The New Consumer Data Revolution: Who Pays Whom?
Reflection on Google

- Market Capitalization (May 6, 2008, in billion USD)
  - IBM: 171
  - Google: 184
  - Microsoft: 273

- Who here works for Google?
  - We all do!

- What does Google know about you?
Google data sources

- Search
  - Query sequence
  - Choice set
  - Which link the user chooses
  - Queries and their refinements, clicks on links and returns
  - Use individually and in the aggregate

  ▶ Use: Search quality
  ▶ Use: Trends

- Gmail
  - Content and social relationships
  - Opening up the social graph (Orkut)
  - Response behavior: model

  ▶ Use: relevance ranking
  ▶ Use: Discovery
  ▶ Spam
    - Mark as spam
Google data sources

- **Toolbar**
  - Since 2002
  - Follow user across sites
  - Always with user, give data to get data

- **Notes**
  - Annotate web pages

- **Analytics**
  - Javascript embedded by site owner
  - Gives insight e.g., on how popular a given link is
  - Insight in what people do going through websites

- **Reader**
  - Read, Mark, Save, Tag
  - Forward to friend

- **Docs**
  - Content and social relationships
Google data sources

- **Google groups**
  - Vs Wiki / Google docs

- **Talk**
  - E.g., chat on webpage

- **Maps**
  - Intent
  - Repository of 3D objects
  - UGC, being able to change info (200m)

- **Location**
  - Phone: My Location feature
  - GPS

- **Calendar**
  - Future
- **Design tools for home**
  - Creates qualified leads for e.g., window manufacturer

- **Checkout**
  - From the cradle to the grave
  - For purchases of higher priced items, process of learning about product dimensions etc
  - Payment systems
  - Google checkout is a low margin business that drives the high-margin business of ads
Past, present and future

Who knows about your relationships with others?
How does Google (or Amazon) differ from IBM (or Microsoft)?

Past, present and future

Past
- Given data, get me insights

Present
- Given problem, get me data

Future: Consumer data revolution
- I give you data, you give me money!

Framework

Architectures of experimentation (Web1)
- Give set of possible actions, do experiments

Architectures of participation (Web2)

Architectures of interaction (Web3)
Data and business

Key: Users now actively contributing data
  - Four stages

The data business

Consumers and conversations

Summary

- Data companies pay employees to create data
- Implicitly generated data (traces)
  - Experiments
- Users contribute data
  - Incentives
  - Data about themselves, about others, about relationships
  - (not binary)
- Users expect value for themselves
The Consumer Data Revolution

- The First Data Revolution
  - Example: Amazon.com
  - Mainly implicit data, some explicit data

- The Second Data Revolution
  - Example: Facebook
  - Mainly explicit data
  - Customers start interacting

- The Consumer Data Revolution
  - How their attitudes to their info changes
Business Machines: Communication, Computation and Data

- **1970’s**
  - “Experts” learn a language the computer understands
  - Digitizing back office
  - 10M people

- **1980’s**
  - Front office interacts with back office
  - 100M people

- **1990’s**
  - Customers interact with firm
  - Search: 1bn people poking at stuff

- **2000’s**
  - 1bn people poking at stuff
  - 100M people producing stuff
    - Peer-production and collaboration
    - Customers interact with customers

- **Now**
  - Discovery in addition to search
    - Serendipity: Discover what not searched
  - People in addition to pages
    - Social commerce
  - Mobile in addition to PC (and paper)
    - Continuous partial attention
  - Model current situation plus history
    - Sensing
Who will select the best picture?

See other businesses along 10th St

Interactive Map and Directions

View larger image

Share your own customer images
Collect relevant customer feedback

Capture context automatically

**Suggestion Box**
Your comments can help make our site better for everyone. If you've found something incorrect, broken, or frustrating on this page, let us know so that we can improve it. Please note that we are unable to respond directly to suggestions made via this form.

**If you need help with an order, please contact Customer Service.**

**Please mark as many of the following boxes that apply:**

- [ ] Product information is missing important details.
- [ ] Product information is incorrect. Propose corrections using our Online Catalog Update Form.
- [ ] The page contains typographical errors.
- [x] The page takes too long to load.
- [ ] The page has a software bug in it.
- [ ] Content violates Amazon.com’s policy on offensive language.
- [ ] Product offered violates Amazon.com’s policy on items that can be listed for sale.

