

# Google

Steward of Sensitive Information



Google Search

I'm Feeling Lucky

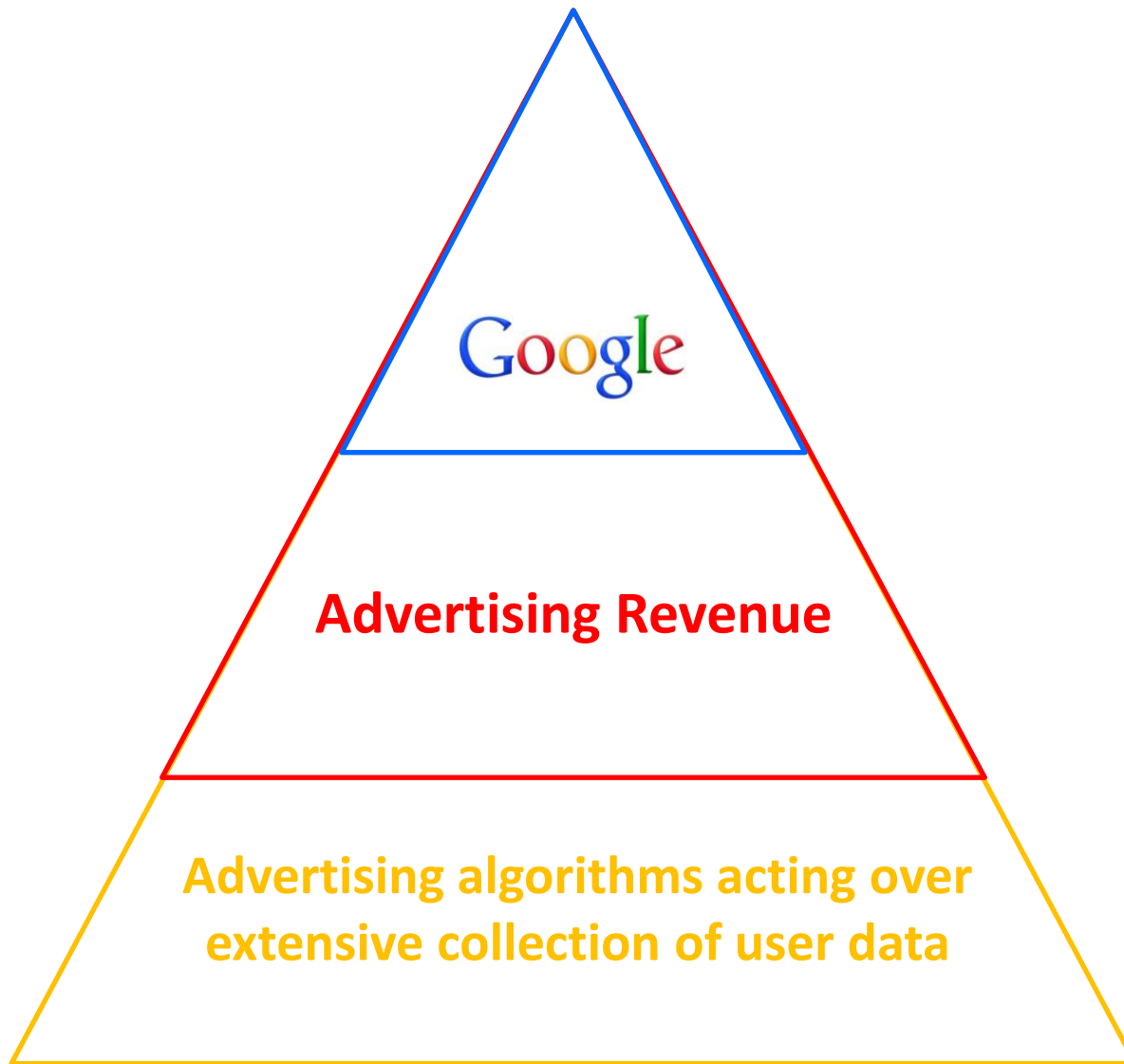
# A Contradiction?

*Slogan: Don't be evil*

but...

