



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

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?

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.

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.

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

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.**

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.

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

Data Mining Sources

Home About Products Services Markets Contact

10 December 2010, Time:06:16 PM

Homepage What is C6?

C6 is the world leading provider of intelligence information providing both positive and negative information enabling you to take a true risk based approach to managing fraud and money laundering compliance risks.

Since our creation C6 has become a market leader in KYC, KYS and EDD. We have a firm commitment to our customers to ensure that our cutting edge data sets are continually updated, we currently add between 10-15,000 new articles a day across the globe.

Our datasets represent the best of breed sources enabling our clients to make informed corroborated decisions utilising a single interface.

Find out [more information about C6](#).

Latest News

Know Your Customer



C6 was designed originally to provide the ultimate tool to Know Your Customer. This is C6's main use across the world and has an enviable reputation for efficiency and factual documentary support enabling swift confirmation of identity and intent.

Anti Money Laundering



C6 monitors global Bank, Police, Regulatory and Government sites to ensure up to the minute information. The system is updated minute by

Useful information

- iBrief
- tBrief
- Whitepapers

Useful information

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

September 12, 2013

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

Data Mining Sources



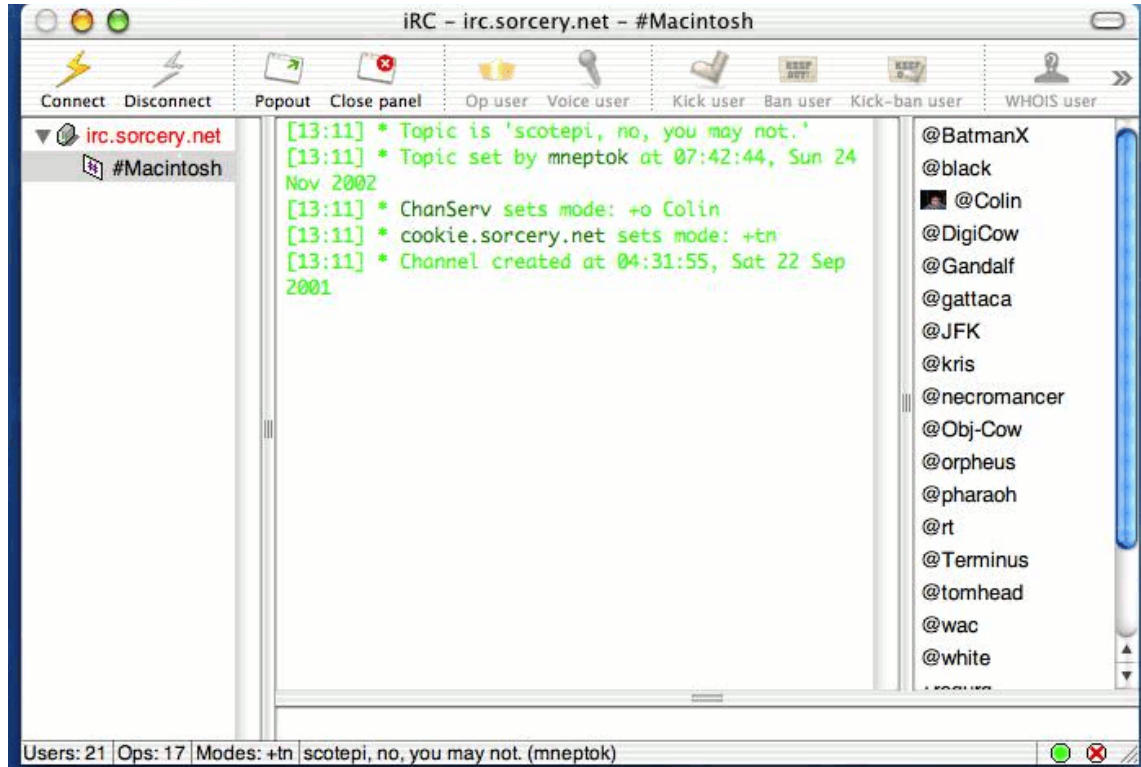
Social Networks

- 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.