**Comments or Examples:**
Examples: Missing information such as dimensions and model number, typos, inaccuracies, etc.

I went to the bathroom and came back, and the page was still loading!!

我去了浴室，又回来，网页却还在下载！！
Let Amazon.com access your camera and microphone?

- November 2007
  - Adobe Flash installed on 820 million Internet-connected computers and mobile devices
**IMMI**

- Listening into your room
  - every 30 seconds,
  - for 10 seconds.
Data at the heart of the Digital Networked Economy

- Clicks
  - Music
  - Dating

Try now!
The Mix Maker for your MP3s

More information on version 2.7.1

TiVo Owners

© 2003 people & data | www.peer2peer.com | +49 174 90 65 096 | +86 138 1819 3800
Data collection and analysis (Amazon.com)

- **Level**
  - **Customer**
    - 消费者
  - **Orders**
    - 订单
  - **Session aggregates**
    - 访问总计
  - **Clicks**
    - 点击

- **Data collected 2007**
  - 100 MB
  - 10 GB
  - 1 TB
  - 100 TB

- **Amount of data**
  - 数据量

© people & data | www.2eigend.com | +1 650.906.5906 | +49 174.906.5906 | +86 138.1819.8390

28
Data in the Digital Networked Economy

- Clicks
  - Selection
- Intention
  - Search
- Attention
  - Engagement
- Interaction
  - Relationship
- Situation
  - Discovery
- Geolocation
  - Mobile phones
  - GPS
Navigation with two-way internet connectivity (dash.net)

Dash Express™

Dash Express is the first navigation system with two-way, Internet connectivity. It’s packed with features that give you the smartest way to get from A to B and information about everything in between.
A data revolution, not a software revolution

- Mapping companies did not realize that users can add value...
  - Example: NAVTEQ
    - acquired by Nokia for USD 8.6B in Oct 2007
    - spent USD 300M to reach breakeven
    - 1000 employees driving with GPS...
    - ... vs 100M GPS-enabled Nokia phones alone sold in 2008

- ...vs Amazon.com realizing early on that users can add value
  - E.g., by reviewing books

- ... vs Google enabling external developers to build services using company’s data

- Q: When will airlines, banks etc follow?
"Pay as you drive" insurance (Norwich Union)

"Pay as you drive" insurance, which is a new and fairer way to insure your car, is designed for people who don’t use their car as much as others. You simply pay variable costs based on when and where and how far you drive plus a fixed monthly fee.

You could save up to 30%*

Car Insurance is now fairer by miles

- If you drive less than 6000 miles, "Pay As You Drive"™ could save you up to 30% on your car insurance.
- A Global Positioning System (GPS) is fitted to your car for free. The GPS allows us to work out how you use your car, so you can enjoy a fairer premium.
Give data to get data about applicants (JobScore)

- Use information about candidates from interviews at other companies

Hire Great People
Build your pipeline, stay organized, and find great talent, all for free!

Take a Tour or Sign Up

Already have an account? Sign In
The basis of the Digital Networked Economy is *Data*.
Fast innovation by *experimentation* (the scientific method)
The five levels

- Community
- Interaction
- Participation
- Experimentation
- Data
## Web1.0 vs Web2.0

<table>
<thead>
<tr>
<th></th>
<th>Web1.0</th>
<th>Web2.0</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Read-only</strong></td>
<td>Read-only</td>
<td>Read + write</td>
</tr>
<tr>
<td></td>
<td>只读网页</td>
<td>可读写网页</td>
</tr>
<tr>
<td><strong>High barrier of exit (stickiness)</strong></td>
<td>高的退出壁垒（粘性）</td>
<td>低的进入壁垒</td>
</tr>
<tr>
<td><strong>Pretest and validate</strong></td>
<td>预先测试和验证</td>
<td>发起和学习；推进和请求</td>
</tr>
<tr>
<td><strong>Taxonomy (controlled directory)</strong></td>
<td>基于超链接的探索（体现出作者的控制）</td>
<td>“Tagsonomy” (not controlled)</td>
</tr>
<tr>
<td></td>
<td>Discovery based on hyperlinks (expressing control of author)</td>
<td>Discovery based on social relations (trust, reputation)</td>
</tr>
<tr>
<td></td>
<td>基于超链接的探索（体现出作者的控制）</td>
<td>基于社会关系的发现（信任，信誉）以及其他人的元数据。</td>
</tr>
<tr>
<td><strong>Push: Supply driven.</strong></td>
<td>推式广告：供给驱动</td>
<td>拉式发现：需求驱动</td>
</tr>
<tr>
<td><strong>Target</strong></td>
<td>目标定位</td>
<td>搜索</td>
</tr>
<tr>
<td><strong>Search</strong></td>
<td>搜索</td>
<td>搜索</td>
</tr>
<tr>
<td><strong>Discover</strong></td>
<td>发现</td>
<td>发现</td>
</tr>
</tbody>
</table>

