A Brief History of Information*

James Caverlee
Assistant Professor
Computer Science and Engineering
Texas A&M University

*(plus lots of unsolicited advice)*
• Our Research:
  • Web-scale information management
  • Distributed data-intensive systems
  • Social computing.
• Enable efficient and trustworthy information sharing and knowledge discovery over dynamic, heterogeneous, and massive-scale networked information systems.
Dynamic, Heterogeneous, and Massive-Scale Networked Information Systems
How do you find answers to questions?
3rd Century BC

• Library of Alexandria
• 500,000 volumes
• catalogs and classifications
• “controlled vocabularies”
We could not calculate directions between Egypt (Library of Alexandria) and College Station, TX 77843.
Uh oh ...
Brief History of “Information”

• First concordance of the Bible

• invention of the inverted list data structure

“River”:
Before the Web ... 

- Vannevar Bush’s Memex 1945
No Internet ... we had Sneakernet
Okay ... so the world has changed a lot
Not just devices ... also how they are connected: The Web is Ubiquitous
Every two days now we create as much information as we did from the dawn of civilization up until 2003, according to Google’s Eric Schmidt. That’s something like five exabytes of data, he says. [TechCrunch]
Democratization of Computation
Big Data + Big Computation
One example project ...
Eyjafjallajökull, April 2010
Air Travel Chaos Deepens Into Weekend

Passengers camped out as they waited for the resumption of air travel in Frankfurt on Friday. More Photos »

By ALAN COWELL, NICOLA CLARK and LIZ ROBBINS
Published: April 15, 2010

PARIS — The menacing cloud of ash from a volcano in Iceland moved eastward across Northern Europe, grounding hundreds of flights and stranding thousands of travelers. 

This article is by Alan Cowell, Liz Robbins and Nicola Clark.

Multimedia

PARIS — The menacing cloud of ash from a volcano in Iceland moved eastward across Northern Europe, grounding hundreds of flights and stranding thousands of travelers.
Results for #roadsharing

1 more results since you started searching. Refresh to see them.

roadsharing: #getmehome #roadsharing #ashtag From Brussels, Belgium To Rome, Italy http://tinyurl.com/y4n7o8s (expand)
2 minutes ago from API · Reply · View Tweet

soniasoy: #getmehome #roadsharing #ashtag From Manila (PH) to Valencia (SP)? xD Maybe boatsharing?
2 minutes ago from TweetDeck · Reply · View Tweet

roadsharing: #getmehome #roadsharing #ashtag From Milan, Italy To Madrid, Spain http://tinyurl.com/y4gwun (expand)
4 minutes ago from API · Reply · View Tweet

rs_twit: visit roadsharing.com for lifts from FR to ES #roadsharing #ashtag #stranded
4 minutes ago from API · Reply · View Tweet

roadsharing: #getmehome #roadsharing #ashtag From Stuttgart, Germany To Venice, Italy http://tinyurl.com/y3zpp9e (expand)
6 minutes ago from API · Reply · View Tweet

rs_twit: visit roadsharing.com for lifts from PT to FR #roadsharing #ashtag #stranded
6 minutes ago from API · Reply · View Tweet

Slide from Corinne Weisgerber
Shift from “community” to “crowd”

• Established, well-formed **communities**

• Membership is typically static and long-lived

• Often, explicitly declared

• Fundamental shift to **crowds**

• Ad-hoc collections of users reflecting real-time interests and affiliations

• Organic, short-lived, self-organized

• Often, implicitly defined
Crowds cut across application domains

- Identification and tracking of online “hotspots” as they arise in real-time
- Disasters, terror attacks, civil uprisings
- Social media analytics, advertising
- Emergency informatics
- Public health
- Politics and governance
- ...

• Need a common infrastructure and research agenda:

• **How?**
  
  • How do crowds form and evolve? How do we detect and track the dynamics of crowds on the real-time web?

  • **Challenge:** 100s of millions of users + highly-dynamic/bursty interactions place huge demands on traditional methods.

• **Where?**
  
  • How are users and crowds distributed across space? What are the efficient and effective methods of geolocating and tracking users and crowds on the real-time web?

  • **Challenge:** slow adoption of location-based social media features (e.g., <26% of Twitter users include city name)
Crowds: How?
Crowds: Where?

- Location + geography impact community formation [e.g., Liben-Nowell et al. Geographic routing in social networks]

- Good news: Location-based social media (Facebook Places, Google Latitude, Foursquare) on the rise!
On average we find that the location estimates converge quickly (needing just 100s of tweets), placing 51% of Twitter users within 100 miles of their actual location.

