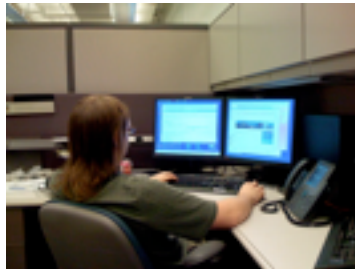


# Cybercasing the Joint: On the Privacy Implications of Geo-Tagging



**Gerald Friedland, Robin Sommer**  
**International Computer Science Institute**  
**Berkeley, CA**  
**[fractor,robin@icsi.berkeley.edu](mailto:fractor,robin@icsi.berkeley.edu)**

# What is Geotagging?



Jan\_Joubert's\_Gat\_Bridge.jpg (2/3)

File Edit View Go Bookmarks Image Tools Help

Jan\_Joubert's\_Gat\_Bridge.jpg Properties

General Comment Photo Data (EXIF) Histogram

GPS Coordinates

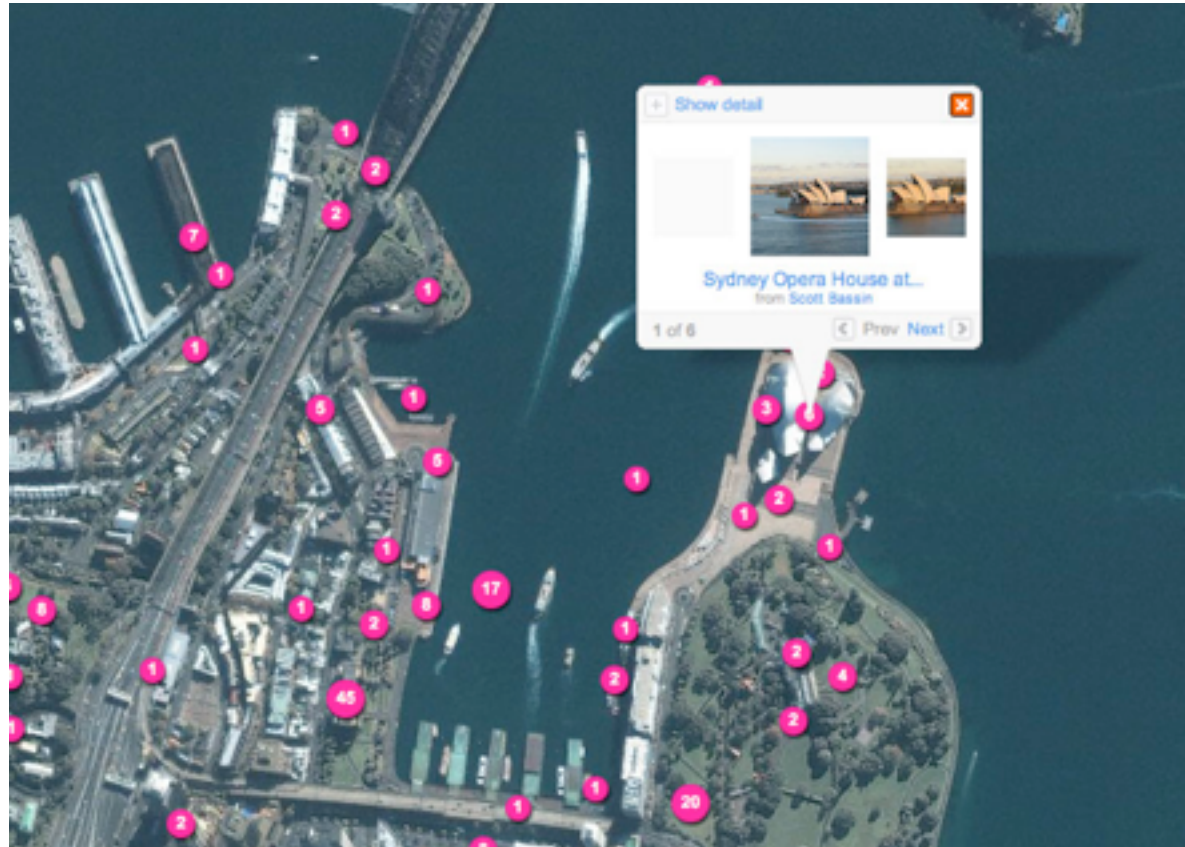
GPSLatitudeRef	S
GPSLatitude	33.00, 56.00, 17.82
GPSLongitudeRef	E
GPSLongitude	19.00, 9.00, 39.85
GPSAltitudeRef	0x00
GPSAltitude	0.00
GPSVersionID	0x00, 0x00, 0x02, 0x02

Image Structure

← Previous    Next →    ✕ Close

Source: Wikipedia

# Why Geo-Tagging?



**Allows easier clustering of photo and video series as well as additional services.**

# Why Geo-Tagging?

Part of location-based service hype:





# Support for Geo-Tags

Social media portals provide APIs to connect geo-tags with metadata, accounts, and web content.

