Dr. Alfred Z. Spector  
VP, Research and Special Initiatives  
Google, Inc.  
Asia, September 2009
Abstract

At its core, Google's mission is to organize the world's information and make it universally accessible and useful. The breadth of this mission, coupled with our services based delivery model, provides Google great opportunities to perform research and to innovate in many areas of technology. In this presentation, I'll summarize our approach to innovation and the results we have achieved in many domains; for example, translation, speech, and vision. I'll also discuss some of our focus areas moving forward, and our general approach to research organization. I'll conclude by discussing our interactions with the research world around us, a world with which we desire strong, mutually beneficial connections.
Outline

1. Google’s Mission
2. Our Technical Approach
3. Innovation
4. Research @ Google
5. Key Themes
   A. Totally Transparent Processing
   B. The Rule of Distributed Computing
   C. Hybrid, not Artificial, Intelligence
6. Some Other Research Projects
7. Relationship with Academe and other Research Organizations
8. Summary and Perspective on Computer Science
Google's Mission and Google Research

Organizing the world’s information and
Making it universally accessible and useful.

A research organization optimized for in situ work
Search and more...
How (1): A Focus on Services

• Google primarily delivers services to achieve it’s mission
• Implications
  • Lower cost of Development
  • Economies of scale
    • Lower cost of Installation
    • Lower cost of Operation
  • Resilience
  • Location transparency
  • Service Integration
  • Aggregated user feedback
  • …
How (2): The Google Common Distributed System

• **Vast:**
  – Data in the cloud
  – Processing in the cloud
  – Global Usage
  *(Feasible due to large clusters using decades of distributed computing research)*

• **Implications:**
  – Economies of scale from *shared* infrastructure
  – Low barriers to product launch
  – Decentralized development more feasible
<table>
<thead>
<tr>
<th>Metric</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.3 searches per user</td>
<td>0.55060</td>
</tr>
<tr>
<td>480,000,000 total pageviews</td>
<td>2.7M</td>
</tr>
<tr>
<td>108 milliseconds</td>
<td>-0.54</td>
</tr>
<tr>
<td>2,800,000,000 views</td>
<td>-0.0000339</td>
</tr>
<tr>
<td>425,440.01</td>
<td>56.76%</td>
</tr>
<tr>
<td>9995.55</td>
<td>17.35</td>
</tr>
<tr>
<td>108 seconds/search</td>
<td>1.3 searches per user</td>
</tr>
<tr>
<td>$4.78 RPM</td>
<td></td>
</tr>
<tr>
<td>$7,660,400</td>
<td></td>
</tr>
<tr>
<td>$0.303 CPA</td>
<td></td>
</tr>
<tr>
<td>$1607.44</td>
<td></td>
</tr>
<tr>
<td>10,400</td>
<td></td>
</tr>
<tr>
<td>6.55</td>
<td></td>
</tr>
<tr>
<td>108</td>
<td></td>
</tr>
<tr>
<td>How (3): Empiricism - Let Measurement &amp; Feedback Rule</td>
<td></td>
</tr>
</tbody>
</table>
Key Change: Holistic Approach To Design

Not:

Focus

User

Computer Service

But:

Focus

Vast # of Users

Vast Computer Services

• Implications
  • Users and computers doing more than either could individually.
  • Virtuous circle from: Data and Processing, Reach, Feedback in a virtuous circle.
How Do We Innovate…

• Commitment to advancing technology
  – Rich domain of work due to our mission
  – Some exciting, grand challenge problems

• Cultural agreement that getting a concept into production is often as challenging and fun as its initial invention

• Technical leverage
  – The Google Common Distributed System
  – A Focus on Services
  – Holistic Approach to Design
Google’s Research Mission

To innovate, and to catalyze innovation, and to learn in ways that collectively help Google achieve its mission

Implications:

– Operation in areas relevant to Google
– Broad applicability of many areas of CS and related areas
– A diversified portfolio (various points on risk/reward curve)
– Strong relationship to academic community
– Strong emphasis on publication
Google’s Innovation Culture is Different

• Focus on **talent**

• **Distributed** across the organization:
  
  • Impacting Google necessitates broad, diverse involvement in science and engineering
  
  • Research is done both in our research team and in our engineering organization, organized opportunistically

• Teams benefit greatly:
  
  – From mutual talent
  
  – From Google’s comparative advantages to our scale and broad use
  
  – From service-based architecture ("ease" of working *in vivo*)
Publications

• Publishing/presenting is important
• Publications at [www.research.google.com](http://www.research.google.com)

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of Publications</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004</td>
<td>15</td>
</tr>
<tr>
<td>2005</td>
<td>54</td>
</tr>
<tr>
<td>2006</td>
<td>122</td>
</tr>
<tr>
<td>2007</td>
<td>197</td>
</tr>
<tr>
<td>2008</td>
<td>216</td>
</tr>
<tr>
<td>2009</td>
<td>250+ (p)</td>
</tr>
</tbody>
</table>
Google Publications as of Early 2008

From our web site: http://research.google.com/pubs/papers.html
Key Systems Papers, for Example

Google Cluster Infrastructure

GFS (Google File System)

MapReduce Programming Model for generating & processing large data sets

BigTable: A Distributed Storage System for Structured Data
Chang, et al. OSDI 06.

• Many more papers in other areas, of course
Selected Publications

- **Audiovisual Celebrity Recognition in Unconstrained Web Videos**, Sargin, Aradhye, Moreno, Zhao
- **Automatic Speech and Speaker Recognition: Large Margin and Kernel Methods**, Keshet, Bengio
- **Bid Optimization for Broad Match Ad Auctions**, Even dar, Mansour, Mirrokni, Muthukrishnan, Nadav
- **Dependency Parsing**, Kubler, McDonald, Nivre
- **Detecting The Origin Of Text Segments Efficiently**, Abdel-Hamid, Behzadi, Christoph, Henzinger
- **Domain Adaptation with Multiple Sources**, Mansour, Mohri
- **LSH Banding for Large-Scale Retrieval with Memory and Recall Constraints**, Covell, Baluja
- **OpenFst: An Open-Source, Weighted Finite-State Transducer Library and its Applications to Speech and Language**, Riley, Allauzen, Jansche
Selected Publications

• Outclassing Wikipedia in Open-Domain Information Extraction: Weakly-Supervised Acquisition of Attributes over Conceptual Hierarchies, Pasca
• Solving Maximum Flow Problems on Real World Bipartite Graphs, Negruseri, Pasoi, Stanley, Stein, Strat
• The Unreasonable Effectiveness of Data, Halevy, Norvig, Pereira
• Using the Doubling Dimension to Analyze the Generalization of Learning Algorithms, Bshouty, Li, Long
• Google's Deep-Web Crawl, Madhavan, Ko, Kot, Ganapathy, Rasmussen, Halevy
• Webtables: Exploring the power of tables on the web, Cafarella, Halevy, Wang, Zhang
Selected Publications