• ... one more thing
Three easy steps to success (?)

- Go to talks
- Own your education
- Write, write, write
The End of Anonymity, The Beginning of Privacy

Vitaly Shmatikov
Associate Professor
Department of Computer Science
University of Texas at Austin

4:10 p.m., Monday March 28, 2011
Room 124, Bright Building

Abstract

The Internet economy relies on the collection and aggregation of personal data on an ever-increasing scale. Information about our tastes, purchases, searches, browsing history, social relationships, health history, genetics, and so forth is shared with advertisers, marketers, and researchers, who use it to predict individual behavior and provide personalized product offerings, recommendations, and even clinical treatments.

I will survey privacy issues caused by massive aggregation of personal information. After demonstrating that the existing methods for "anonymizing" the data fail to provide meaningful privacy protection, I will describe new approaches to privacy-preserving computation. This includes Airavat, a new system for large-scale data analysis which integrates mandatory access control and differential privacy.
TOMORROW HAPPENS HERE.

danah boyd gives the Opening Remarks at the 2010 SXSW Interactive Festival.

What can we say? You came, and you brought it. We want to thank each and every one of you for helping to make SXSW® Interactive 2010 one of our best years ever. The panels, the parties, the 13th Annual Web Awards, the ScreenBurn at SXSW® Arcade, the Film and Interactive Trade Show and Exhibition, and Microsoft
The Python Tutorial

Release: 2.6
Date: March 13, 2010

Python is an easy to learn, powerful programming language. It has efficient high-level data structures and a simple but effective approach to object-oriented programming. Python's elegant syntax and dynamic typing, together with its interpreted nature, make it an ideal language for scripting and rapid application development in many areas on most platforms.

The Python interpreter and the extensive standard library are freely available in source or binary form for all major platforms from the Python Web site, http://www.python.org/, and may be freely distributed. The same site also contains distributions of and pointers to many free third party Python modules, programs and tools, and additional documentation.

The Python interpreter is easily extended with new functions and data types implemented in C or C++ (or other languages callable from C). Python is also suitable as an extension language for customizable applications.

This tutorial introduces the reader informally to the basic concepts and features of the Python language and system. It helps to have a Python interpreter handy for hands-on experience, but all examples are self-contained, so the tutorial can be read off-line as well.

For a description of standard objects and modules, see The Python Standard Library, The Python Language Reference gives a more formal definition of the language. To write extensions in C or C++, read Extending and Embedding the Python Interpreter and Python/C API Reference Manual. There are also several books covering Python in depth.

This tutorial does not attempt to be comprehensive and cover every single feature, or even every commonly used feature. Instead, it introduces many of Python's most noteworthy features, and will give you a good idea of the power and flexibility of the language.
iPhone Application Development (Winter 2010)

Alan Caniniastro and Josh Shaffer

Last Modified: March 19, 2010
Tracks in Tracks: 44

DESCRIPTION

Tools and APIs required to build applications for the iPhone platform using the iPhone SDK. User interface designs for mobile devices and unique user experiences using multitouch technologies. Object-oriented design using model-view-controller pattern, memory management, Objective-C programming language. iPhone APIs and tools including Xcode, Interface Builder and Instruments on Mac OS X. Other topics include: core animation, Touch ID networking, mobile device power management and performance considerations.

Prequisites: C language and programming experience at the level of 10681 or X. Recommended: UNIX, object-oriented programming, graphical tools.

Offered by Stanford’s School of Engineering, the course will last ten weeks and include both the lecture videos and PDF documents. A new lecture will be posted each week after each class meeting. Subscribe to this course, and automatically receive new lectures as they become available.

Released with a Creative Commons BY-NC-ND license.

