GovRisk

The International Governance & Risk Institute

Innovative Use of Data Mining to Detect Collusion & Corruption

Colin Ehren

Managing Partner, C&SE

Tackling Corruption and Collusion in Public Procurement: Latin America and the Caribbean

September 12, 2013

Panama City, 2013

GovRisk
The International Governance & Risk Institute

Overview

What is Data Mining?

- Generic Term - A process used by Governments, Organisations and Companies to turn raw data into useful information.
- Data Mining relies on sources of information.
  - What does your organisation know?
  - What does your organisation NOT know?

Colin Ehren
candse@gmail.com
Copyright C&SE 2013

GovRisk
The International Governance & Risk Institute

Suspicious Bidding Patterns

Identification Issues / Opportunities

- What is 'Suspicious'?
  - May not be apparent in current bidding process.
- Internal records re bidding companies.
  - Intelligence / Analysis Resource.
  - Investigation Resource.
- May not be apparent in Single Organisation data.
- Government Department records.
- Foreign Government records.
  - Regional Information Sharing initiatives.

Colin Ehren
candse@gmail.com
Copyright C&SE 2013

GovRisk
The International Governance & Risk Institute

Suspicious Bidding Patterns

Know Your Staff

- What does your organisation know about your staff?
- What does it NOT know?
- Statutory Asset Declarations?
- Vetting Procedures?
  - Family relationships: Parents, Children, Siblings, etc.
- Professional Standards Intelligence Units?
  - Co-Worker relationships: Boss, Supervisor, Staff.

Colin Ehren
candse@gmail.com
Copyright C&SE 2013

GovRisk
The International Governance & Risk Institute

Suspicious Bidding Patterns

Potential Corruption Indicators

- Expenditure Profile > Income Profile.
- Assets not compatible with known income.
  - Property Ownership
  - Motor Vehicles
  - Jewellery
  - Clothing
  - Collectibles
  - Boats
  - Holidays
  - Schooling

Colin Ehren
candse@gmail.com
Copyright C&SE 2013

GovRisk
The International Governance & Risk Institute

Identifying Relationships

- The existence of a relationship is not necessarily suspicious.
  - Companies often have working relationships with others.
  - Staff often have relatives working in Industry.
- Have the companies declared the relationship when bidding for a contract?
- Has the member of Staff declared the relationship if over-seeing the procurement process?
- The identification of a relationship can be the first indicator of Corruption or Collusion.

Colin Ehren
candse@gmail.com
Copyright C&SE 2013

GovRisk
The International Governance & Risk Institute

Data Mining Sources

- Organisation/Government internal records.
- Partner Organisation/Government records.
- Official Company registers / records.
- Official Financial or Banking records.
  - Companies and their Directors/Staff.
  - Your Staff.

Colin Ehren
candse@gmail.com
Copyright C&SE 2013

GovRisk
The International Governance & Risk Institute

Data Mining Sources

- Online / Internet
  - Databases for Background & Asset checks.
  - Subscription services for Public & Court records.
  - Social / Business Networks
  - Investigative Bloggers.
  - Anti-Corruption NGO reporting.
  - Transparency International
  - UNODC, Interpol, StAR

Colin Ehren
candse@gmail.com
Copyright C&SE 2013

GovRisk
The International Governance & Risk Institute

Data Mining Sources

Useful information

- iBrief
- tBrief
- Whitepapers

Useful information

- Adds 10-15,000 articles per day.
- Offshore Intelligence datasets.
- Suspect Address data

Colin Ehren
candse@gmail.com
Copyright C&SE 2013

GovRisk
The International Governance & Risk Institute

Data Mining Sources

GovRisk
The International Governance & Risk Institute

Social Networks

- Wikipedia.com - 196 active Social Networking sites.
- E-crm.co.uk – 700+ sites listed.
- List ‘modern’ social networks.
- They take no account of traditional social networking on the Internet.