- **Cost-efficient Dragonfly Topology for Large-scale Systems**, Kim, Dally, Scott, Abts,
- *Detecting influenza epidemics using search engine query data*, Ginsberg, Mohebbi, Patel, Brammer, Smolinski, Brilliant
- *Discriminating the relevance of web search results with measures of pupil size*, Oliveira, Aula, Russell
- A discriminative kernel-based model to rank images from text queries, Grangier, Bengio
- **Boosting with Structural Sparsity**, Duchi, Singer
- *Affiliation Networks*, Lattanzi, Sivakumar
- *On Sampling-based Approximate Spectral Decomposition*, Kumar, Mohri, Talwalkar
Disk drive failures are a significant problem

- Datasheets do not tell whole story; MTTF is not enough
- Significant and valuable previous work, but insufficient
- Knowledge would help
- Conventional wisdom --- Is it true?
  - Typical disk drive failure rate: < 1% per year
  - Temperature: increases failures
  - Utilization: increases failures
System Health Infrastructure

- Data collected on every machine, periodically
- Stored indefinitely
- Analysis done offline
Temperature

Failure rates per average temperature

- No strong indication that high temp [ ] high failures.
Chromium

Join us in an open-source browser project to help move the web forward

Chromium is the open-source project behind Google Chrome. Take a first look at Chromium and invite you to join us in our effort to:

- Help build a safer, faster, and more stable web for all Internet users to experience the web
- Create a powerful platform for developing a new generation of web applications

Built for the open web

All the code in the project is open source, including JSL, a new JavaScript virtual machine. You can get Chromium's source code and contribute to the project.

Built by the web community

Join our developer discussion group to learn about Chromium and contribute to it further.

- Make Chromium better by finding bugs and filing bug reports.
- Submit patches for known bugs.

Ready to learn more? Read the documentation and follow the Getting Started guide.

The Chromium blog

- Graphics in Chrome
- Site size
  - Google Chrome uses a library called Skia, which is also the graphics engine behind Google's Android mobile OS. The two...
- Context and Overview
  - In user-exposed contexts, 'browser' refers to the frame of an application - the toolbars, sidebars, and buttons that...
- Responsiveness for Modules and the Browser
  - In my last post, I wrote about how we handle I/O in the browser process to keep the main thread of Google Chrome free.

Developer discussion groups

- chromium-discuss
  - Landing on the Air
  - As it seems only once a year after everyone. Set that most of the web relies on hack systems and not in Microsoft's...
- chromium-on-embedded
  - You're still crashes on tablet pc
  - Thank you for the follow-up. Okland
- chromium-mock-up
  - Why is it that everything good has to come out on the Mac? I don't know how many times I wished Transmission was.

©2008 Google - Code Site - Terms of Service - Privacy Policy - Feedback - Site Directory

Google Code offered in: English - Portuguese - Proceed - Español - 氷ペ
open source project
How will you shape it?

Get the source code »

Market Updates!
Android Market is an open service that makes it easy for developers to distribute applications to handsets.
Learn more »

Developers
Android enables developers to easily create great mobile applications.
Learn more »

Open Source
Android Open Source Project provides access to the entire platform source and enables developers to contribute.
Learn more »

In the News
T-Mobile G1 Now Available
T-Mobile
Android Open Source Project
Android
Open Handset Alliance announces the launch of Android Open Source Project
Open Handset Alliance
The Evolutionary Path Forward to New Accomplishments

• Application mix will continue to grow in unpredictable ways:
  – Four areas in flux today: publishing, education, healthcare and government

• Systems will evolve: ubiquitous high performance networking, distributed computing, new end-user devices, ...

• Three truly big results brewing:
  1. Totally Transparent Processing
  2. Ideal Distributed Computing
  3. Hybrid, Not Artificial, Intelligence
Totally Transparent Processing
Totally Transparent Processing

<table>
<thead>
<tr>
<th>D: The set of all end-user access devices</th>
<th>L: The set of all human languages</th>
<th>M: The set of all modalities</th>
<th>C: The set of all corpora</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal Computers</td>
<td>Current languages</td>
<td>Text</td>
<td>The normal web</td>
</tr>
<tr>
<td>Phone</td>
<td>Historical languages</td>
<td>Image</td>
<td>The deep web</td>
</tr>
<tr>
<td>Media Players/Readers</td>
<td>Other forms of human notation</td>
<td>Audio</td>
<td>Periodicals</td>
</tr>
<tr>
<td>Telematics</td>
<td>Possible language specialization</td>
<td>Video</td>
<td>Books</td>
</tr>
<tr>
<td>Set-top Boxes</td>
<td>Formal languages</td>
<td>Graphics</td>
<td>Catalogs</td>
</tr>
<tr>
<td>Appliances</td>
<td>...</td>
<td>Other sensor-based data</td>
<td>Blogs</td>
</tr>
<tr>
<td>Health devices</td>
<td>...</td>
<td>...</td>
<td>Universal Geodata</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>Scientific datasets</td>
</tr>
</tbody>
</table>

\[ \forall d \in D, \forall l \in L, \forall m \in M, \forall c \in C \]
Types of Transparent Processing

- Search, of many forms
- Transformational Communication
- Navigation and Suggestion
- Information Fusion

Some Google Examples:
- Universal search
- Voice Search
- Find Similar, applied to images
- Google Translate, particularly in mash-ups
- Combining images and maps
- Audio transcription
- Images and 3d models
Fluidity Among the Modalities

Text | Voice
---|---
Image | Video

Many New Modalities Here or Coming
Query completion before: Used a fixed dictionary, e.g., in emacs, bash, T9, etc.

Query suggestion today: Model queries with query logs, serve them dynamically

Technical challenges:

- response-time, coverage, freshness, corpus dependency (YouTube, image, mobile)
- domain dependent: rea -> real madrid good suggestion in Spain
- diversity (danger of popularity), filtering out duplicates, inappropriate results, etc.

Impact: Made possible by scale,

- the richer the query log corpus, the better
- the faster the response time, the better
Voice Search

For voice search, just bring the phone to your ear and speak.

Really, no buttons required!