*“If you have something that you don't want anyone to know, maybe you shouldn't be doing it in the first place”*

*--Eric Schmidt*



Google

**Advertising Revenue**

**Advertising algorithms acting over  
extensive collection of user data**

**Privacy Responsibilities  
They are now shaping the web**

# Breaking Down the Privacy Policy

- Google *collects* everything it can
  - Information you volunteer, data from other services, UDIs, location data, etc.
- What Google *shares* depends on your preferences
  - Opt-in for sharing personal info with outside companies
  - Google shares non-identifying aggregate information
- Google *uses* info within itself and its own services
  - Different policies for different services

# Important Definitions

- From Google Privacy Policy:

## **Personal information**

This is information which you provide to us which personally identifies you, such as your name, email address or billing information, or other data which can be reasonably linked to such information by Google.

## **Sensitive personal information**

This is a particular category of personal information relating to confidential medical facts, racial or ethnic origins, political or religious beliefs or sexuality.

# Gmail Case Study

## Scope of the Issue

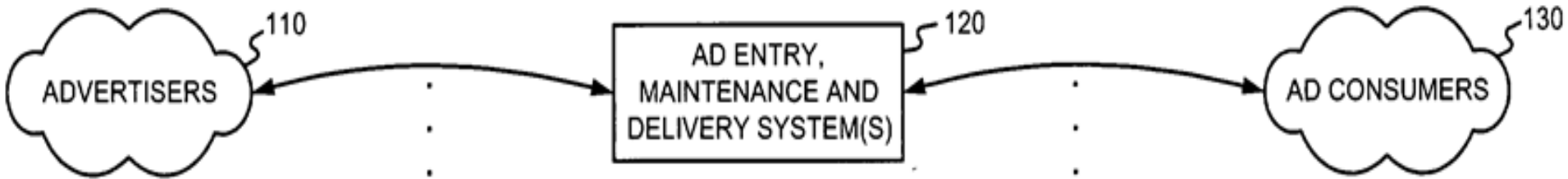
- 425 million users of Gmail
- 72 major universities use Google Apps for email, including 7 of the 8 Ivies
- Large number of K-12 schools, public, non-public, and charter also use Google Apps

# Gmail Privacy Policy Specifics

- Google uses additional information from other Google products to aid in ad targeting
  - You can *opt-out* of this practice

**You cannot opt-out of email-mining**

# Google Targeted Advertising



- Advertisers send ad info to system
  - Available ads
  - Serving constraints
- Advertisers receive info on ad use
- Consumer sends ad request and document details
- Ad system chooses ads to put in document
- Ad consumers send information back to the system



# Gmail Content Extraction

1. Internal (from the email data):  
Google mines all email data for key words
2. External (from inferences based on email data)



Google uses proprietary process on email data to make inferences that are also used in add selection

# Internal v. External Data

- Internal Data

- Email body
- Subject line
- Sender name & email
- Receiver name & email
- Email attachments

- External Data

- Sender location (from IP address or time-zone)
- Concepts derived from body text
- Concepts derived from attachments