Colin Ehren
candse@gmail.com
Copyright C&SE 2013

GovRisk
The International Governance & Risk Institute

Social Networks

- Internet Relay Chat
- Usenet Talk Groups
- Google Groups / Yahoo Groups
- E-mail
- Dedicated Discussion Forums
- Dating & Reunion Sites

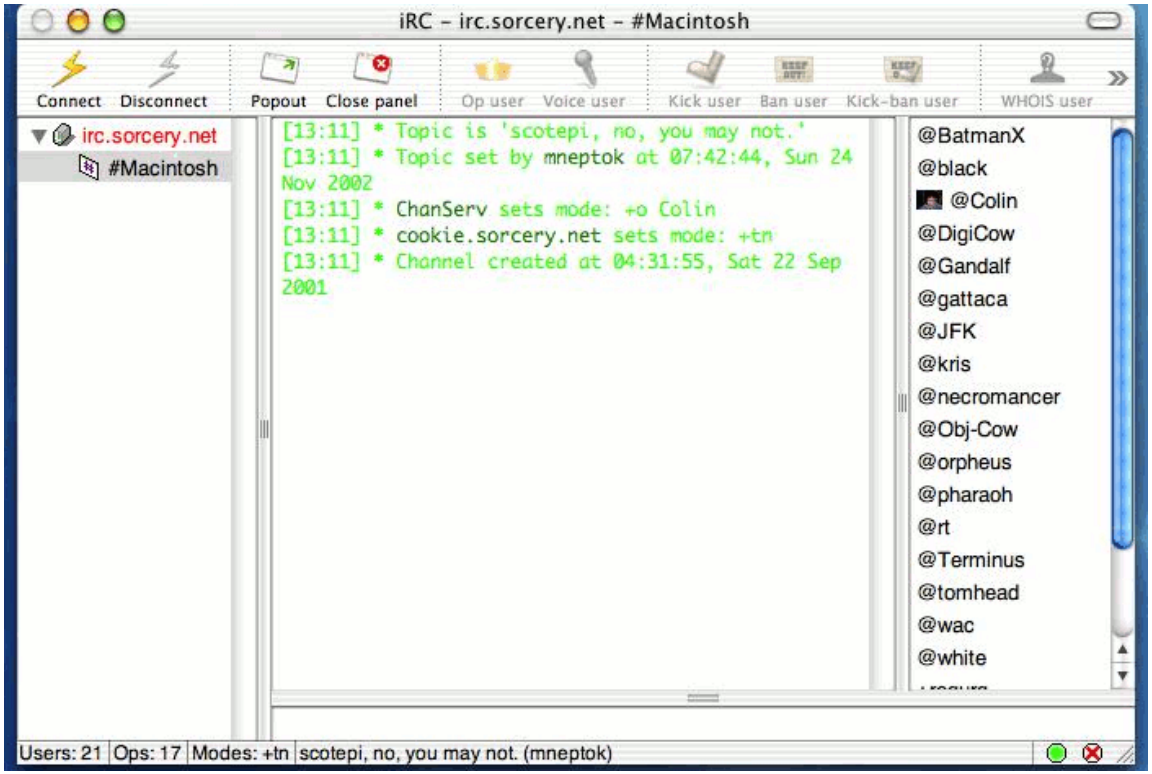
Colin Ehren
candse@gmail.com
Copyright C&SE 2013

GovRisk
The International Governance & Risk Institute

IRC (Chat Rooms)

Mibbit

MICE iRC



- Who is online

Colin Ehren
candse@gmail.com
Copyright C&SE 2013

GovRisk
The International Governance & Risk Institute

Groups / Discussions

- Public Posting / Forum Service
- Multiple access points both free and account based.
- Very wide range of subjects.
- Multiple Adult themed providers.
- General content providers
  - Google Groups
    - Conversations – Manchester Utd
    - Jellore – User Profile
  - Yahoo Groups
    - Search – Manchester Utd
    - Groups – PoshnBecksPictures