Watch a video to learn more.
Challenges and Rationale for Success

Technically this is very challenging:
- Huge vocabulary
- Variability in accent
- Background noise

What makes this possible:
- Scalable technology
- Data inputs: Query logs, voice logs
- Compute power
## Choosing a data source

<table>
<thead>
<tr>
<th>Web Queries</th>
<th>Speech Transcriptions</th>
<th>Business Listing Databases</th>
</tr>
</thead>
<tbody>
<tr>
<td>huge N</td>
<td>tiny N</td>
<td>wide coverage</td>
</tr>
<tr>
<td>typed, not spoken</td>
<td>perfectly matched to task</td>
<td>little info for popularity</td>
</tr>
<tr>
<td>user in a different setting</td>
<td>slow, expensive, manual(!)</td>
<td>little info for how callers ask</td>
</tr>
<tr>
<td>Google advantage</td>
<td>chicken and egg</td>
<td>where most efforts like this start</td>
</tr>
</tbody>
</table>
Combination works best: utterance ROC curves: incl. rejection

(better)

combination

web queries

speech transcriptions

business listing databases
The Benefits of Unsupervised Training
Transcriptions in Google Voice
Google Translate

**RBMT** – Rules-based machine translation
**SMT** – Statistical (data-driven) machine translation

**2001-2004**
- RBMT – 3rd Party
  - 2001
    - French
    - Italian
    - Spanish
    - Portuguese
    - German
  - 2004
    - Chinese
    - Japanese
    - Korean

**2006**
- RBMT – 3rd Party
- SMT – Google
  - Arabic
  - Chinese (S)
  - Chinese (T)
  - Dutch
  - French
  - German
  - Greek
  - Hungarian
  - Indonesian
  - Italian
  - Japanese
  - Korean

**2007**
- SMT – Google
  - Arabic
  - Bulgarian
  - Catalan
  - Chinese (S)
  - Chinese (T)
  - Croatian
  - Czech
  - Danish
  - Dutch
  - Finnish
  - French
  - German
  - Greek
  - Hindi
  - Indonesian
  - Italian
  - Japanese

**2008**
- SMT – Google
  - Arabic
  - Bulgarian
  - Catalan
  - Chinese (S)
  - Chinese (T)
  - Croatian
  - Czech
  - Danish
  - Dutch
  - Estonian
  - Finnish
  - French
  - Galician
  - German
  - Greek
  - Hebrew
  - Hindi
  - Indonesian
  - Italian
  - Japanese
  - Korean
  - Latvian
  - Lithuanian
  - Norwegian
  - Polish
  - Portuguese
  - Romanian
  - Russian
  - Slovak
  - Slovenian
  - Spanish
  - Swedish
  - Ukrainian
  - Vietnamese

**2009**
- SMT – Google
  - Albanian
  - Arabic
  - Bulgarian
  - Catalan
  - Chinese (S)
  - Chinese (T)
  - Croatian
  - Czech
  - Danish
  - Dutch
  - Estonian
  - Filipino
  - Finnish
  - French
  - Galician
  - German
  - Greek
  - Hebrew
  - Hindi
  - Indonesian
  - Italian
  - Japanese
  - Korean
  - Latvian
  - Lithuanian
  - Norwegian
  - Polish
  - Portuguese
  - Romanian
  - Russian
  - Serbian
  - Slovak
  - Slovenian
  - Spanish
  - Swedish
  - Thai
  - Turkish
  - Ukrainian
  - Vietnamese
  - NEW
    - Afrikaans
    - Belarusian
    - Icelandic
    - Irish
    - Macedonian
    - Malay
    - Swahili
    - Welsh
    - Yiddish

**NEW**
- Maltese
- Norwegian
- Polish
- Portuguese
- Russian
- Serbian
- Slovak
- Slovenian
- Spanish
- Swedish
- Thai
- Turkish
- Ukrainian
- Vietnamese
Web Translation

Google Translate


Korean | English | Translate

Web Translation

“Non-dominantly and non-dominantly”

Chairman of the Joint Chiefs of Staff

Government over the 14th chairmen of the Joint Chiefs of Staff (Bar. 30, 64 years, PB), Commander of 3rd grade

Assignment of the start of a typical irregular recruitment, admissions control a typical letter of recommendation from teachers and students. The key word to the students to write stories that are coming in. Within a short period of the recommendation statement, students write something physically difficult, the students write a letter of recommendation.

Joint Chiefs Chairman Koo yul-kwang notes more: 2009-09-15

Government over the 14th chairmen of the Joint Chiefs of Staff (Bar. 30, 64 years, PB), Commander of 3rd grade.

64 31.12 runners term international change? 2009-06-15

Army Chief of Staff for Personal July 14 middle-aged young person was carried. Yoonchul (64 years, 23 years) was Secretary of Defense Joint Chiefs of Staff Chairman Kim Taehoon. Secretary of Defense for the past three days demoted new appointment military personnel in high places, the predictions for the major said. Group practices, according to the 29-30 posted from 64 years old expected to be in two.

[Following Briefing] 7.6 billion in new growth engines, supporting professional training? 2009-09-15

Labor and Human Resources Development Corporation and the new growth engines in the field of professional training in the program next year, up from 7.6 billion was to support the professional association and an agreement was signed 14 days. Concluded the Korea Financial Investment Agency Association (Green banking industry), Korea IT Industry Promotion Association (the software industry...

64 31.12 runners term international change? 2009-06-15

• Policy and Strategic Planning areas carries Han yul-kwang changpyo export. Field and policy perspectives and insights in the field were combining. The Defense Ministry and army headquarters, strategic department's various business and policy reporting, military bureaucracy that has several management plans. The style is positive and has an affinity. To the Japanese...

Picture: “chastity would like gettyimagesenlarge?” 2009-08-13

Park Young: Tokyo correspondent "From the political future will, like chastity." Yoko

Tsunohata, party of Japan's next prime minister Ichibai the 1st day of a private office in Chicago by...
Cross-language search
In order to help prevent flooding, the MTA has begun to install new street furniture on Hillside Avenue in Queens. The furniture raises the grates above sidewalk level to prevent water, silt, mud, and debris from cascading into tunnels.
Actualizaciones de Noticias

"Abril 22 - CEREMONIA DE INAUGURACIÓN"

La ceremonia se celebrará en el Auditorio A, justo después de la Web Grupo 20º aniversario.

Con el fin de asistir a la ceremonia de inauguración, los participantes deben usar la Conferencia Insignia, de haber completado el proceso de inscripción en el registro de contador situado en el Palacio Municipal de Congresos del 20 de abril en adelante.