Social Networks

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

IRC (Chat Rooms)



- Who is online

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

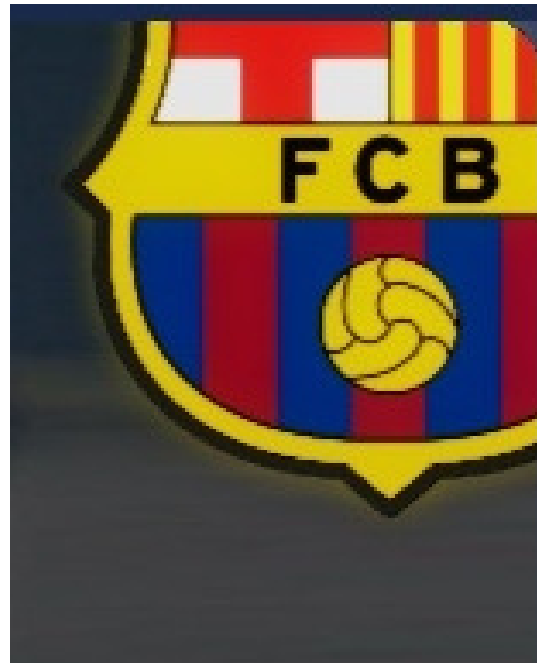
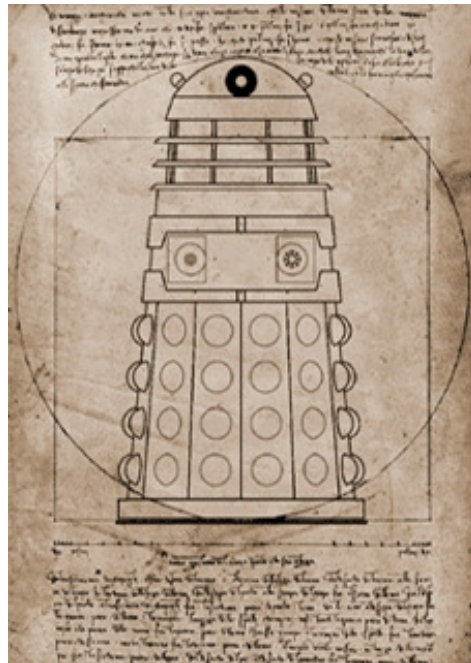
E-Mail

- 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 eastrrmmtao104.cox.net (eastrrmmtao104.cox.net [68.230.1
      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 c
Received: from eastrrmimpo01.cox.net ([68.1.16.119])
      by eastrrmmtao104.cox.net
      (InterMail vM.7.08.02.01 201-2186-121-102-20070209) with ESMTP
      id <20071003033817.EGTR14054.eastrrrmmtao104.cox.net@eastrrmir
      for <aseem1234+test@gmail.com>; Tue, 2 Oct 2007 23:38:17 -0400
Received: from aseem ([72.204.154.191])
      by eastrrmimpo01.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

Discussion Forums



September 12, 2013


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

Dating & Reunion



muslimmatch.com
connecting muslims worldwide

Chat Forum Gallery



AdultFriendFinder.com
The World's Largest Sex & Swinger Personals Community

Username: Password:

Home **JOIN NOW!** Member Login Browse Chat Affiliates Magazine Blog Help

Join the World's Largest **sex & swinger** Personals Community

Fast and easy – signing up, that is! It only takes seconds to register! Login now, and have a date by tonight!

Next >> 1 2 3 4 5

Join for Free!

Search Members

I am/We are a

seeking a

between the ages of and


Country

State/Province

Zip Code (U.S. only)

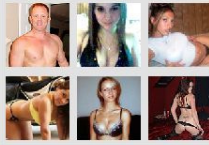
Search

Hottest Member Videos



Explore the erotic lifestyle of millions of people through thousands of **sexy webcam videos**.

New Members



18,937,266 **members!**

Meet people looking for sex in Maidstone

109,529 Online NOW!*

1,785,708 Members in **United Kingdom**

150,087 New Photos this week!