Colin Ehren
candse@gmail.com
Copyright C&SE 2013

GovRisk
The International Governance & Risk Institute

E-Mail

- Headers

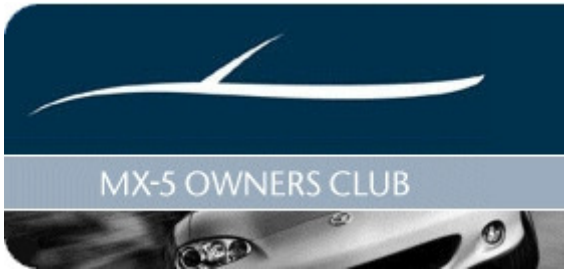
```
Delivered-To: aseem1234+test@gmail.com
Received: by 10.115.22.20 with SMTP id z20cs54925wai;
        Tue, 2 Oct 2007 20:38:18 -0700 (PDT)
Received: by 10.70.112.4 with SMTP id k4mr13653349wxc.1191382696912;
        Tue, 02 Oct 2007 20:38:16 -0700 (PDT)
Return-Path: <aseem1234@cox.net>
Received: from eastrmmtao104.cox.net (eastrmmtao104.cox.net [68.230.2
        by mx.google.com with ESMTP id h13si74220wxd.2007.10.02.20.38
        Tue, 02 Oct 2007 20:38:16 -0700 (PDT)
Received-SPF: pass (google.com: domain of aseem1234@cox.net designate
Authentication-Results: mx.google.com; spf=pass (google.com: domain
Received: from eastrmimpo01.cox.net ([68.1.16.119])
          by eastrmmtao104.cox.net
          (InterMail vM.7.08.02.01 201-2186-121-102-20070209) with ES
          id <20071003033817.EGTR14054.eastrmmtao104.cox.net@eastrmir
          for <aseem1234+test@gmail.com>; Tue, 2 Oct 2007 23:38:17 -0
Received: from aseem ([72.204.154.191])
        by eastrmimpo01.cox.net with bizsmtp
        id vfeD1X00Q482NXc0000000; Tue, 02 Oct 2007 23:38:15 -0400
From: "Aseem Kishore" <Aseem1234@cox.net>
To: <aseem1234+test@gmail.com>
Subject: Test
Date: Tue, 2 Oct 2007 22:38:51 -0500
```

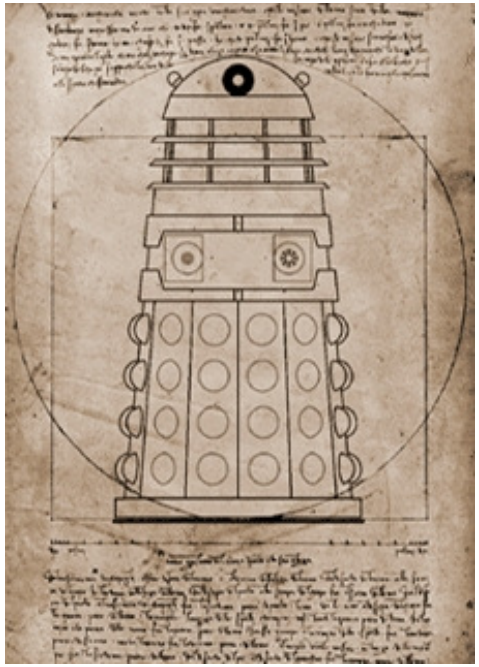
- Tracking – aseem kishore, aseem1234, 72.204.154.191
- Personal Spammers – Email Hunters

