integral to product development