3,630 chatting now

14,089 using IM now

Join for **FREE!**



Convicts Reunited

Log In



Friends Reunited



BULLIES REUNITED

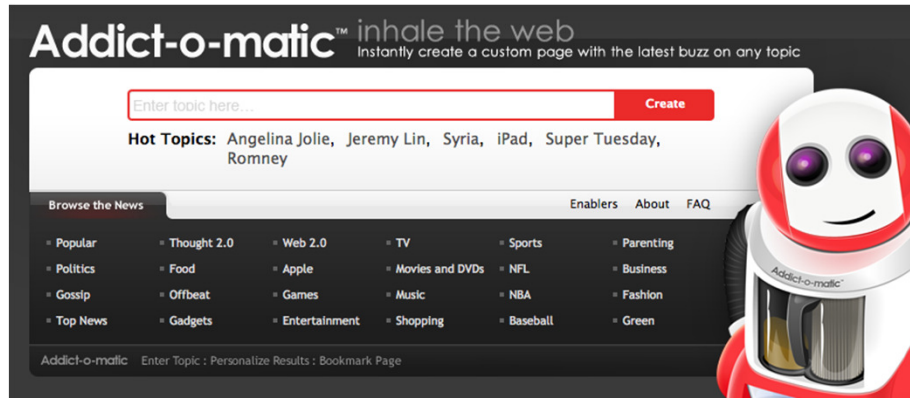


Forces Reunited
The British Armed Forces Community



exjw REUNITED

Social Media Monitoring

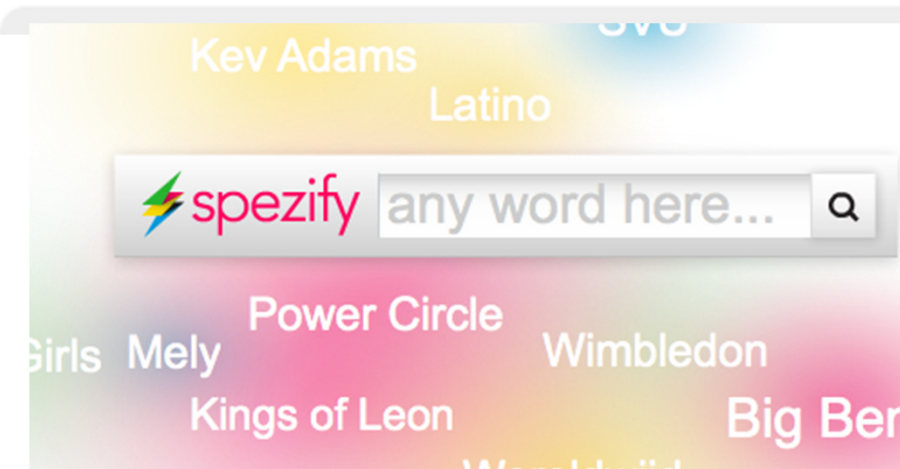


TOPSY

Real-time search for the social web

socialmention*

Real-time social media search and analysis:



KURRENTLY.com



September 12, 2013

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

Data Mining Sources

Bloomberg.com | Businessweek.com | Bloomberg TV

Bloomberg Businessweek

Global Economics

Companies & Industries

Politics & Policy

SILOBreaker
NEWS SEARCH ANALYSIS INSIGHT

 **DUE DIL**
Reinventing Due Diligence

September 12, 2013

opencorporates

The Open Database Of
The Corporate World

We have information on
59,123,360 companies

Search

search companies search officers

N
C
R
f
e

N
I
n

 SearchPeopleDirectory

White Pages

Yellow Pages

Reverse Phone

AdChoices 

[▶ Phone Book](#)

[▶ Phone Number](#)

Home » [Phone Books](#) » [South America & Central America Phone Book](#)

South America & Central America Phone Books

Search by

Phone Books
by Continent

Dialing Codes
by Continent

Reverse Phone
by Continent

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

Data Mining Sources



Visualisation



TouchGraph **NAVIGATOR**™ 

LinkedIn  **Maps**



Any Questions?

September 12, 2013

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

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