# Gmail Case Study

## Patent Figures

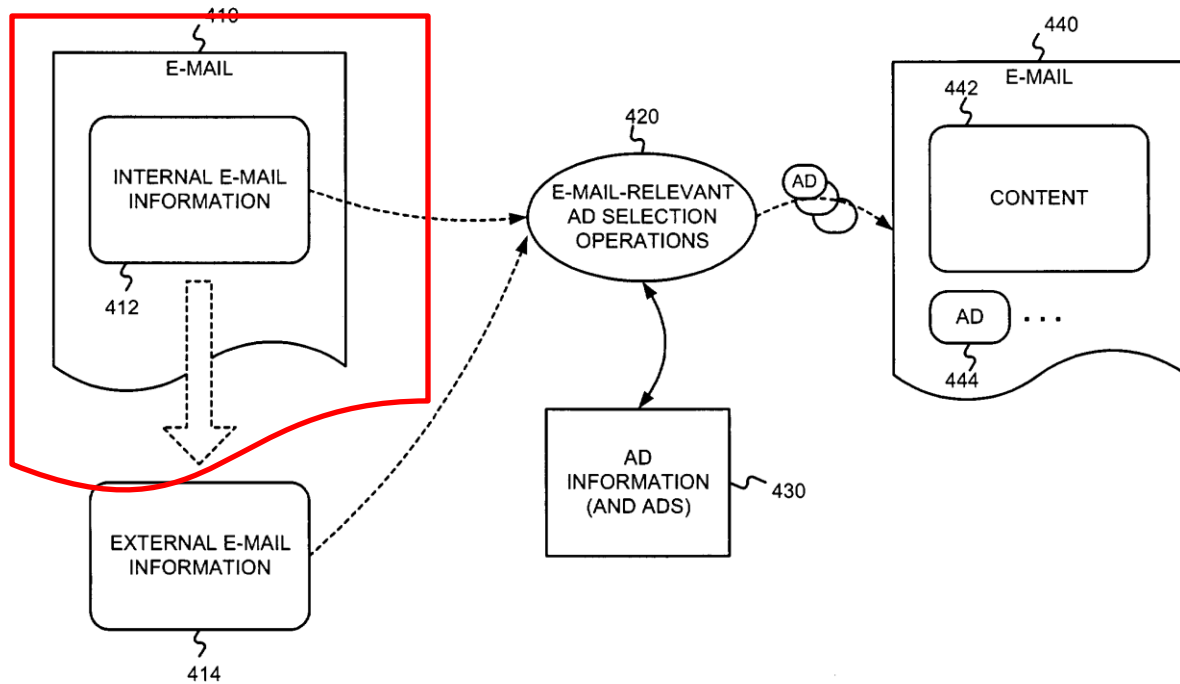


FIGURE 4

# Gmail Case Study

## Patent Figures

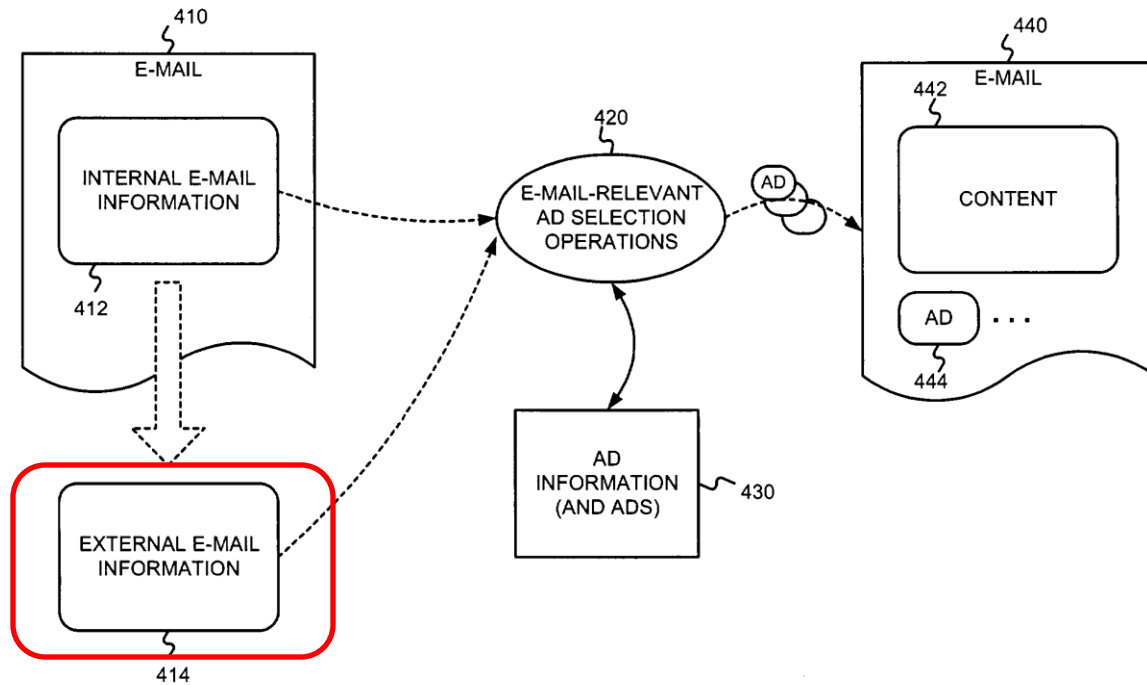


