From E-Business ...
... to ME-Business

Andreas S. Weigend, Ph.D.
People & Data
http://www.weigend.com

Fall 2006

Some trends

- How has the behavior and the attitude of customers shifted in response to new technologies and new incentive structures on the web?
- From algorithmic search to social search/ *Leverage technology and people to have your toughest questions answered*
- From global and local irrelevance to glocal relevance/ *When does location matter?*
- From the money economy via the intention economy to the attention economy/ *Why an economy? Because it can be measured!*
- From traditional market research to observing the behavior of all individuals/ *Why sample if you can have all?*
- Specifically, the effect of the long tail
  - Production: From controlled production for the masses to uncontrolled production by the masses/ *Why?*
  - Consumption: From traditional marketing push to individual discovery/ *How can you help serendipity along?*
Selection

- One item a day
  - woot.com

- A few items a week
  - tchibo.com: “A new experience every week”

  or

- 100,000 designs
  - zazzle.com
Recommendations

- Expert recommendations (Web 0.0)

  or

- Statistical recommendations (Web 1.0)
  - Amazon.com

  or

- Individual recommendations (Web 2.0)
  - thisnext.com
<table>
<thead>
<tr>
<th>1986</th>
<th>1996</th>
<th>2006</th>
</tr>
</thead>
<tbody>
<tr>
<td>Read-only web</td>
<td>Read-write web</td>
<td></td>
</tr>
<tr>
<td>“Consumer”, “User”</td>
<td>“Producer”</td>
<td>(Lower barrier of entry)</td>
</tr>
<tr>
<td>(Increase barrier of exit)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Static</td>
<td>Temporal</td>
<td></td>
</tr>
<tr>
<td>Taxonomy (controlled)</td>
<td>“Tagsonomy” (uncontrolled)</td>
<td></td>
</tr>
<tr>
<td>Pretest and validate</td>
<td>Launch and learn</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Broadcast email, newsletters</th>
<th>Discovery based on hyperlinks (expression of author’s attention and control)</th>
<th>Discovery based on social relations (trust, reputation), and metadata by other humans. RSS.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Target</strong></td>
<td><strong>Search</strong></td>
<td><strong>Discover</strong></td>
</tr>
<tr>
<td>Push advertising:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Supply driven.</td>
<td>Pull discovery:</td>
<td>Demand driven</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pay for space</th>
<th>Pay for presentment</th>
<th>Pay for click, action</th>
</tr>
</thead>
</table>

### Examples

- **Money economy → Intention economy → Attention economy**
  - Information economy → Recommendation economy; Platform economy
  - Use attention stream: Attensa
  - Use context: Cleverset

- **Paying a few experts to create content → User generated content**
  - Implicit data + explicit data; Metadata
  - Examples: Music, Tags
  - Use socio-, psycho-, demographics: Nugg.ad
  - Currencies (itunes vs Spiralfrog)

- **Algorithmic search → Social search**
  - People answer questions: Google Answers, MSN, Yahoo
  - Use info in files on your computer to determine relevance: Illumio

- **Discovery → Social discovery**
  - Head → Long tail
  - URLs: Del.icio.us
  - Photos: Flickr

- **Global → Local**
  - Use mobile: Socialight
  - Use WLAN: Jambo.net
Increase of communication

• 0. Business metrics and data collection
  ▶ An iterative process

• 1. Data Analysis
  ▶ Data mining: Description, prediction

• 2. Architectures of Experimentation
  ▶ A/B test, active learning, survey design...

• 3. Architectures of Participation
  ▶ Remember, share, discover
  ▶ Empower and incentivize people to contribute
  ▶ Self-expression

• 4. Architectures of Interaction
From Faith to Data

- The Era of Faith
  - Massive investments into cathedrals etc.
  - Unclear ROI (Return on Investment)
  - No feedback, or long feedback cycle

- The Era of Data
  - Massive investments into measuring, networking, storing
  - ROI measurable
  - Short feedback cycle
  - Experiments

Turning behavior into data

- Revealed preferences
  - Music
  - Search
  - Online trading
  - Online dating

© 2006 Weigend Associates LLC | www.weigend.com

© 2007 People & Data | Weigend Associates LLC
Information at www.weigend.com
Everything can and will become data
任何东西都够能且一定会变成数据

- Additional sources of data about people
人类的其它数据来源

  - Movement
移动
  - Mobile phones
手机
  - GPS

© 2007 People & Data | Weigend Associates LLC
Information at www.weigend.com
Everything can and will become data