Portal	%	Total
YouTube (estimate)	3.0	3M
Flickr	4.5	180M

Allows easy search, retrieval, and ad placement.



# Problems

People are unaware of

1. geo-tagging
2. resulting inference possibilities:
  - a. high resolution of sensors
  - b. large amount of geo-tagged data
  - c. easy-to-use APIs allow fast retrieval



# Related Work

A cartoon illustration of a robber with a beard, wearing a tan cap, a black mask covering his eyes, and a black and white striped shirt. He is carrying a large green money bag with a white dollar sign on it. The character is set against a light blue circular background.

**PLEASE ROB ME**

**Raising awareness about over-sharing**

Check out our [quest blog post](#) on the CDT website.

A stylized map graphic with white lines on a blue background. Two red location pins are placed on the map, each with a white circle containing a black 'X' inside its top, indicating a location of interest or a warning.

“Be careful when using social location sharing services, such as FourSquare.”



# Related Work


**I Can Stalk U**  
Raising awareness about inadvertent information sharing


Home   How   Why   About Us   Contact Us

If you came from the Social Media White Noise Podcast, see [our response](#)

---

**What are people *really* saying in their tweets?**

 [albeeta](#): I am currently nearby 10456 Riverside Dr Los Angeles CA  
less than a minute ago · [Map Location](#) · [View Tweet](#) · [View Picture](#) · [Reply to albeeta](#)

 [chrisbeanland](#): I am currently nearby <http://maps.google.com/?q=51.5385,-0.133>  
less than a minute ago · [Map Location](#) · [View Tweet](#) · [View Picture](#) · [Reply to chrisbeanland](#)

**Links**

- Mayhemic Labs
- PaulDotCom
- SANS ISC
- Electronic Frontier Foundation
- Center for Democracy & Technology

[How did you find me?](#)

Mayhemic Labs, June 2010: “Are you aware that Tweets are geo-tagged?”





# Can you do real harm?

- **Cybercasing: Using online (location-based) data and services to mount real-world attacks.**
- **Three Case Studies:**



# Case Study 1: Twitter

- Pictures in Tweets can be geo-located
- From an undisclosed celebrity we found:
  - Home location (several pics)
  - Where the kids go to school
  - The place where he/she walks the dog
  - “Secret” office
- Systematic search: [picfog.com](http://picfog.com)

# Celebs unaware of Geo-Tagging

twitpic

[Click here to login or](#)



Working with the very talented Adam Hamilton on creating a new album. My best, Bill

Source: ABC News <sup>11</sup>



# Celebs unaware of Geotagging

## EXIF IFD1

- Compression {0x0103} = JPEG compression (6)
- X-Resolution {0x011A} = 4718592/65536 ==> 72
- Y-Resolution {0x011B} = 4718592/65536 ==> 72
- X/Y-Resolution Unit {0x0128} = inch (2)
- Y/Cb/Cr Positioning (Subsampling) {0x0213} = centered / center of pixel array (1)
- Embedded thumbnail image:



## EXIF GPS IFD

- GPS Version ID {0x00} = 0x02,0x02,0x00,0x00
- GPS Latitude Reference {0x01} = N
- GPS Latitude {0x02} = 34/1,12/1,3/1 [degrees, minutes, seconds] ==> 34° 12' 3" == 34.200833°
- GPS Longitude Reference {0x03} = W
- GPS Longitude {0x04} = [redacted] [degrees, minutes, seconds] ==> [redacted]' == [redacted]





# Google Maps shows Address...

Google maps

34.200833, [redacted]