*Web1.0* and *Web2.0* are two major versions of the web, with *Web2.0* being more interactive and user-friendly.
Architectures of Participation: Some Attributes of Web2

- **User focus (Me-business, not E-Business)**
  - User is at the center of Web2
    (not the company)
  - VRM (Vendor Relationship Management)
    not CRM (Customer Relationship Management)

- **System engineered for feedback**
  - System engineered to improve over time by leveraging user data
    (not deteriorating over time)

- **Network effects**
  - Demand-side economies of scale
    (not only supply-side economies of scale)

- **Transparency**
  - Google Maps: Create API for developers to use the data
    (not increase security)
Participation

- Why do people participate?
  - Spread their ideas
  - Belonging
  - ...
  - ...

- Incentives
  - Pay people (e.g., Yahoo directory, Mahalo)
  - Get volunteers (e.g., Wikipedia)
  - Create self-interest (e.g., BitTorrent)
Communication

- (0) Broadcast
- (1) Collect
- (2) Experiment
- (3) Contribute
- (4) Interact
- (5) Broadcast
<table>
<thead>
<tr>
<th>Sources</th>
<th>Metrics</th>
<th>Applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collect</td>
<td>Wall Street</td>
<td>Trading models</td>
</tr>
<tr>
<td>Experiment</td>
<td>Clicks</td>
<td>Company centric</td>
</tr>
<tr>
<td>Contribute</td>
<td>Profile</td>
<td>User centric</td>
</tr>
<tr>
<td>Interact</td>
<td>Links</td>
<td>Relationship centric</td>
</tr>
</tbody>
</table>
The Data Business: Who Pays Whom?

- Value of data proportionally to their impact on decisions

- **Traditional data business**
  - **Acxiom**
    - Who do I send my advertisements to?
    - (economics of communication)
  - **Maps**
    - What route do I take?
    - (discussed earlier)
  - **OAG**
    - What flight suits me best?
    - (user generated, screen-scraped)
  - **MLS (Multiple listing service)**
    - What house do I buy?
    - (user generated: Zillow, Redfin, Trulio)
  - **Contacts**
    - Who do I sell to? Hire?
    - (LinkedIn, jigsaw)
Buy and trade business cards (Jigsaw)

- 7M business cards (2007.12)
- 200k members
- Community corrects

Jigsaw contacts are purchased with points.
You can earn points by:
- Adding Contacts** (5 Points+)
- Correcting Contacts (5 Points+)
- Referring others (125 Points)

** All contacts are added anonymously.

Jigsaw has two choices for participation:
1) PLAY
   Trade contacts for Contacts
2) PAY
   Subscribe for a low monthly fee to receive contacts
   If you need more information, call 1-877-Jigsaw9.

Loss of control
Where do people go for information?

- 20 years ago: Produced news
  - Publicized opinion

- 10 years ago: Search
  - Search: Already know what you are looking for
  - Discovery: Find something you didn’t know you were looking for (Serendipity)
  - Blogs, Wikipedia

- Now: Social network
  - Discover via people you are connected to
  - “Generation C”
Conversational Marketing

- Conversation
  - Between whom?

Company

downcasting

Individuals
The Consumer Data Revolution: Who Pays Whom?

- What does Google know about you?
  - Data and business
  - Key: Users now actively contributing data
    - Four stages
  - The data business
  - Consumers and conversations

Summary
- Data companies pay employees to create data
- Implicitly generated data (traces)
  - Experiments
- Users contribute data
  - Incentives
  - Data about themselves, about others, about relationships
  - (not binary)
- Users expect value for themselves
Thank you!