Por razones de seguridad y espacio limitado, se aplicarán las siguientes restricciones:

- Contra el registro el 22 de abril se abrirá a 07:45h. A partir de las 10:00 a las 11:30 será cerrado.
- Lugar de acceso, será prohibido de 10:15 a 10:30h. (Incluso a los delegados el uso de una tarjeta de identificación WWW2009)

- 2009/04/07 relacionado de datos disponibles para WWW2009 documentos | Navegar ponencias y pósteres en EPrints
- 2009/04/07 Mejor Papel / Cartel nominaciones anunciadas

WWW2009 Programa de un vistazo

<table>
<thead>
<tr>
<th>20 de abril</th>
<th>21 de abril</th>
<th>22 de abril</th>
<th>23 de abril</th>
<th>24 de abril</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Lunes</strong></td>
<td><strong>Martes</strong></td>
<td><strong>Miercoles</strong></td>
<td><strong>Jueves</strong></td>
<td><strong>Viernes</strong></td>
</tr>
<tr>
<td>Temas</td>
<td>Temas</td>
<td>Temas</td>
<td>Temas</td>
<td>Temas</td>
</tr>
<tr>
<td>Talleres</td>
<td>Talleres</td>
<td>Talleres</td>
<td>Talleres</td>
<td>Talleres</td>
</tr>
<tr>
<td>W4A</td>
<td>W4A</td>
<td>W4A</td>
<td>W4A</td>
<td>W4A</td>
</tr>
</tbody>
</table>
2/28/09: Your Weekly Address
Impact of data - More data is better …

+weblm = LM trained on 219B words of web data

News data
Challenges in Image processing

**Visual Semantics**
- Recognition (people, landmarks, objects)
- Machine Learning

**Correspondence**
- Matching images and videos
- Image mosaicing

**Geometry**
- Ego-motion estimation
- Multi-view stereo

**Image Processing**
- Maps from aerial imagery
- OCR in all the world’s languages
Image Analysis in Image Search

- Image Search helps users find the image they want quickly.
- Understanding the actual content of an image is critical.
- We've been focusing more and more on analyzing images.
- This has been rolling out over the last year.
  - Both as user visible filters
  - Behind the scenes in our back-ends.
- Genre filters like clip art / line drawings / color are great examples
  - [flowers], line drawings, clip art, photo, face
  - [porsche], red, green, yellow, orange, ...

Google
Similar Images in Image Search

- Google has just launched a "Similar Images" feature.
- Accessed by clicking on the similar images link under an image.
- It can also be accessed via preview thumbnails in the result frame.
- We think this will create a major shift in how to search for images.
- Searching for images can now become a navigational experience, where the text (or voice) query acts as a starting point.
Similar Images in Image Search

• A variety of features are used to determine visual similarity.
Similar Images in Image Search

Refine by the content of a specific image.

Different "faces" of Paris...
Content-based retrieval of sound

Auditory front end based on stable models from long ago, with new feature extraction ideas.

PAMIR multi-label retrieval (MLR) for the trainable back-end retrieval.

What about sound segmentation or separation?
“Sapsucker” (woodpecker) representations
Newer Modalities

- 3D Graphics
- Maps and Geography
- Timelines
- Music
- Etc.
Maps/Earth as a Modality
Totally Transparent Processing In-Process…

Many New Modalities Here or Coming
Ideal Distributed Computing
Distributed Computing is 30 years old. But, not very deeply understood until recently

- The application mix
- The true nature of global, open systems:
  - Implications on systems, applications, mix and match.
- The implications of operations at true scale
  - E.g., work on programming & runtimes predominated system mgmt.
- The complexity of the architecture that would result
  - We tend to assume, if we can conceive it, it’s okay.
- The collection of further abstractions that would build on fundamentals then known

- In summary there was a limitation of understanding of (truly) large-scale, open integrated distributed systems
Cloud Computing Architecture

The “Cloud”

Distributed Computing Infrastructure

Operating System

Computer Cluster

Search
Maps
Translate
Apps...

Search
Maps
Translate
Apps...

Search
Maps
Translate
Apps...

Internet
All manner of networking hardware

Billions of endpoints
Optimization Applications

Vast opportunities for applying optimization in large systems:
- There are two traditional fields:
  1. Long Term Planning:
     - Network planning, capacity planning.
  2. Short Term Operational Decisions:
     - Scheduling of tasks, dynamic allocation.
- We want to introduce new types of applications:
  1. On Line Optimization.
  2. Data Checking:
     - Verify consistency and validate data/config file.
  3. Dynamic repair:
     - Find the closest feasible solution after an incident (computer broke down).
     - Continual optimization (i.e. replan after a driver missed a turn).
Fusion Tables: a collaborative database in the cloud

An online database that will make it easier for users to

<table>
<thead>
<tr>
<th>action</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>share data with others</td>
<td>uploading tables and inviting collaborators</td>
</tr>
<tr>
<td>explore their data</td>
<td>using filters &amp; aggregates and visualizing the data on maps, timelines, chars, etc.</td>
</tr>
<tr>
<td>combine datasets</td>
<td>creating merged tables from multiple base tables</td>
</tr>
<tr>
<td>discuss their data</td>
<td>participating in threaded discussions</td>
</tr>
<tr>
<td>publish as web properties</td>
<td>making public datasets and embedding visualization on external pages</td>
</tr>
</tbody>
</table>

- leverage the reliability and performance of the Google infrastructure for data management applications
Fusion Tables Example (1)