Search Maps

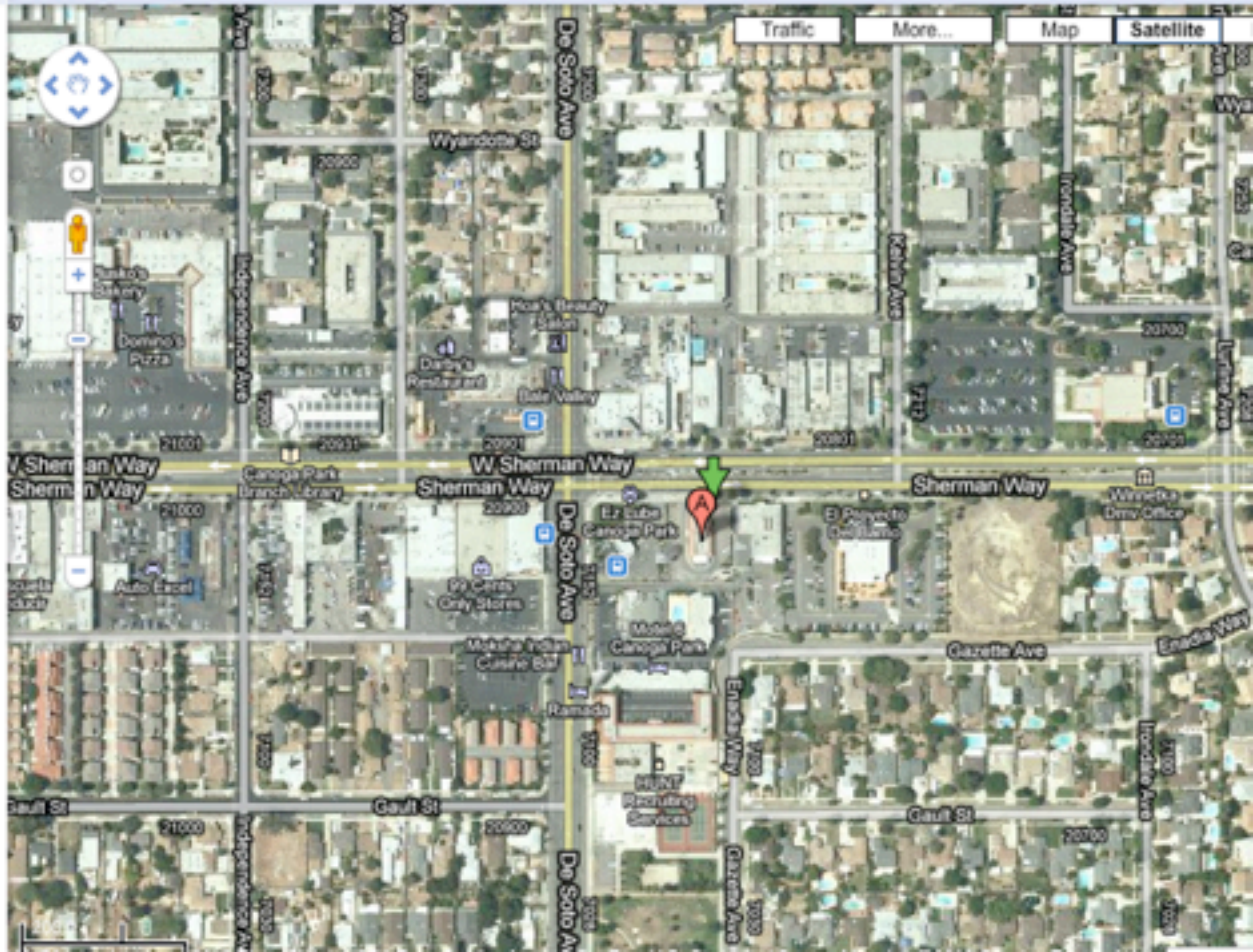
Show search options

[Get Directions](#) [My Maps](#)

[Print](#) [Send](#)



[Directions](#) [Search nearby](#) [Save to...](#) [more](#)



