Search Among Secrets

Douglas W. Oard
iSchool and UMIACS
University of Maryland, College Park, USA
Open Government

“Information maintained by the Federal Government is a national asset. My Administration will take appropriate action, consistent with law and policy, to disclose information rapidly in forms that the public can readily find and use.”

Presidential Memorandum, January 21, 2009
- 62,320 messages on a private server
- Term-list hits yielded 30,490 as “work-related”
  - Provided to the State Department, initially on paper
  - Added emails from State Department archives
    - Of ~1 billion emails per year, ~50,000 are archived (0.005%)
- Complex multi-agency manual review process
  - Target: 1,000 documents per week
1. Review a batch (1,000 messages) per week
2. Apply proposed redactions to each batch
3. Send batch to subject matter experts (SMEs) for consultation
4. SMEs return batch to FOIA office
5. FOIA office incorporates review recommendations from SMEs
6. FOIA office refers to other agencies implicated in emails
7. Other agencies review
8. Other agencies send feedback to FOIA office
9. FOIA office incorporates appropriate recommendations
10. FOIA office sends batch to Office of Legal Counsel (OLC) for review
11. OLC reviews
12. OLC sends feedback to FOIA office
13. FOIA office incorporates OLC recommendations
14. Divergent recommendations among reviewing entities reviewed
15. Final review
Exclusive: U.S. to shift 50 staff to boost office handling Clinton emails
WASHINGTON | BY ARSHAD MOHAMMED

The U.S. State Department plans to move about 50 workers into temporary jobs to bolster the office sifting through Hillary Clinton's emails and grappling with a vast backlog of other requests for information to be declassified, officials said on Tuesday.

The move illustrates the huge administrative burden caused by Clinton's decision to use a private email address for official communications as secretary of state and a judge's ruling in a Freedom of Information Act (FOIA) lawsuit that they be released.
Jeb Bush email dump includes Social Security numbers

A trove of 250,000 emails released by prospective 2016 presidential candidate Jeb Bush includes the sensitive personal information of several Florida residents, leaving them vulnerable to identity thieves.

A scan of the email dump by technology blogs The Verge and Gizmodo revealed names, emails and in some cases, Social Security numbers of Bush’s correspondents. Many appear to be normal Florida residents unaware their messages to the then-governor would eventually become public.
A Story in Three Parts

➢ The need for search among secrets

• Protecting things

• Protecting parts of things
What’s Different Now?

Formal

Informal
Hourglass Model of Mission Control

3

1

1

16

19

Spacecraft
“Capsule” Communicator
Flight Director
Flight Controllers
Staff Support Rooms (SSR)
ISSUE NO. 2: Prioritizing the Declassification Review of Historically Significant Information.

There is no satisfactory means at present of identifying historically significant information within the vast body of information that is being reviewed and declassified. Accordingly, no priority is given to the declassification and release to the public of such information.
“Value-Sensitive” Prioritization

Such a system might operate as follows: A board consisting of prominent historians, academicians, and former Government officials would be appointed by the Archivist to determine which events or activities of the U.S. Government should be considered historically significant from a national security and foreign policy standpoint for a particular year.

PIDB Improving Declassification report, December, 2007
U.S. Freedom of Information Act

“identify the subject(s) or record(s) as clearly and specifically as possible -- for example, all previously released National Intelligence Estimates (NIEs) on the former Soviet Union's space program.”

E-Discovery

Clinton White House

32 million emails

National Archives

search request

Tobacco Policy

80,000

hired 25 persons

for 6 months …

200,000

32 million emails
An E-Discovery Process Model

- Formulation
  - Production request
  - Acquisition
    - Collection
      - Review for Relevance
        - Responsive ESI
          - Production
            - Sense-making
              - Insight
A Story in Three Parts

- The need for search among secrets

  ➢ Protecting things

- Protecting parts of things
Automated Privilege Classification

○ CM-AS-A  □ CM-NAS
Indicative Terms for Privilege

agreement

credit
ENA
payment
form
seller
mark
litigation
trading
event
delivery
amounts
swap
obligations
contract
ISDA
support
party

legal
court
without
provided

transaction
language
respect
agreements
transactions
claim
claims
damages
non

provisions
counsel
section
Estimating Privilege Propensity
Machine-Assisted Privilege Review

Date: Wed, 25 Apr 2001 03:43:00
From: J.Bennett <J.Bennett@GM.jpg>
To: Harry.Kingerski@enron.com, Jeff.Dasovich@enron.com, Leslie.Lawner@enron.com, t.johnso8@enron.com, James.D.Steffes <James_D.Steffes@enron.com>, Scott.Stoness (E-mail) <sstoness@enron.com>, Sue.Mara (E-mail) <smara@enron.com>
Cc: MDay <MDay@GMSSR.com>
Subject: Advice Letter Protest on SCE PX Credit

Attached is the revised version of the protest of SCE's advice letter on the PX credit. Please provide all comments ASAP. If you cannot get them to me by 12:30 PDT, then send to Mike Day for incorporation.