FIGURE 4

# Gmail Case Study

## Patent Figures

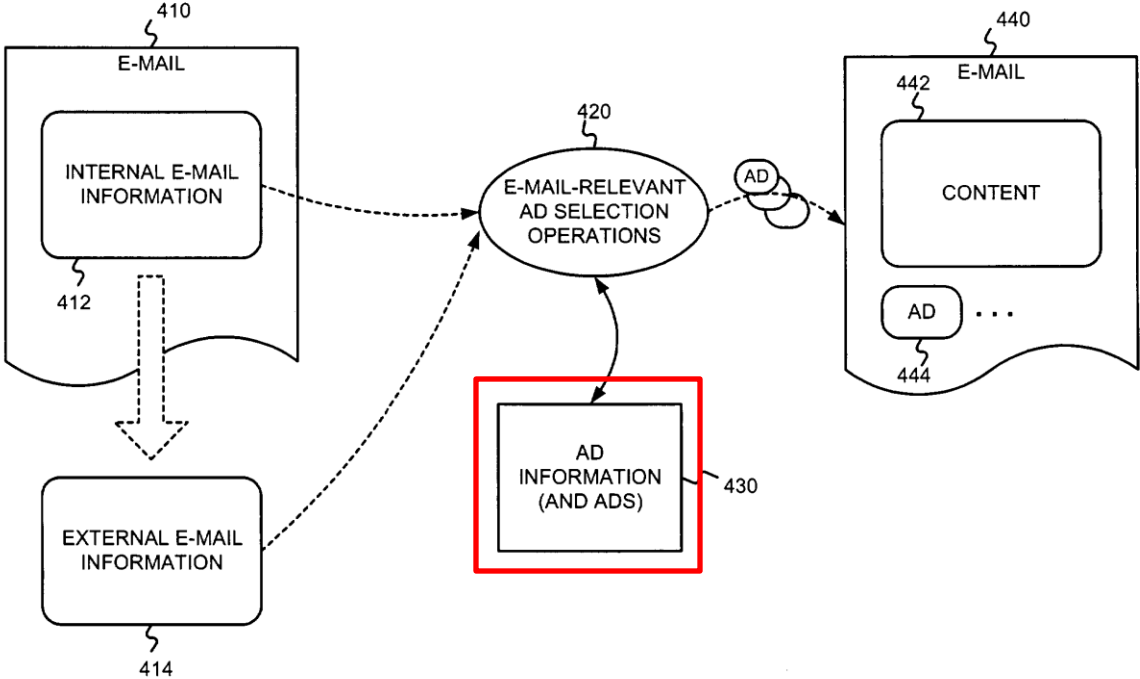


FIGURE 4

# Gmail Case Study

## Patent Figures