<table>
<thead>
<tr>
<th>Country</th>
<th>Year</th>
<th>Total Fresh</th>
<th>Per-capita W</th>
<th>Domestic use</th>
<th>Industrial Use</th>
<th>Agricultural u</th>
<th>Domestic u</th>
<th>Industrial U</th>
<th>Agricultural u</th>
<th>Source</th>
<th>2005 Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Algeria</td>
<td>2000</td>
<td>6.07</td>
<td>164.78</td>
<td>21.98</td>
<td>13.19</td>
<td>64.83</td>
<td>40.61</td>
<td>24.38</td>
<td>119.79</td>
<td></td>
<td>32.85</td>
</tr>
<tr>
<td>Angola</td>
<td>2000</td>
<td>0.35</td>
<td>21.96</td>
<td>23</td>
<td>17</td>
<td>60</td>
<td>5.05</td>
<td>3.73</td>
<td>13.17</td>
<td></td>
<td>15.94</td>
</tr>
<tr>
<td>Benin</td>
<td>2001</td>
<td>0.13</td>
<td>15.4</td>
<td>32</td>
<td>23</td>
<td>45</td>
<td>4.93</td>
<td>3.54</td>
<td>6.93</td>
<td></td>
<td>8.44</td>
</tr>
<tr>
<td>Botswana</td>
<td>2000</td>
<td>0.19</td>
<td>107.34</td>
<td>41</td>
<td>18</td>
<td>41</td>
<td>44.01</td>
<td>19.32</td>
<td>44.01</td>
<td></td>
<td>1.77</td>
</tr>
<tr>
<td>Burkina Faso</td>
<td>2000</td>
<td>0.8</td>
<td>60.47</td>
<td>13</td>
<td>1</td>
<td>86</td>
<td>7.86</td>
<td>0.6</td>
<td>52</td>
<td></td>
<td>13.23</td>
</tr>
<tr>
<td>Burundi</td>
<td>2000</td>
<td>0.29</td>
<td>38.41</td>
<td>16.91</td>
<td>6</td>
<td>77</td>
<td>6.5</td>
<td>2.3</td>
<td>29.58</td>
<td></td>
<td>7.55</td>
</tr>
<tr>
<td>Cameroon</td>
<td>2000</td>
<td>0.99</td>
<td>60.66</td>
<td>18.11</td>
<td>8.05</td>
<td>73.84</td>
<td>10.99</td>
<td>4.88</td>
<td>44.79</td>
<td></td>
<td>16.32</td>
</tr>
<tr>
<td>Cape Verde</td>
<td>2000</td>
<td>0.02</td>
<td>39.22</td>
<td>7</td>
<td>2</td>
<td>91</td>
<td>2.75</td>
<td>0.78</td>
<td>35.69</td>
<td></td>
<td>0.51</td>
</tr>
<tr>
<td>Central African Republic</td>
<td>2000</td>
<td>0.03</td>
<td>7.43</td>
<td>80</td>
<td>16</td>
<td>4.49</td>
<td>5.94</td>
<td>1.19</td>
<td>0.34</td>
<td></td>
<td>4.04</td>
</tr>
<tr>
<td>Chad</td>
<td>2000</td>
<td>0.23</td>
<td>23.59</td>
<td>17</td>
<td>0</td>
<td>83</td>
<td>4.01</td>
<td>0</td>
<td>19.58</td>
<td></td>
<td>9.75</td>
</tr>
<tr>
<td>Comoros</td>
<td>1999</td>
<td>0.01</td>
<td>12.5</td>
<td>48</td>
<td>5</td>
<td>47</td>
<td>6.63</td>
<td>5.88</td>
<td>0.8</td>
<td></td>
<td>0.8</td>
</tr>
<tr>
<td>Democratic Republic</td>
<td>2000</td>
<td>0.36</td>
<td>6.26</td>
<td>53</td>
<td>17</td>
<td>31.47</td>
<td>3.32</td>
<td>1.06</td>
<td>1.97</td>
<td></td>
<td>57.55</td>
</tr>
<tr>
<td>Congo, Republic of</td>
<td>2000</td>
<td>0.03</td>
<td>7.5</td>
<td>59.27</td>
<td>29</td>
<td>12</td>
<td>4.45</td>
<td>2.18</td>
<td>0.9</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>Cote D'Ivoire</td>
<td>2000</td>
<td>0.93</td>
<td>51.24</td>
<td>24</td>
<td>11.82</td>
<td>64.91</td>
<td>12.3</td>
<td>6.06</td>
<td>33.26</td>
<td></td>
<td>18.15</td>
</tr>
<tr>
<td>Djibouti</td>
<td>2000</td>
<td>0.02</td>
<td>25.32</td>
<td>84</td>
<td>0</td>
<td>16</td>
<td>21.27</td>
<td>0</td>
<td>4.05</td>
<td></td>
<td>0.79</td>
</tr>
<tr>
<td>Egypt</td>
<td>2000</td>
<td>68.3</td>
<td>922.6</td>
<td>7.62</td>
<td>6</td>
<td>86</td>
<td>70.27</td>
<td>55.36</td>
<td>793.44</td>
<td></td>
<td>74.03</td>
</tr>
</tbody>
</table>
Fusion Tables Example (2)

Water Usage by Peter Gleick

<table>
<thead>
<tr>
<th>Country</th>
<th>Year</th>
<th>Total Fresh</th>
</tr>
</thead>
<tbody>
<tr>
<td>Algeria</td>
<td>2000</td>
<td>8.07</td>
</tr>
<tr>
<td>Angola</td>
<td>2000</td>
<td>0.35</td>
</tr>
<tr>
<td>Benin</td>
<td>2001</td>
<td>0.13</td>
</tr>
<tr>
<td>Botswana</td>
<td>2000</td>
<td>0.19</td>
</tr>
<tr>
<td>Burkina Faso</td>
<td>2000</td>
<td>0.8</td>
</tr>
<tr>
<td>Burundi</td>
<td>2000</td>
<td>0.29</td>
</tr>
<tr>
<td>Cameroon</td>
<td>2000</td>
<td>0.99</td>
</tr>
<tr>
<td>Cape Verde</td>
<td>2000</td>
<td>0.02</td>
</tr>
<tr>
<td>Central African Republic</td>
<td>2000</td>
<td>0.03</td>
</tr>
<tr>
<td>Chad</td>
<td>2000</td>
<td>0.23</td>
</tr>
<tr>
<td>Comoros</td>
<td>1999</td>
<td>0.01</td>
</tr>
<tr>
<td>Congo, Democratic Republic</td>
<td>2000</td>
<td>0.36</td>
</tr>
<tr>
<td>Congo, Republic of</td>
<td>2000</td>
<td>6.26</td>
</tr>
<tr>
<td>Cote D'Ivoire</td>
<td>2000</td>
<td>7.5</td>
</tr>
<tr>
<td>Djibouti</td>
<td>2000</td>
<td>25.32</td>
</tr>
<tr>
<td>Egypt</td>
<td>2000</td>
<td>68.3</td>
</tr>
</tbody>
</table>

---

Share this table:
- Invited people
  - As viewers - can see and comment on the data
  - As collaborators - can also edit the data
  - As owners - can also invite people to view or collaborate
  - alonhalev@gmail.com

Viewers (1)
- hagonzal@gmail.com

Collaborators (0)

Owners (1)
- jayant@gmail.com

Separate email addresses with commas.
- Send email invitations
- Anyone may view this table

Transferring data from dslist.0.tableserver.fusiontables.vb.borg.google.com...
Fusion Tables Example (3)