* The above graph plots Privilege propensity
Annotations Increase Recall

S₁ Judgments as Ground Truth

RECALL ↑

PRECISION ↓
Balancing Relevance and Sensitivity
Evaluation

• Our goal is to show some relevant documents
  – To do that, we must risk showing some sensitive docs

• Precision-oriented search could limit the risk

• Loss function from operational risk management
  – Severity, mitigation, continuous improvement
## Discounted Cumulative Gain

### Table: Discounted Cumulative Gain

<table>
<thead>
<tr>
<th></th>
<th>Highly Relevant (h)</th>
<th>Moderately Relevant (m)</th>
<th>Not Relevant</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RETRIEVED</strong></td>
<td>+G(_h)</td>
<td>+G(_m)</td>
<td>0</td>
</tr>
<tr>
<td><strong>NOT RETRIEVED</strong></td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

\[ DCG_k = \sum_{i=1}^{k} \frac{g_i}{d_i} \]
## Cost-Sensitive DCG

### Not Retrieved

<table>
<thead>
<tr>
<th></th>
<th>Highly Relevant (h)</th>
<th>Moderately Relevant (m)</th>
<th>Not Relevant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not Retrieved</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

### Retrieved

\[
DCG_k = \sum_{i=1}^{k} \frac{g_i}{d_i}
\]

### Fine to Show

<table>
<thead>
<tr>
<th></th>
<th>Highly Relevant (h)</th>
<th>Moderately Relevant (m)</th>
<th>Not Relevant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fine to Show</td>
<td>(+G_h)</td>
<td>(+G_m)</td>
<td>0</td>
</tr>
</tbody>
</table>

### Somewhat Sensitive (s)

<table>
<thead>
<tr>
<th></th>
<th>Highly Relevant (h)</th>
<th>Moderately Relevant (m)</th>
<th>Not Relevant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Somewhat Sensitive (s)</td>
<td>(-C_s)</td>
<td>(-C_s)</td>
<td>(-C_s)</td>
</tr>
</tbody>
</table>

### Very Sensitive (v)

<table>
<thead>
<tr>
<th></th>
<th>Highly Relevant (h)</th>
<th>Moderately Relevant (m)</th>
<th>Not Relevant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very Sensitive (v)</td>
<td>(-C_v)</td>
<td>(-C_v)</td>
<td>(-C_v)</td>
</tr>
</tbody>
</table>

### CS - DCG

\[
CS - DCG_k = \sum_{i=1}^{k} \left( \frac{g_i}{d_i} + c_i \right)
\]

### Not Retrieved

<table>
<thead>
<tr>
<th></th>
<th>Highly Relevant (h)</th>
<th>Moderately Relevant (m)</th>
<th>Not Relevant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fine to Show</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

### Somewhat Sensitive (s)

<table>
<thead>
<tr>
<th></th>
<th>Highly Relevant (h)</th>
<th>Moderately Relevant (m)</th>
<th>Not Relevant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Somewhat Sensitive (s)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

### Very Sensitive (v)

<table>
<thead>
<tr>
<th></th>
<th>Highly Relevant (h)</th>
<th>Moderately Relevant (m)</th>
<th>Not Relevant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very Sensitive (v)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
A Story in Three Parts

• The need for search among secrets

• Protecting things

➢ Protecting parts of things
Sudheendra Hangal, Providing Access to Historical Email Archives for Historical Research,

*Personal Digital Archiving, 2013*
Date: March 2, 2001 11:24am
From: Katbarc@
To: creeley@
Subject:

..... Robert,
..... 
..... Central Michigan (...'....... ..., ...
..... ....) ....... ....... ....... ...
..... ............... ....... ....
..... ............... ....... ....

..... Portland

.... $,... ...'....... ....... ....... ...
.... $,...

....

Kathryn

Kathryn Barcos
The Steven Barclay Agency

.....

Petaluma, ..

.....

Jun 03, 1961; Kennedy Administration

Redaction 1/30

danger. The USSR does not foment revolution but the United States always looks for outside forces whenever certain upheavals occur. One example of USSR's determination not to interfere in internal affairs of other countries is Iran, an ally of the United States. The Soviet Union dose not want a revolution there and dose not do anything in that country to promote such a development. However, the people of that country are so poor that the country has become a volcano and changes are bound to occur sooner or later. The Shah will certainly be overthrown. By supporting the Shah, the United States generates adverse feelings toward the United States among the people of Iran and, conversely, favorable feelings toward the USSR. This, of course, is to the US's own disadvantage. The Soviet Union does not sympathize with dictators or tyranny. This is the crux of the matter. No agreement seems to be possible on this
Evaluating Redaction

• Re-identification attacks
• K-anonymity
  – Goal: Set a lower bound on uncertainty
  – Method: Suppress identifying attributes
  – Challenges: weak ties, future data
• Differential privacy
  – Goal: public data for preliminary analysis
  – Method: preserve approximate statistics (e.g., IDF)
  – Challenges: low counts, high dimensionality
Consequences for IR Evaluation

• 20\textsuperscript{th} Century IR Evaluation
  – Perfect retrieval is our ideal
    • All and only the relevant documents
  – We measure how close to the ideal we come
    • Precision, Recall, MAP, NDCG, ...

• 21\textsuperscript{st} Century IR Evaluation
  – Maximizing access to relevant content is our goal
  – Protecting sensitive content expands search space
Conclusion

• IR systems are too good
  – They could find what they shouldn’t
  – To prevent this, we do not index some things
  – What we don’t index, we can’t find

• Two lines of research are needed
  – Techniques for search among secrets
  – Techniques for evaluation search among secrets
For More Information

• FnTIR survey on E-Discovery
  – http://ediscovery.umiacs.umd.edu

• Stanford ePADD (Email Redaction)
  – Demo: http://epadd.stanford.edu/
  – Code: https://library.stanford.edu/projects/epadd

• The Redaction Engine
  – http://www.history-lab.org/