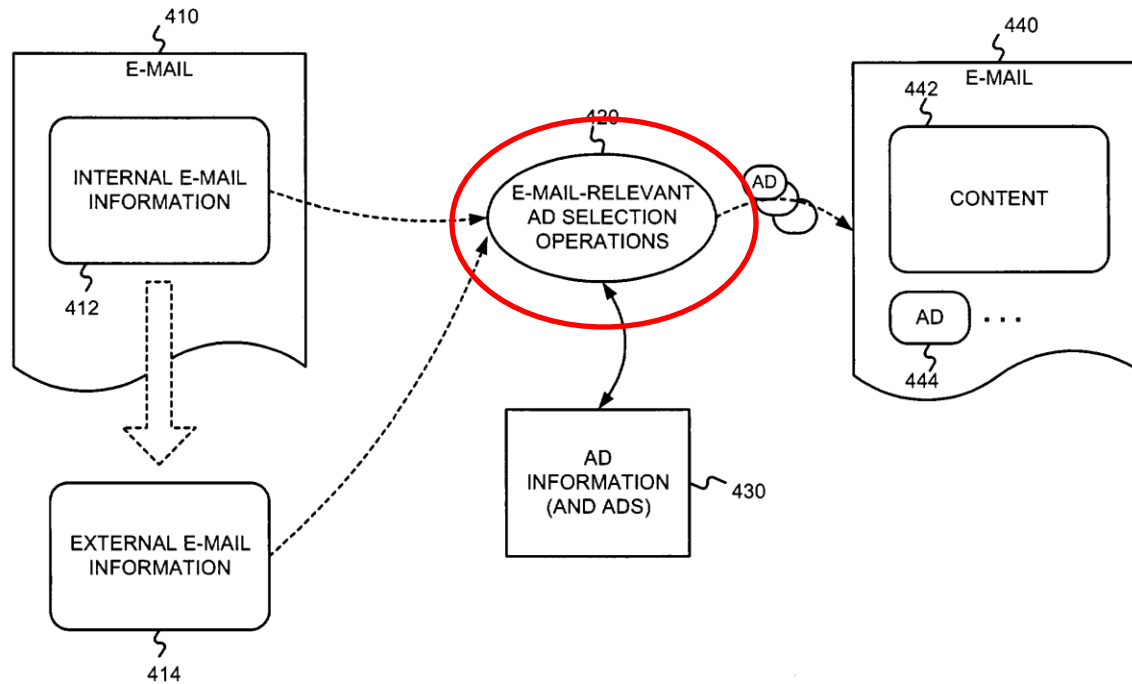


FIGURE 4

# Gmail Case Study

## Patent Figures

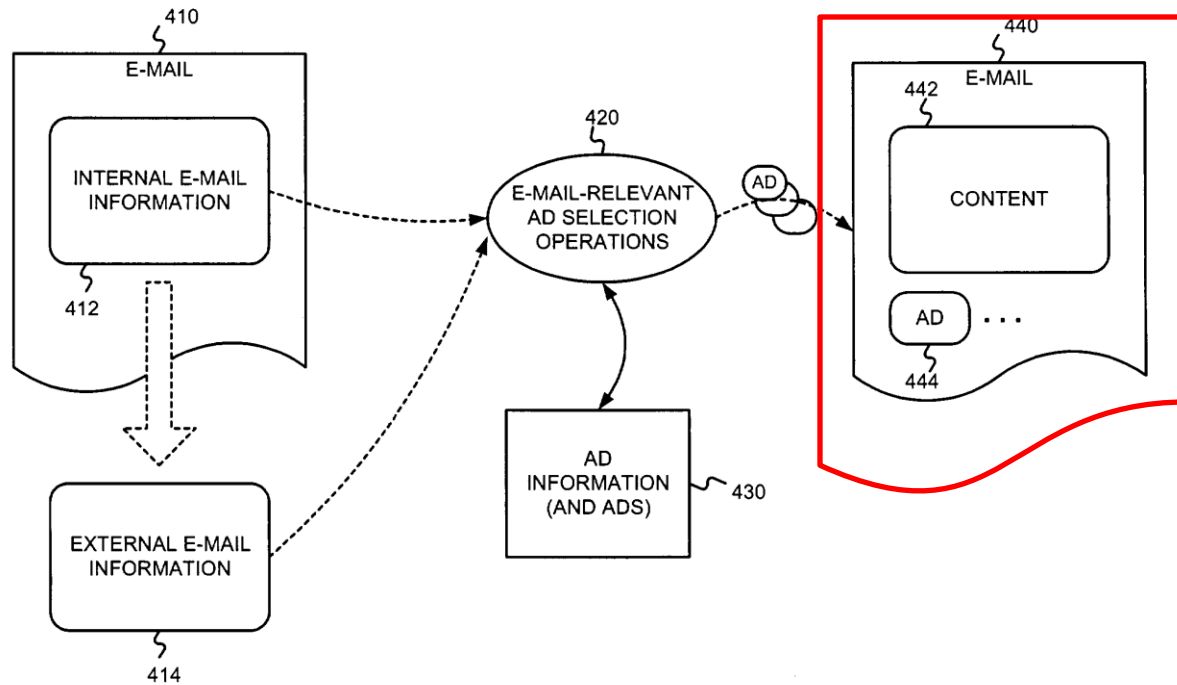
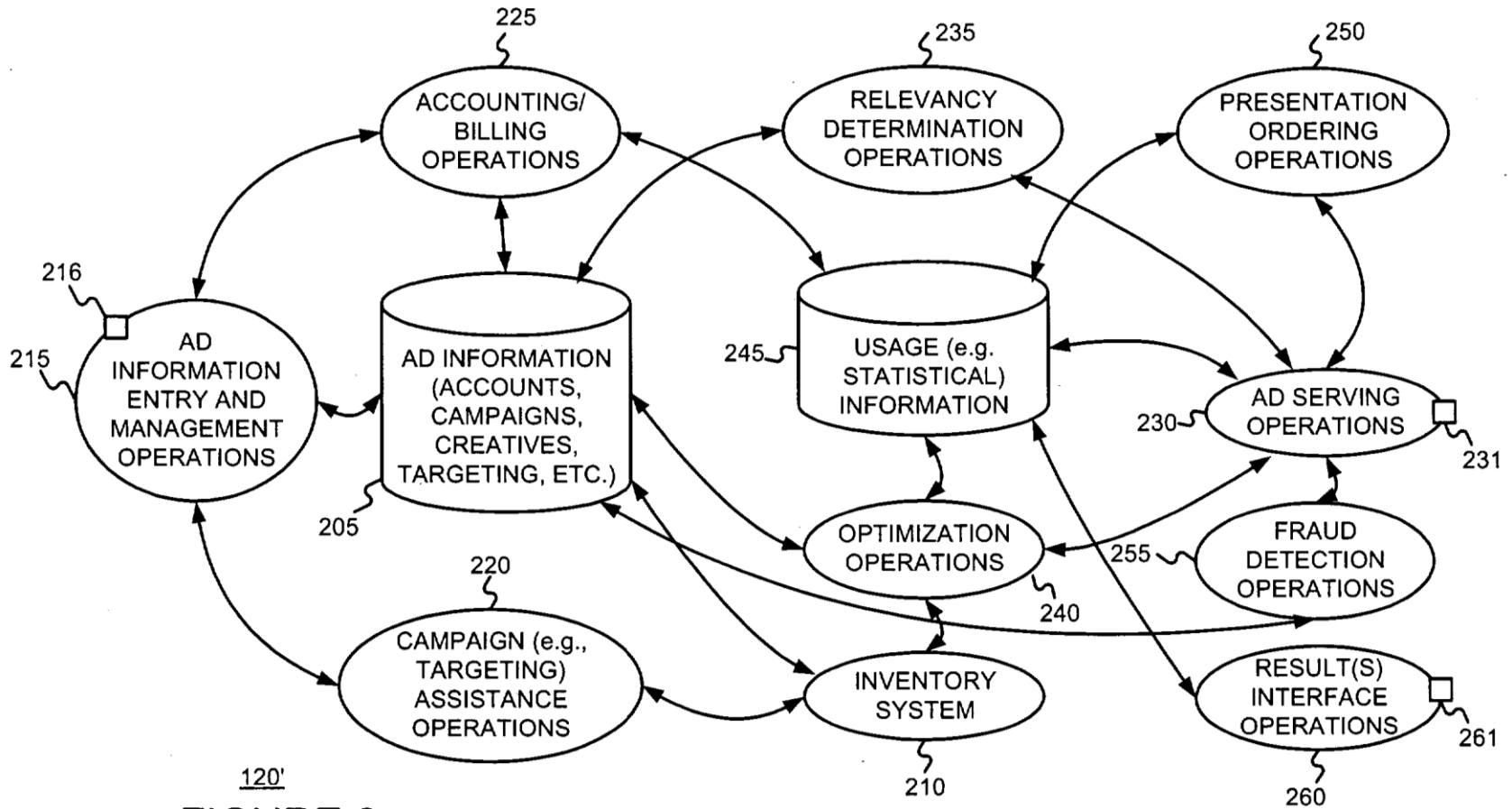