<table>
<thead>
<tr>
<th>Country</th>
<th>Year</th>
<th>Total Fresh</th>
<th>Per-capita Wa</th>
<th>Domestic u</th>
<th>Industrial u</th>
<th>Agricultural</th>
<th>Domestic us</th>
<th>Industrial Use</th>
<th>Agricultural</th>
<th>Source</th>
<th>2005 Pops</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belize</td>
<td>2000</td>
<td>0.15</td>
<td>555.56</td>
<td>7</td>
<td>73</td>
<td>20</td>
<td>38.89</td>
<td>405.56</td>
<td>111.11</td>
<td>i</td>
<td>0.27</td>
</tr>
<tr>
<td>Canada</td>
<td>1996</td>
<td>44.72</td>
<td>1385.81</td>
<td>19.56</td>
<td>68.67</td>
<td>11.77</td>
<td>271.05</td>
<td>951.63</td>
<td>163.14</td>
<td>o</td>
<td>32.27</td>
</tr>
<tr>
<td>Singapore</td>
<td>1975</td>
<td>0.19</td>
<td>43.88</td>
<td>45</td>
<td>51</td>
<td>4</td>
<td>19.75</td>
<td>22.38</td>
<td>1.76</td>
<td>c</td>
<td>4.33</td>
</tr>
<tr>
<td>Austria</td>
<td>1999</td>
<td>3.67</td>
<td>448.11</td>
<td>35.09</td>
<td>63.92</td>
<td>0.99</td>
<td>157.23</td>
<td>286.43</td>
<td>4.46</td>
<td>m</td>
<td>8.19</td>
</tr>
<tr>
<td>Belgium</td>
<td>1998</td>
<td>7.44</td>
<td>714.01</td>
<td>13.3</td>
<td>85.47</td>
<td>1.22</td>
<td>94.97</td>
<td>610.29</td>
<td>8.75</td>
<td>k</td>
<td>10.42</td>
</tr>
<tr>
<td>Bulgaria</td>
<td>2003</td>
<td>6.92</td>
<td>895.21</td>
<td>3.02</td>
<td>78.22</td>
<td>18.76</td>
<td>27</td>
<td>700.25</td>
<td>167.97</td>
<td>m</td>
<td>7.73</td>
</tr>
<tr>
<td>Czech Republic</td>
<td>2002</td>
<td>1.91</td>
<td>186.89</td>
<td>40.76</td>
<td>57.1</td>
<td>2.14</td>
<td>76.18</td>
<td>106.71</td>
<td>4.01</td>
<td>m</td>
<td>10.22</td>
</tr>
<tr>
<td>Finland</td>
<td>1999</td>
<td>2.33</td>
<td>443.81</td>
<td>13.65</td>
<td>83.69</td>
<td>2.66</td>
<td>60.56</td>
<td>371.43</td>
<td>11.82</td>
<td>m</td>
<td>5.25</td>
</tr>
<tr>
<td>France</td>
<td>2000</td>
<td>33.16</td>
<td>548.1</td>
<td>15.72</td>
<td>74.48</td>
<td>9.8</td>
<td>86.18</td>
<td>408.2</td>
<td>53.71</td>
<td>m</td>
<td>60.51</td>
</tr>
<tr>
<td>Germany</td>
<td>2001</td>
<td>38.01</td>
<td>459.67</td>
<td>12.36</td>
<td>67.85</td>
<td>19.79</td>
<td>56.8</td>
<td>311.9</td>
<td>90.97</td>
<td>m</td>
<td>82.69</td>
</tr>
<tr>
<td>Hungary</td>
<td>2001</td>
<td>21.03</td>
<td>2082.18</td>
<td>9.24</td>
<td>58.67</td>
<td>32.09</td>
<td>192.48</td>
<td>1221.54</td>
<td>688.16</td>
<td>m</td>
<td>4.45</td>
</tr>
<tr>
<td>Iceland</td>
<td>2003</td>
<td>0.17</td>
<td>566.67</td>
<td>34.02</td>
<td>65.85</td>
<td>0.13</td>
<td>192.8</td>
<td>373.12</td>
<td>0.74</td>
<td>m</td>
<td>0.3</td>
</tr>
<tr>
<td>Ireland</td>
<td>1994</td>
<td>1.18</td>
<td>284.34</td>
<td>22.62</td>
<td>77.36</td>
<td>0.02</td>
<td>64.33</td>
<td>219.96</td>
<td>0.05</td>
<td>m</td>
<td>4.15</td>
</tr>
</tbody>
</table>
Fusion Tables Example (4)

intensity map visualization
Fusion Tables Example (5)

### Renewable Fresh Water Resources

<table>
<thead>
<tr>
<th>Country</th>
<th>Annual Renewable Water Resources (km^3/3yr)</th>
<th>Year of Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Algeria</td>
<td>14.3</td>
<td>1997</td>
</tr>
<tr>
<td>Angola</td>
<td>184.0</td>
<td>1997</td>
</tr>
<tr>
<td>Benin</td>
<td>25.8</td>
<td>2001</td>
</tr>
<tr>
<td>Botswana</td>
<td>14.7</td>
<td>2001</td>
</tr>
<tr>
<td>Burkina Faso</td>
<td>17.5</td>
<td>2001</td>
</tr>
<tr>
<td>Burundi</td>
<td>3.6</td>
<td>1997</td>
</tr>
<tr>
<td>Cameroon</td>
<td>285.5</td>
<td>2003</td>
</tr>
<tr>
<td>Cape Verde</td>
<td>0.3</td>
<td>1990</td>
</tr>
<tr>
<td>Central African Republic</td>
<td>144.4</td>
<td>2003</td>
</tr>
<tr>
<td>Chad</td>
<td>43.0</td>
<td>1997</td>
</tr>
<tr>
<td>Comoros</td>
<td>1.2</td>
<td>2003</td>
</tr>
<tr>
<td>Congo</td>
<td>832.0</td>
<td>1997</td>
</tr>
<tr>
<td>Congo, Democratic Republic (formerly Zaire)</td>
<td>1283.0</td>
<td>2001</td>
</tr>
<tr>
<td>Cote D’Ivoire</td>
<td>81.0</td>
<td>2001</td>
</tr>
<tr>
<td>Djibouti</td>
<td>0.3</td>
<td>1997</td>
</tr>
<tr>
<td>Egypt</td>
<td>86.8</td>
<td>1997</td>
</tr>
<tr>
<td>Equatorial Guinea</td>
<td>26.0</td>
<td>2001</td>
</tr>
<tr>
<td>Eritrea</td>
<td>6.3</td>
<td>2001</td>
</tr>
<tr>
<td>Ethiopia</td>
<td>110.0</td>
<td>1997</td>
</tr>
<tr>
<td>Senegal</td>
<td>124.0</td>
<td>1997</td>
</tr>
</tbody>
</table>

**Source:** Pacific Institute
Fusion Tables Example (6)

- Select table to merge with
- Select key
- Create a new merged table
**Water Resources and Usage**