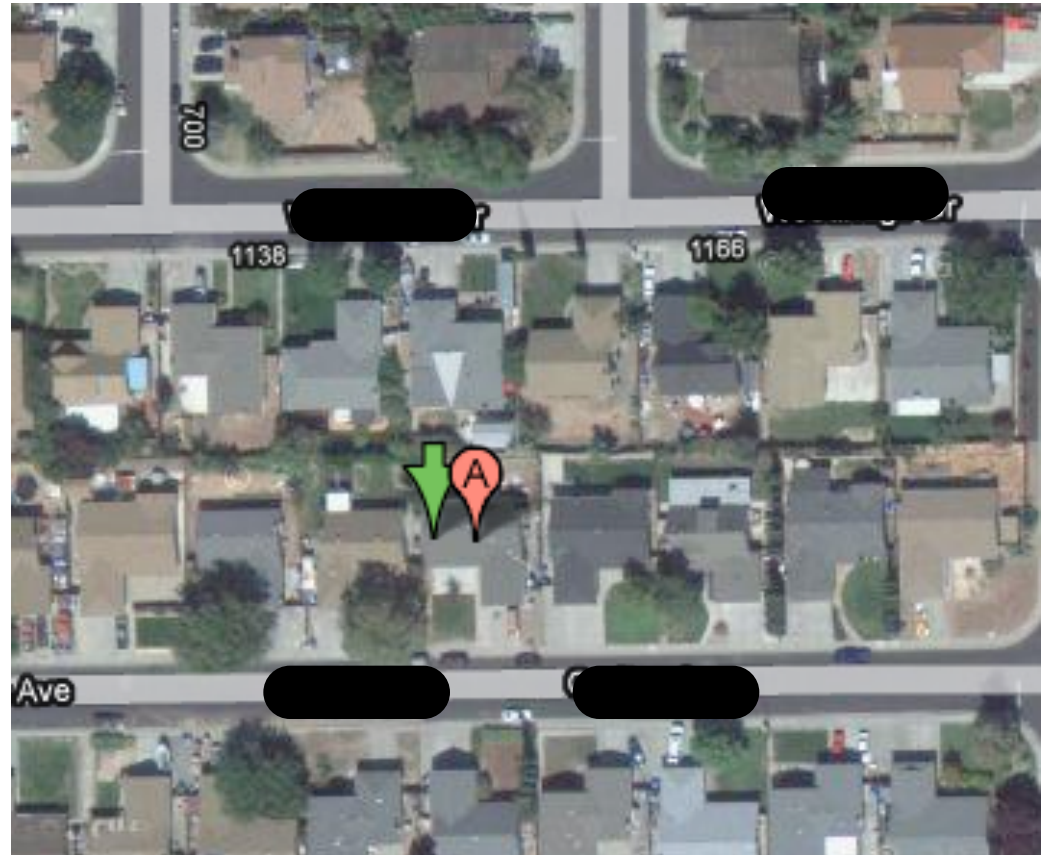




# Case Study 2: Craigslist

- Many ads with geo-location otherwise anonymized
- Sometimes selling high-valued goods, e.g. cars, diamonds
- Sometimes “call Sunday after 6pm”
- Multiple photos allow interpolation of coordinates for higher accuracy

# Craigslist: Real Example



# Geo-Tagging Resolution



iPhone 3G picture



Google Street View

Measured accuracy:  $\pm 1\text{m}$

# People are Unaware of Geo-Tagging

“For Sale” section of Bay Area Craigslist.com:  
4 days: 68729 pictures total, 1.3% geo-tagged

#	<i>Model</i>	#	<i>Model</i>
414	iPhone 3G	6	Canon PowerShot SD780
287	iPhone 3GS	3	MB200
98	iPhone	2	LG LOTUS
32	Droid	2	HERO200
26	SGH-T929	2	BlackBerry 9530
20	Nexus One	1	RAPH800
9	SPH-M900	1	N96
9	RDC-i700	1	DMC-ZS7
6	T-Mobile G1	1	BlackBerry 9630