FIGURE 4

# “Invention” in the Network



120'  
FIGURE 2



# Gmail and Profiling

- Google has suggested that it does not create profiles on users
- The patent for data extraction and ad targeting suggests otherwise
  - “information about the sender (...derived from pervious interactions with the sender)”
  - “information about a recipient...based on a profile about the sender”
- Potentially a serious threat to privacy

# Gmail and the Surveillance Morass



Ties to CIA and NSA

Secondary use of gmail data and search data

# Gmail Case Study

- What can you do?

If you get an e-mail from a Gmail account and you wish not to reply consider explaining something like this:

*Dear Friend,*

*I have received your e-mail, but due to privacy concerns, I don't want to send my response to your Gmail account. Please give me another e-mail address where I can reach you. If you don't have another e-mail address, consider the following free e-mail accounts with generous storage which do not pose the same privacy risks:*

- *Rediffmail(1GB + no content extraction)*
- *Walla(1GB + no content extraction)*
- *Spymac(1GB + no content extraction)*
- *Aventure-mail(2GB + no content extraction)*

*For more information on the privacy risks posed by Gmail, see <http://www.epic.org/privacy/gmail/faq.html>.*

*Sincerely,*

*Concerned Citizen*

- Serious infringement upon your contacts – even if they're not gmail subscribers

# Integration of Services

## Implications of “Real Name” Policy

Gmail, Google+, YouTube,...

Real name policy now enforced

Discussion point: Why is this troubling?

# Street View Case Study

## OUTRAGE?!?

- Starting in 2007
  - Compromising images
  - Privacy invasion: recording of private spaces
  - Snooped on WiFi data
    - Packet payload
    - Location data
- Does the merit of this service outweigh privacy concerns?