Colin Ehren
candse@gmail.com
Copyright C&SE 2013

GovRisk
The International Governance & Risk Institute

Discussion Forums



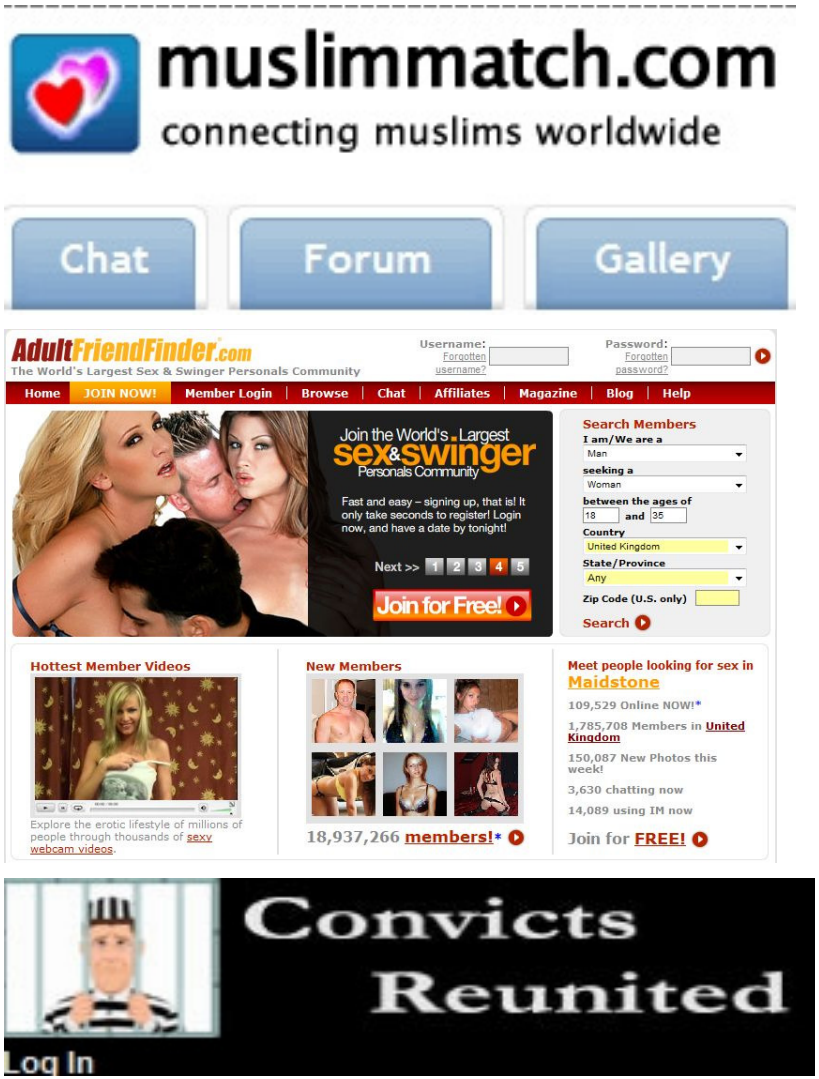


September 12, 2013

Colin Ehren
candse@gmail.com
Copyright C&SE 2013

GovRisk
The International Governance & Risk Institute

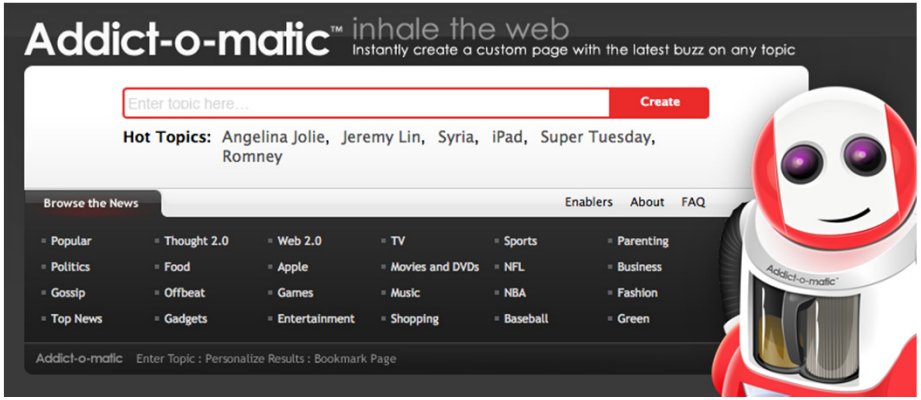
Dating & Reunion



Colin Ehren
candse@gmail.com
Copyright C&SE 2013

GovRisk
The International Governance & Risk Institute

Social Media Monitoring



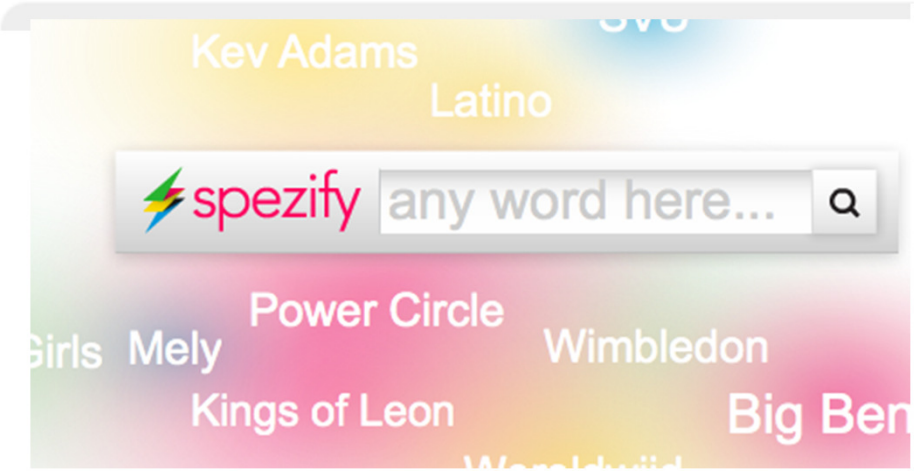
TOPSY

Real-time search for the social web

socialmention*

Real-time social media search and analysis:

KURRENTLY.com



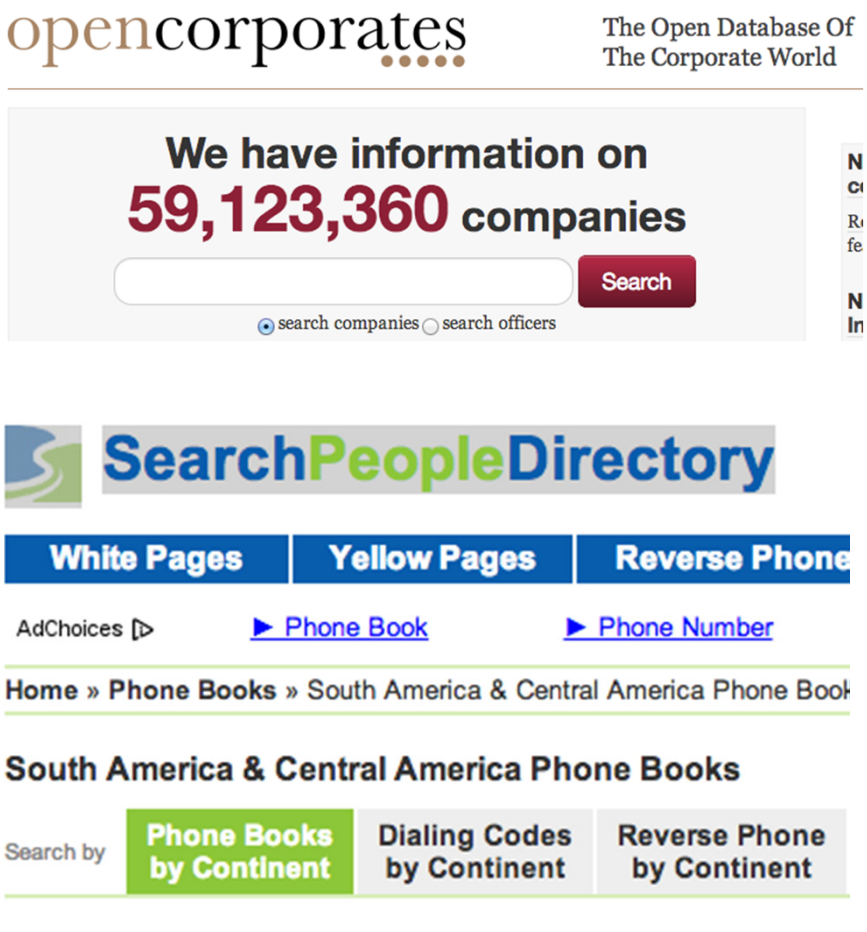
WHOS TALKIN

September 12, 2013

Colin Ehren
candse@gmail.com
Copyright C&SE 2013

GovRisk
The International Governance & Risk Institute

Data Mining Sources



September 12, 2013

Colin Ehren
candse@gmail.com
Copyright C&SE 2013

GovRisk
The International Governance & Risk Institute

Data Mining Sources

September 12, 2013

Colin Ehren
candse@gmail.com
Copyright C&SE 2013

GovRisk
The International Governance & Risk Institute

Visualisation

TouchGraph NAVIGATOR

LinkedIn Maps

Colin Ehren
candse@gmail.com
Copyright C&SE 2013

GovRisk
The International Governance & Risk Institute

Any Questions?

September 12, 2013

Colin Ehren
candse@gmail.com
Copyright C&SE 2013

GovRisk
The International Governance & Risk Institute

Innovative Use of Data Mining to Detect Collusion & Corruption

Colin Ehren
Managing Partner, C&SE

candse@gmail.com
+44 (0)7941 338 449

Tackling Corruption and Collusion in Public Procurement: Latin America and the Caribbean

September 12, 2013

Panama City, 2013

Colin Ehren
candse@gmail.com
Copyright C&SE 2013