# Case Study 3: YouTube

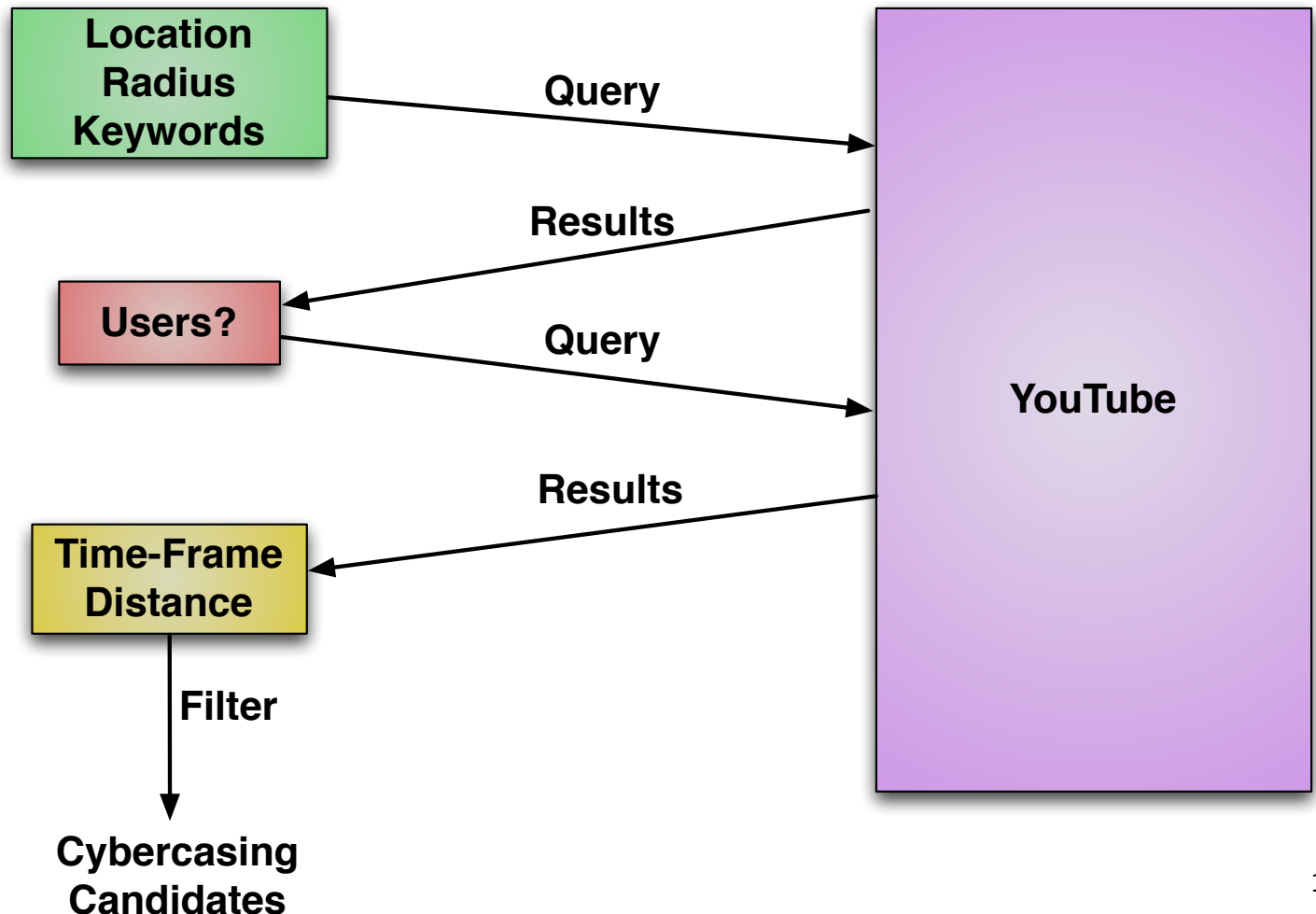
- Once data is published, the Internet keeps it (in potentially many copies).
- APIs are easy to use and allow quick retrieval of large amounts of data
- Even simple inference algorithms (across different websites) allow for cybercasings.

Can we find people on vacation in YouTube?



# Cybercasing on YouTube

Experiment: Cybercasing using the YouTube API (240 lines in Python)





# Cybercasing on YouTube

## Input parameters

**Location:** 37.869885, -122.270539

**Radius:** 100km

**Keywords:** kids

**Distance:** 1000km

**Time-frame:** this\_week



# Cybercasing on YouTube

## Output

Initial videos: 1000 (`max_res`)

➡ User hull: ~50k videos

➡ Vacation hits: 106

➡ Cybercasing targets: >12



# Cybercasing on YouTube

INTI  
COMF  
I N :

First Day of [REDACTED] Vacation

1 videos

Subscribe

O  
In  
↓  
↓  
↓





# Solutions?





# Solutions?

- Better Education
- More secure default values
- Blurring
- Scrubbing
- Privacy-preserving APIs and policies

# Proposal: Opt-In with Choice of Accuracy



Mockup of a privacy-improved iPhone dialog

# Conclusion

- Geo-location offers great opportunities and we should continue to explore them
- However it can pose real-world risks
- Therefore, we should:
  - Raise the awareness on privacy issues
  - Discuss policies and interfaces



# Questions?

- Are you concerned?
- What is a good trade-off between privacy and utility?
- How can we design policies and APIs to implement the trade-off?