- Additional sources of data about people
  - Movement
    - Mobile phones
    - GPS
  - Identity of person
    - DNA analysis
    - Available data:
      - 30,000 genes
      - 1 billion base pairs
      - 10亿碱基对
  - Brain activity
    - Neurorkmarketing
    - fMRI analysis of response to stimuli
      - fMRI
    - RFIDs (Radio frequency identifiers)
      - RFIDs
- Unique identifiers for physical objects
- Facts
  - Price: 2 cent
  - Size: 2mm
- Opportunities
  - Supply chain
    - Wal-Mart: $8 billion savings expected per year by using RFIDs
  - Market research
    - Dynamic pricing
- Fears
  - Loss of privacy
  - Dictatorship
- Reality
  - It will happen: Big business
  - Need: Understand pros and cons and make conscious decisions
- Compare to web sites

© 2006 Weigend Associates LLC | www.weigend.com
Storage is free
免费存储

Dramatic drop in price
价格大幅下降

Exponential increase in storage
存储量呈指数增长

Why now?
为什么现在发生？

Data collected implicitly: Dramatic growth over time
隐秘采集数据：随时间推移急剧增长

Data collected explicitly: Amount constant over time
公开采集数据：随时间推移变化不大

© 2007 People & Data | Weigend Associates LLC
Information at www.weigend.com
Why now?
为什么现在发生？

Communication
通信

- Malthus's Law of Information:
马尔萨斯信息定律:
  - New information content is doubling every year
    新信息内容每年翻一番
  - Time spent on information consumption is constant
    信息消费时间几乎不变

Large e-business company: Amount of data created per year
大型电子商务公司年均数据产量

- Level
  - New data per year
    每年新数据量
  - Customer
    消费者
    100 MB
  - Orders
    订单
    10 GB
  - Session aggregates
    访问总计
    1 TB
  - Clicks
    点击
    100 TB
1. Re-Define
(重新)定义
   - Objectives and benchmarks
     目标和基准

2. Measure
测量
   - Collect, store, manage the data
     收集、储存和管理数据

3. Describe
描述
   - Exploratory data analysis
     探测性的数据分析

4. Predict and evaluate
预测和评估
   - Probabilistic models
     概率模型

5. Decide, act, and evaluate
决策，行为和评估

Define Objectives
定义目标

- Stock price
  股票价格

- Profit
  利润

- Number of items sold
  销售数量

- Number of visits
  访问量

- Rate of conversion
  行动转化率

- Customer acquisition
  赢得消费者

- Customer retention
  留住消费者

- Customer satisfaction
  消费者满意度
Measure

- Orders
- Overall use of the site
  - Buying vs selling
  - Searching vs browsing
  - Engagement: Reviews, etc.
- Customer service contacts
  - E-mail, phone
- Surveys
  - Satisfaction
  - Intentions / Goals / Modalities

Why is it hard?

- Even simple behavioral analysis requires significant infrastructure

Reporting → Behavioral analysis and predictive modeling

Cost center → Profit center
2. Experimentation

- 1. Data Analysis
  - Data mining: Description, prediction

- 2. Architectures of Experimentation
  - A/B test, active learning, survey design...

- 3. Architectures of Participation
  - Remember, share, discover
  - Empower and incentivize people to contribute
  - Self-expression

- 4. Architectures of Interaction
Also want to make recommendations!

**Complements**
*(buy in addition to)*

- Customers who bought X also bought Y

**Substitutes**
*(buy instead of)*

- Customers who bought X also bought Z

Also want to make recommendations!

Also want to make recommendations!
Result: Right vs Left

- **Metrics**
  - **Conversion rate**: Percentage of visits placing an order
    - 转化率：百分之几的人访问页面并购买
  - **Order size**: Number of additional (from the second page) items put in cart
    - 订单大小：（从第二页起）新购多少商品

- **Result**
  - "Your Shopping Cart" on right is about 1% better than on left
    - "Your Shopping Cart"置于右侧比置于左侧的效果提高1%

- **Some details**

<table>
<thead>
<tr>
<th>Details</th>
<th>All customers</th>
<th>Existing customers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cart-adds from 2nd page</td>
<td>+0.6%</td>
<td>+0.8%</td>
</tr>
<tr>
<td>Wishlist-adds</td>
<td>+1.4%</td>
<td></td>
</tr>
<tr>
<td>DVD Cart-adds</td>
<td></td>
<td>+0.8%</td>
</tr>
<tr>
<td>DVD ($)</td>
<td>+1.1%</td>
<td>+1.0%</td>
</tr>
</tbody>
</table>

Guiding the customer to the ultimate purchase

What similar items do customers ultimately buy after viewing this item?
Long-term Effects

- Displaying Google’s “sponsored links” within Amazon.com’s site?
  - Initial effect positive...
  - ...but how to model whether this helps or hurts the company in the long run?
  - ...but how to model whether this helps or hurts the company in the long run?