Pacific Institute, Peter Gleick

<table>
<thead>
<tr>
<th>Country</th>
<th>Annual Renewal</th>
<th>Year of Est</th>
<th>Year</th>
<th>Total Fresh</th>
<th>Per-capita Waters</th>
<th>Domestic use (%)</th>
<th>Industrial Use (%)</th>
<th>Agricultural use (%)</th>
<th>2005 Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Afghanistan</td>
<td>55.9</td>
<td>1997</td>
<td>2000</td>
<td>23.26</td>
<td>546.97</td>
<td>1.70</td>
<td>26.74</td>
<td>11.18</td>
<td>62.08</td>
</tr>
<tr>
<td>Albania</td>
<td>41.7</td>
<td>2001</td>
<td>2000</td>
<td>1.71</td>
<td>546.33</td>
<td>26.74</td>
<td>11.18</td>
<td>62.08</td>
<td>3.13</td>
</tr>
<tr>
<td>Algeria</td>
<td>14.3</td>
<td>1997</td>
<td>2000</td>
<td>6.07</td>
<td>184.78</td>
<td>21.98</td>
<td>13.19</td>
<td>64.83</td>
<td>32.85</td>
</tr>
<tr>
<td>Angola</td>
<td>184.0</td>
<td>1987</td>
<td>2000</td>
<td>0.35</td>
<td>61.96</td>
<td>23</td>
<td>17</td>
<td>60</td>
<td>15.94</td>
</tr>
<tr>
<td>Antigua and Barbuda</td>
<td>0.1</td>
<td>1990</td>
<td>1990</td>
<td>0.01</td>
<td>62.5</td>
<td>20</td>
<td>0.08</td>
<td>20</td>
<td>0.08</td>
</tr>
<tr>
<td>Argentina</td>
<td>814.0</td>
<td>2000</td>
<td>1999</td>
<td>29.19</td>
<td>753.29</td>
<td>17</td>
<td>9.49</td>
<td>74.02</td>
<td>38.75</td>
</tr>
<tr>
<td>Armenia</td>
<td>10.5</td>
<td>1997</td>
<td>2000</td>
<td>2.95</td>
<td>976.82</td>
<td>29.95</td>
<td>4.38</td>
<td>65.67</td>
<td>3.02</td>
</tr>
<tr>
<td>Australia</td>
<td>398.0</td>
<td>1995</td>
<td>2000</td>
<td>24.06</td>
<td>1193.45</td>
<td>14.72</td>
<td>10.03</td>
<td>75.25</td>
<td>20.16</td>
</tr>
<tr>
<td>Austria</td>
<td>84.0</td>
<td>2005</td>
<td>1999</td>
<td>3.67</td>
<td>448.11</td>
<td>35.09</td>
<td>63.92</td>
<td>0.99</td>
<td>8.19</td>
</tr>
<tr>
<td>Azerbaijan</td>
<td>30.3</td>
<td>1997</td>
<td>1999</td>
<td>17.25</td>
<td>2051.13</td>
<td>4.81</td>
<td>27.65</td>
<td>67.53</td>
<td>8.41</td>
</tr>
<tr>
<td>Bahamas</td>
<td>nd</td>
<td>nd</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bahrain</td>
<td>0.1</td>
<td>1997</td>
<td>2000</td>
<td>0.3</td>
<td>410.96</td>
<td>39.65</td>
<td>3</td>
<td>56.81</td>
<td>0.73</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>1210.6</td>
<td>1999</td>
<td>2000</td>
<td>78.4</td>
<td>559.86</td>
<td>3.19</td>
<td>0.85</td>
<td>96.16</td>
<td>141.82</td>
</tr>
<tr>
<td>Barbados</td>
<td>0.1</td>
<td>2003</td>
<td>2000</td>
<td>0.09</td>
<td>333.33</td>
<td>33.37</td>
<td>44.07</td>
<td>22</td>
<td>0.27</td>
</tr>
<tr>
<td>Belarus</td>
<td>58.0</td>
<td>1997</td>
<td>1998</td>
<td>2.79</td>
<td>285.86</td>
<td>23.44</td>
<td>47</td>
<td>30.07</td>
<td>9.76</td>
</tr>
<tr>
<td>Belgium</td>
<td>20.8</td>
<td>2005</td>
<td>1998</td>
<td>7.44</td>
<td>714.01</td>
<td>13.3</td>
<td>85.47</td>
<td>1.22</td>
<td>10.44</td>
</tr>
<tr>
<td>Belize</td>
<td>16.6</td>
<td>2000</td>
<td>2000</td>
<td>0.15</td>
<td>555.56</td>
<td>7</td>
<td>73</td>
<td>20</td>
<td>0.27</td>
</tr>
<tr>
<td>Benin</td>
<td>25.8</td>
<td>2001</td>
<td>2001</td>
<td>0.13</td>
<td>15.4</td>
<td>32</td>
<td>23</td>
<td>45</td>
<td>8.44</td>
</tr>
<tr>
<td>Bhutan</td>
<td>95.0</td>
<td>1997</td>
<td>2000</td>
<td>0.43</td>
<td>199.07</td>
<td>5</td>
<td>1.1</td>
<td>94</td>
<td>2.16</td>
</tr>
<tr>
<td>Bolivia</td>
<td>622.5</td>
<td>2000</td>
<td>2000</td>
<td>1.44</td>
<td>155.86</td>
<td>13.29</td>
<td>7</td>
<td>81</td>
<td>9.18</td>
</tr>
</tbody>
</table>
### Water Resources and Usage