<table>
<thead>
<tr>
<th>Name</th>
<th>Time</th>
<th>Artist</th>
<th>Album</th>
<th>Price</th>
<th>GET MOVIE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Introduction to Mac...</td>
<td>47:07</td>
<td>Alan Caniniastro</td>
<td>iPhone Application Development (CS:193...</td>
<td>Free</td>
<td>GET MOVIE</td>
</tr>
<tr>
<td>2. Lecture 1 Slides (January 5,...</td>
<td>1:13:31</td>
<td>Josh Shaffer</td>
<td>iPhone Application Development (CS:193...</td>
<td>Free</td>
<td>GET MOVIE</td>
</tr>
<tr>
<td>2. Objective-C and Foo...</td>
<td></td>
<td>Alan Caniniastro</td>
<td>iPhone Application Development (CS:193...</td>
<td>Free</td>
<td>GET MOVIE</td>
</tr>
<tr>
<td>4. Building an Applicat...</td>
<td>55:44</td>
<td>Alan Caniniastro</td>
<td>iPhone Application Development (CS:193...</td>
<td>Free</td>
<td>GET MOVIE</td>
</tr>
<tr>
<td>5. Custom Classes, Ob...</td>
<td>1:07:22</td>
<td>Alan Caniniastro</td>
<td>iPhone Application Development (CS:193...</td>
<td>Free</td>
<td>GET MOVIE</td>
</tr>
<tr>
<td>6. Lecture 1 Slides (January 12,...</td>
<td></td>
<td>Josh Shaffer</td>
<td>iPhone Application Development (CS:193...</td>
<td>Free</td>
<td>GET MOVIE</td>
</tr>
<tr>
<td>7. iPhone Application Development (CS:193...</td>
<td></td>
<td>Alan Caniniastro</td>
<td>iPhone Application Development (CS:193...</td>
<td>Free</td>
<td>GET MOVIE</td>
</tr>
<tr>
<td>8. Lecture 4 Slides (January 14,...</td>
<td></td>
<td>Alan Caniniastro</td>
<td>iPhone Application Development (CS:193...</td>
<td>Free</td>
<td>GET MOVIE</td>
</tr>
<tr>
<td>9. Xcode, Finding Help,...</td>
<td>44:56</td>
<td>Paul Salzman and...</td>
<td>iPhone Application Development (CS:193...</td>
<td>Free</td>
<td>GET MOVIE</td>
</tr>
</tbody>
</table>
Google App Engine

Run your web apps on Google's infrastructure.
Easy to build, easy to maintain, easy to scale.

Java™ Language Support
App Engine recently unveiled its second language: Java. This release includes our Java runtime, integration with Google Web Toolkit, and a Google Plugin for Eclipse, giving you an end-to-end Java solution for AJAX web applications. The Java runtime is now available for anyone to use, so please give it a try and send us your feedback.

- Get the full scoop in our blog post.
- Click over to YouTube to watch our Campfire One announcements.

Getting Started
1. Sign up for an App Engine account.
2. Download the App Engine SDK.
3. Read the Getting Started Guide.
4. Check out sample apps in the App Gallery.

Watch and Learn
Developing and deploying on Google App Engine. Watch Now

links from AppEngine.reddit.com
Pre-release 1.3.2 SDKs available
We host a large number of technical talks at Google, many of which are videotaped and made available for external viewing. You can search our Tech Talks below, and we’ve also made a selection of other talks at Google available.

Google Technology Roundtable

Hear from Google Engineers and Scientists

- Large Scale Search System Infrastructure and Search Quality, Jeff Dean, and Amit Singhal
- Map Reduce, Jeff Dean, Sanjay Ghemawat, Jerry Zhao, and Matt Austern
- Applications of Human Language Technology, Mike Cohen and Franz Och

Latest additions

Conscious Understanding: What is its Physical Basis?
Google Tech Talk March 10, 2010 ABSTRACT Presented by Sir Roger Penrose. Powerful arguments can be given, to support the case that the quality of ...

Mar 17, 2010
• Blog?

• Text?

• With a critic is even better.
  • In this class
  • Where else?
Tar Heels invade Duke for another Cameron clash
By James Caverlee
March 1, 1996

By JAMES CÄVERLEE

For the past four years, the Duke-North Carolina basketball rivalry has produced some of the most memorable games in the country. From the thriller pitting the top two ranked teams in the country two years ago to a double-overtime battle last season to this year's one-point game, the rivalry has been unmatched.

Except for one fact: ever since their first meeting in the 1992-93 season, the Blue Devils have lost every game between the two storied programs. That's six losses in a row to the Tar Heels, matching the longest such streak in head coach Mike Krzyzewski's career.

But Sunday afternoon, the Blue Devils (18-10, 8-7 in the Atlantic Coast Conference) will get a chance to end that streak and lay claim to a share of third place in the conference in Cameron Indoor Stadium before a national television audience.

"This is Duke-Carolina," sophomore Ricky Price said. "They're our arch-rivals and we haven't beaten them in a while, and that's enough for me."
Acknowledgments

2010 Young Faculty Award

Thanks to my students: Zhiyuan Cheng, Brian Eoff, Chiao-Fang Hsu, Krishna Kamath, Said Kashoob, Jeremy Kelley, Elham Khabiri, and Kyumin Lee

For more info: Google “caverlee”