- Offer free shipping
  - Competitors quickly adopt, how to model the overall game?

- Pricing, Promotions...

- Role of price comparison engines (shopping.com, etc.)?
  - Some price comparison engines (e.g., shopping.com) role is what?
"... Please let us know what you plan to do at Amazon.com today"
"......请告诉我今天你计划在亚马逊网站上做什么"

- Free responses, hand-coded into non-exclusive categories
  自愿回复，将回复进行分类编码，归入非一些非互斥的类型
  - Multiple assignments possible
    可能回复多种可能
  - Average: 1.2 categories per response
    平均：每个回复涵盖1.2个种类
      - Number of responses: 1023
        回复数量：1023
      - Response rate: 3.1%
        回复率：3.1%
      - Date: February 11-12, 2003
        日期：2003年二月11-12日

**Why do People Visit?**

<table>
<thead>
<tr>
<th>Percentage</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>35%</td>
<td>Research: Response indicates having a target or topic for which customer wants to gather information in depth (&quot;looking for books about Chile&quot;)</td>
</tr>
<tr>
<td>31%</td>
<td>Browse: Looks at items casually at a more general level than Research (&quot;looking at music/books&quot;, without further specification)</td>
</tr>
<tr>
<td>16%</td>
<td>Buy: Response indicates intent to buy in this visit (&quot;buy a book&quot;)</td>
</tr>
<tr>
<td>10%</td>
<td>Complain: Complains about some feature of the site</td>
</tr>
<tr>
<td>9%</td>
<td>Post-buy: Mentions checking order status, other account activities</td>
</tr>
<tr>
<td>7%</td>
<td>Community: Mentions usage of some feature of community, like reviews</td>
</tr>
<tr>
<td>7%</td>
<td>Price: Mentions getting information about price for specific items</td>
</tr>
<tr>
<td>2.9%</td>
<td>Goldbox: Mentions Goldbox activity</td>
</tr>
<tr>
<td>2.2%</td>
<td>Gift: Indicates looking for a gift for someone else</td>
</tr>
<tr>
<td>2.0%</td>
<td>Sell: Mentions a selling activity</td>
</tr>
<tr>
<td>1.4%</td>
<td>Personalization: Mentions personalization feature (e.g., recommendations)</td>
</tr>
<tr>
<td>0.8%</td>
<td>Used: Mentions finding or transacting a used item</td>
</tr>
</tbody>
</table>
Stated vs Revealed Preferences
自认的偏好与实际的偏好

- Obtain insights by combining individual survey response with click analysis:
结合个人调查问卷的回复与点击情况分析，得出一些结论

Look at those who ended up buying something:
最终购物的网民中

Only about one-half of those making a purchase indicated that they wanted to buy something in this visit
只有1/2事先称计划购物

Look at those who said they wanted to buy something:
最初想来购物的网民中

Only about one-third of those indicating intent to buy ended up making a purchase in that visit
只有1/3最终完成购物

Customer Satisfaction
客户满意度

- Relate satisfaction rating to:
把满意度评价与下列两项联系起来:
  - Visit intent (from survey)
    访问意图（来源于调查问卷）
  - Actions (from clickstream)
    行动（从其点击行为看）

- Intermediate goal
  中期目标
  - Predict satisfaction
    预测满意度

- Ultimate goal
  最终目标
  - Increase satisfaction
    提高满意度

  - Number of responses: 6117
    回复的数量：6117
  - Dates: March 15–24, 2003
    日期：2003年3月15-24日
3. Participation

1. Data Analysis
   - Data mining: Description, prediction

2. Architectures of Experimentation
   - A/B test, active learning, survey design...

3. Architectures of Participation
   - Remember, share, discover
   - Empower and incentivize people to contribute
   - Self-expression

4. Architectures of Interaction

Platform: Yellow Pages

See other businesses along 10th St.

View larger image
Share your own customer images

Walk Right
Walk Left
Make customer feedback trivially easy

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.
- 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!!

[Submit]
Money Economy → Intention Economy → Attention Economy

Click

Only able to click on links given by site.