# Street View Case Study - Wardriving

- Intercepted payload data over unencrypted wifi networks
- Google claim: Accidental inclusion of experimental code
- FCC report: software was *intentionally* engineered to do so
- Google Engineer “Doe”: Marius Milner

# Wardriving – Their Design Document

- Mapping location of wireless access points

*“A typical concern might be that we are logging user traffic along with sufficient data to precisely triangulate their position at a given time, along with information about what they were doing.”*

- Wi-Fi data Google gathered would “be analyzed offline for use in other initiatives”
- **Not so “honest” a mistake...**

# Street View Case Study – Legal

- Legal issues surrounding the images
  - German examples
  - \$7 million settlement
- Petitioning to get images removed
  - Easy for the government
    - Pentagon Example
  - Hard for the individual
    - Boring v. Google



# Boring v. Google

## 2008 – 2010

- Street View images their home, on *private* road
- “Borings were unable to show that Google’s conduct was highly offensive to a person of normal sensibilities”
- No proof of “damage as a result of the trespass”
- Ultimate settlement: \$1 (Ouch)

# \$7 Million Multistate Settlement

## A glimmer of hope?

- CT led 2 year investigation
- “consumers have a reasonable expectation of privacy”
  - What does this really mean?
- Requires:
  - Google employee training policy
    - User privacy and confidentiality
  - National privacy awareness campaign (where is it?)
  - Secure and delete data from 2008 – 2010

# Our Mandate to Google

*“As a dominant force **shaping and changing how consumers use the internet**, Google must also **show leadership in minimizing security and privacy risks to consumers** who take advantage of the internet.*

*In complying with this settlement, Google has the opportunity to **set the bar for the industry** in better educating the public about avoiding and reducing cyber-risks.”*

# Alternate Viewpoints

## Points for Discussion

- International Perspective:
  - Germany's Telemediengesetz Federal Law
  - Anonymous access to online services is a legal requirement
- Is privacy really the be-all-end-all?
- Why aren't others as worried?
- Charlie Nesson (Founder of the Berkman Center) :
  - Focus on openness, not privacy and security
  - People are collectively learning how to use the internet

## Sources:

Computer eye image: <http://www.activistpost.com/2013/02/congress-will-battle-over-internet.html>

Boring v. Google case: <http://www2.ca3.uscourts.gov/opinarch/092350np.pdf>

Street View Wifi passwords scandal:

<http://www.wired.com/threatlevel/2012/05/google-wifi-fcc-investigation/>

Background Criticism on Google: [http://en.wikipedia.org/wiki/Criticism\\_of\\_Google](http://en.wikipedia.org/wiki/Criticism_of_Google)

Fed-Google image: <http://www.theipinionsjournal.com/2006/01/google-refuses-to-join-the-governments-thinkpol-spying-force/>

NSA/CIA ties: <http://www.wired.com/threatlevel/2010/02/google-seeks-nsa-help/>  
<http://www.wired.com/dangerroom/2010/02/from-dont-be-evil-to-spy-on-everyone/>

[http://www.theregister.co.uk/2004/10/28/google\\_buys\\_keyhole/](http://www.theregister.co.uk/2004/10/28/google_buys_keyhole/)

[http://www.slate.com/blogs/moneybox/2013/06/07/us\\_tech\\_giants\\_have\\_many\\_foreign\\_customers.html](http://www.slate.com/blogs/moneybox/2013/06/07/us_tech_giants_have_many_foreign_customers.html)

<http://www.theverge.com/2012/6/28/3123643/gmail-425-million-total-users>

<http://www.edudemic.com/report-vast-majority-of-top-u-s-schools-use-google-apps/>

<http://www.google.com/enterprise/apps/education/customers.html>

The EPIC FAQ: <http://epic.org/privacy/gmail/faq.html>

Gmail Ad Policy: <https://support.google.com/mail/answer/6603?hl=en>

Google Privacy Policy: <http://www.google.com/policies/privacy/>

Google Privacy Policy Key Terms: <http://www.google.com/policies/privacy/>

Google Data extraction and ad choosing patent (in full with diagrams):

Patent: <http://www.google.com/patents/US20040059712>

Pentagon pressures: <http://www.informationweek.com/internet/google/google-caves-to-pentagon-wishes/229212721>

- Multistate Settlement: <http://www.ct.gov/ag/cwp/view.asp?Q=520518&A=2341>