Pacific Institute, Peter Gleick

#### Table: Water Resources and Usage

<table>
<thead>
<tr>
<th>Country</th>
<th>Annual Renewal</th>
<th>Year of Est.</th>
<th>Year</th>
<th>Total Fresh</th>
<th>Per-capita Water</th>
<th>Domestic use %</th>
<th>Industrial Use (%)</th>
<th>Agricultural use (%)</th>
<th>2005 Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Afghanistan</td>
<td>65.0</td>
<td>1997</td>
<td>2000</td>
<td>23.26</td>
<td>778.97</td>
<td>1.78</td>
<td>0.01</td>
<td>98.21</td>
<td>29.86</td>
</tr>
<tr>
<td>Albania</td>
<td>41.7</td>
<td>2001</td>
<td>2000</td>
<td>1.71</td>
<td>546.33</td>
<td>26.74</td>
<td>11.18</td>
<td>62.08</td>
<td>3.13</td>
</tr>
<tr>
<td>Algeria</td>
<td>14.3</td>
<td>1997</td>
<td>2000</td>
<td>6.07</td>
<td>184.78</td>
<td>21.98</td>
<td>13.19</td>
<td>64.83</td>
<td>32.85</td>
</tr>
<tr>
<td>Angola</td>
<td>21.96</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antigua and Barbuda</td>
<td>62.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>23</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Argentina</td>
<td>753.29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>17</td>
<td>9.49</td>
<td>74.02</td>
<td>38.75</td>
</tr>
<tr>
<td>Armenia</td>
<td>976.82</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>29.95</td>
<td>4.38</td>
<td>65.67</td>
<td>3.02</td>
</tr>
<tr>
<td>Australia</td>
<td>1193.45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>14.72</td>
<td>10.03</td>
<td>75.26</td>
<td>20.16</td>
</tr>
<tr>
<td>Austria</td>
<td>448.11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>35.09</td>
<td>63.92</td>
<td>80.99</td>
<td>8.19</td>
</tr>
<tr>
<td>Azerbaijan</td>
<td>2051.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4.81</td>
<td>27.65</td>
<td>67.53</td>
<td>8.41</td>
</tr>
<tr>
<td>Bahamas</td>
<td>0.1</td>
<td>1997</td>
<td>2000</td>
<td>0.3</td>
<td>410.96</td>
<td>39.65</td>
<td>3</td>
<td>56.81</td>
<td>0.73</td>
</tr>
<tr>
<td>Bahrain</td>
<td>1210.6</td>
<td>1999</td>
<td>2000</td>
<td>79.4</td>
<td>559.86</td>
<td>3.19</td>
<td>0.65</td>
<td>96.16</td>
<td>141.82</td>
</tr>
<tr>
<td>Barbados</td>
<td>0.1</td>
<td>2003</td>
<td>2000</td>
<td>0.09</td>
<td>333.33</td>
<td>33.37</td>
<td>44.07</td>
<td>22</td>
<td>0.27</td>
</tr>
<tr>
<td>Belarus</td>
<td>58.0</td>
<td>1997</td>
<td>1998</td>
<td>2.79</td>
<td>285.86</td>
<td>23.44</td>
<td>47</td>
<td>30.07</td>
<td>9.76</td>
</tr>
<tr>
<td>Belgium</td>
<td>20.8</td>
<td>2005</td>
<td>1998</td>
<td>7.44</td>
<td>714.01</td>
<td>13.3</td>
<td>85.47</td>
<td>1.22</td>
<td>10.42</td>
</tr>
<tr>
<td>Belize</td>
<td>18.6</td>
<td>2000</td>
<td></td>
<td>0.15</td>
<td>555.56</td>
<td>7</td>
<td>73</td>
<td>20</td>
<td>0.27</td>
</tr>
<tr>
<td>Benin</td>
<td>25.8</td>
<td>2001</td>
<td>2001</td>
<td>0.13</td>
<td>15.4</td>
<td>32</td>
<td>45</td>
<td>23</td>
<td>8.44</td>
</tr>
<tr>
<td>Bhutan</td>
<td>95.0</td>
<td>1987</td>
<td>2000</td>
<td>0.43</td>
<td>199.07</td>
<td>5</td>
<td>1.1</td>
<td>94</td>
<td>2.16</td>
</tr>
<tr>
<td>Bolivia</td>
<td>621.5</td>
<td>2000</td>
<td>2000</td>
<td>1.44</td>
<td>156.86</td>
<td>13.29</td>
<td>7</td>
<td>81</td>
<td>9.1</td>
</tr>
</tbody>
</table>

*Cell value: 184.0*

**[jayant](mailto:jayant@gmail.com)** (17 seconds ago)

Angola has a lot of water for a country its size!

And I thought it was in the Kalahari!
Where is the research?

Motivation: commercial databases have become far too intimidating for common users
• Not being used even when there is a fit

Redefining how structured data management can be done (on the Web)
• Focusing on supporting common user activities
• Making data management primitives easily available
• Blurring the notion of database boundaries
Excitement in Distributed Systems

- Size of user community
- Storage Scale (requiring various characteristics)
  - E.g., security, privacy, availability,
- Processing Scale
  - High performance batch processing
  - High throughput
  - Low latency
- Rapid dynamics
- Highly variable end-user devices

- Communication Scale
  - Bandwidth
  - Endpoints
- Efficiency
  - Equipment
  - Communication
  - Power
  - Management
- Extensibility
- Compliance
- *And more to come, no doubt*
Ideal Distributed Computing

*Large networked clusters grow in a fully distributed world*

- Arbitrarily high volume transactions
- And, various, partitionable batch process for learning, fusion, etc.
- Network
  - Response-time and bandwidth as needed
- Cluster Processing, or “Cloud Computing” growing ever larger
  - Massive parallelism to hit sweet spot of capital & operating efficiency
- Distributed computing
  - Data sharing, function shipping, as needed
  - Connected and disconnected operation, as seamless as possible
  - Auto balancing of loads between client device and cloud elements
  - Emphasis on manageability (newly, to handle consumers’ many endpoints)
- Significant efficiency gains
Hybrid, Not Artificial, Intelligence
Hybrid, not Artificial, Intelligence

• “Artificial Intelligence” aimed at having computers as capable as people, often in very broad problem domains

• It has proven more useful for computers rather:
  – To extend the capability of people, not in isolation
  – To focus on more specific problem areas

• Aggregation of user responses has proven extremely valuable in learning

• Examples
  – Feedback in Information Retrieval; e.g., in ranking or spelling correction
  – Machine learning; e.g., image content analysis, speech recognition with semi-supervised learning

• Another example of bottom up successes
My Long-held View on Semantics, Syntax, & Learning

• Large scale learning has proven surprisingly effective

• Learning is occurring over increasingly variegated features:
  – Both Semantic
  – And Syntactic, and generated in multiple ways

• In my WWW 2002 (Architecting Knowledge Middleware) and Semantic Web 2005 Keynotes, I referred to this as The Combination Hypothesis

• Today, I would refine this as the combination of approaches and learning from people.
Fantastic Opportunities Abound
Just a Few Opportunities

- New Interfaces and applications with mass customization
  - Implications on every vertical
  - Examples: health, government, entertainment
- Virtually unlimited data storage
- Ever improving system “understanding”
- Increasingly fluid partnership between people and computation
- Fundamental changes in the methods of science
- Opportunities for optimization in many more domains

There are no real limits
We Desire A Stronger Relationship with Academe (1)

Google tremendously values talented people and education

• We have preferred a bottom-up approach:
  – Google collaborating with faculty and students for *mutual* benefit
    • Google: Knowledge of challenging problems, skilled employees, opportunities for internship/sabbatical, money
    • University: Faculty and student skills, breadth/depth of perspective
  – Internships
  – Visiting Faculty
• University Research Grants with moderate funding
  – Roughly 100 worldwide
  – Recommend proposals developed with advice/perspective of Google employees
CluE in the Search for Data-Intensive (Cloud) Computing

Unique relationship between Google, IBM and NSF allows academic computing research community to access large-scale computer cluster for cloud computing.
Google Relationship with Academe (cont.)

• Check out growing sites:
  – //research.google.com
  – //research.google.com/university
  – //code.google.com
  – //code.google.com/edu

• Technology Round Table Videos (new!)
  – http://research.google.com/roundtable/
Summary: Innovation at Google

- Strong commitment to broadly advancing technology due to our mission: some grand challenge problems

- Leverage
  - Scale in processing and information, and Usage
  - Learning and Empiricism

- We try to minimize the distance between Research and Development
  - Recognize that putting ideas into production is often as challenging and fun as inventing the idea

- Google will attempt to use its capabilities beneficially to foster research, education, and advancement broadly

- Go to http://research.google.com for more information and resources
Thank you!