From intention to attention

Search

Express intention. Doesn't depend on result.

Click

Only able to click on links given by site.
Tags are distilled attention, a pure form of attention.

- You are what you tag.  
  书签展示真我
- You are what you are tagged as / who you are tagged by.  
  书签决定你的存在

Example: del.icio.us

Follow a tag and discover a topic  
探索别的网友如何使用书签

Follow a user and discover what he is interested in  
探索该网友的兴趣爱好

Discover other users who have tagged the page  
探索给同一网站贴上书签的其他网友

Google Maps + Flickr + Google
- Blogging maps by lawrenceli ... and 3 other people ... on 2005-05-13 ...

HousingMaps craigslist housing google
- Maps of LawrenceLi and 448 other people on 2005-05-13 ... copy this i

The History of Sampling  
- Discover other users who have tagged the page  
  to music was invented by lawrenceli ... and 259 other people ... on 2005-05-10 ... 

LGBT-CRC of Stanford University - Supporting Stanford
Empower millions of users to contribute

- del.icio.us: tag web pages
- flickr.com: tag photos
- 43things.com: tag your goals

What is happening now?

- Quentin Lee, Filmmaker
  (Drift, Ethan Mao)

Would you like to comment?

Sign up for a free account or sign in if you're already a member.
Levels of Analysis and Actionability

- B. Visit
  - Model: Intention, situation, mode
  - Action: Session-based marketing, occasionalization

- A. Page
  - Model: Content
  - Action: Show ads

Examples of modalities of a visit
- Planned vs impulse visit
- Personal vs job-related visit
- At home vs at work
- Is-in-a-hurry vs has-time-to-kill
- Ready to make a decision
Levels of Analysis and Actionability

- **C. Customer**
  - Model: Demographics, behavior
  - Action: Personalization-based marketing
- **B. Visit**
  - Model: Intention, situation, mode
  - Action: Session-based marketing
- **A. Page**
  - Model: Content
  - Action: Show ads

**Attributes of an individual customer**

- Influencability
- Navigational style
- Early adopter
- Leader vs follower
- Attitude to complexity
- Attitude to technology (vs human)
- Activity across stores
- Degree of price sensitivity
- Degree of time sensitivity
- Median time between clicks
- Level of interest in exploring

© 2006 Weigend Associates LLC | www.weigend.com
Levels of Analysis and Actionability

- **D. Network of customers**
  - **Model:** Apply social network research
  - **Action:** Discounts, better service for opinion leaders

- **C. Customer**
  - **Model:** Demographics, behavior
  - **Action:** Personalization-based marketing

- **B. Visit**
  - **Model:** Intention, situation, mode
  - **Action:** Session-based marketing

- **A. Page**
  - **Model:** Content
  - **Action:** Show ads

---

**D. Network of Customers**

- "Customer Lifetime Value" has two components
  - **Intrinsic value**
  - **Network value**
    - Characterizes amount of influence

- **Application**
  - Amazon.com’s Share the love
  - Send discount to others of item just purchased, and obtain credit towards future purchase if at least one of the recipients also buys that item (within a week)
Chris Anderson's blog: LongTail.com (Aug 16 2006 entry)

The Long Tail
A public diary on themes around my book

Long Tail comment elsewhere...

Nick Carr notes a NYT column that suggests that if all content is "unbundled" and forced to fund itself, the cultural products subsidized by commercial junk may go away. Carr writes: "It's not a sure thing, in other words, that an a la carte menu will end up giving us the widest possible array of choices. Rather than promoting the creation of a "long tail" of diverse products, unbundling may end up pushing even more economic rewards to the "hits," squeezing out a lot of the good stuff." My take: unless he's referring to the number of modern album tracks that...

© 2007 People & Data | Weigend Associates LLC
Information at www.weigend.com Handout page 31
Recommendations create demand in the "long tail"

![Graph showing the "long tail" effect with recommendations]

- 39% of Turns from titles 2001-30,000
- 50% of Turns from titles 1,150-30,000

Source: Neil Hunt, Netflix 2005

© 2007 People & Data | Weigend Associates LLC | Information at www.weigend.com
4. Interaction

- 1. Data Analysis
  - Data mining: Description, prediction

- 2. Architectures of Experimentation
  - A/B test, active learning, survey design...

- 3. Architectures of Participation
  - Remember, share, discover
  - Empower and incentivize people to contribute
  - Self-expression

- 4. Architectures of Interaction