I-INTELLIGENCE

# Why Accuracy and Knowledge Still Beat AI

Skip Schiphorst | Chinese and Arabic OSINT-Instructor | I-Intelligence GmbH

I-INTELLIGENCE

SKOPENOW

- Introductions

  - Skip Schiphorst
    - Sinologist
    - OSINT-Language instructor @ I-Intelligence GmbH
    - Specialty: OSINT instruction in Chinese, Arabic and other languages

Skiposint#9557

@SchiphorstSkip

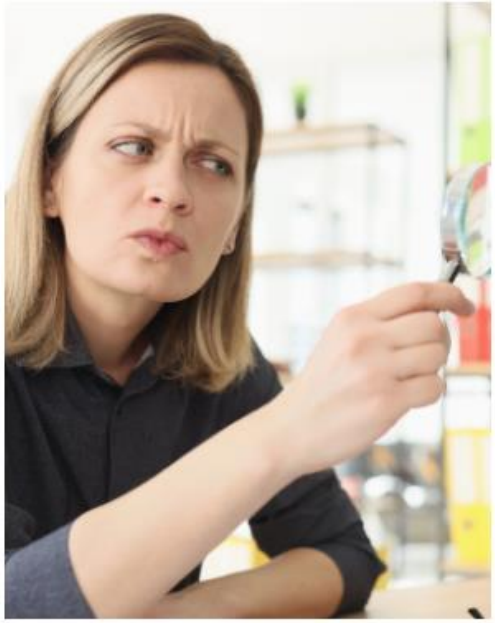
skip-schiphorst-721bb6200

www.i-intelligence.com

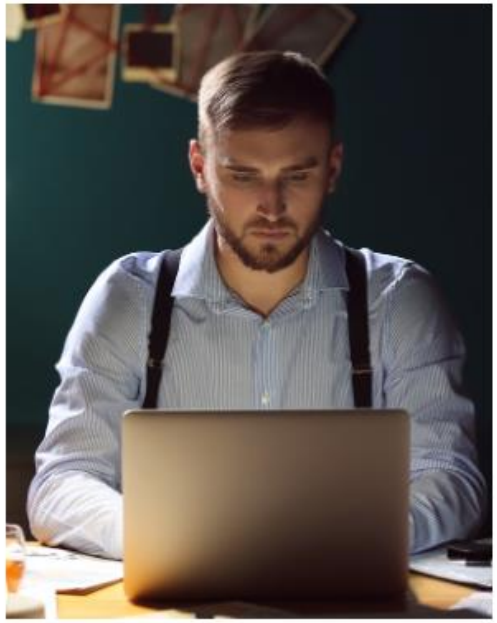
- Why Beat the “Buzzword”?

# Meet The SOCMINT Investigators

Jane



Veronica



Cedric

Jane

Fully reliant on AI-generated data

\+ Knocks out reports fast
\- Sometimes feels like she overlooked something

# Veronica

Old school investigations and sock-puppets for SOCMINT

\+ Detailed reports following every breadcrumb
\- Reports take a long time and often misses connecting links

Cedric

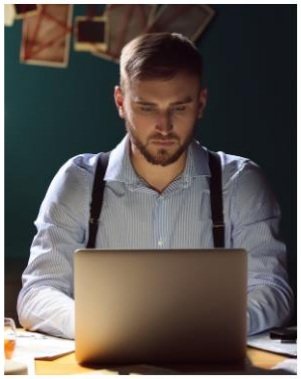
Relies on paid subscriptions for SOCMINT

+ Connects dots and finds new leads rapidly
- Often knows that every data point is worth more time, but doesn’t have “the time” due to the amount of data

# Who’s Going to Research Foreign Language Content?

Jane
AI

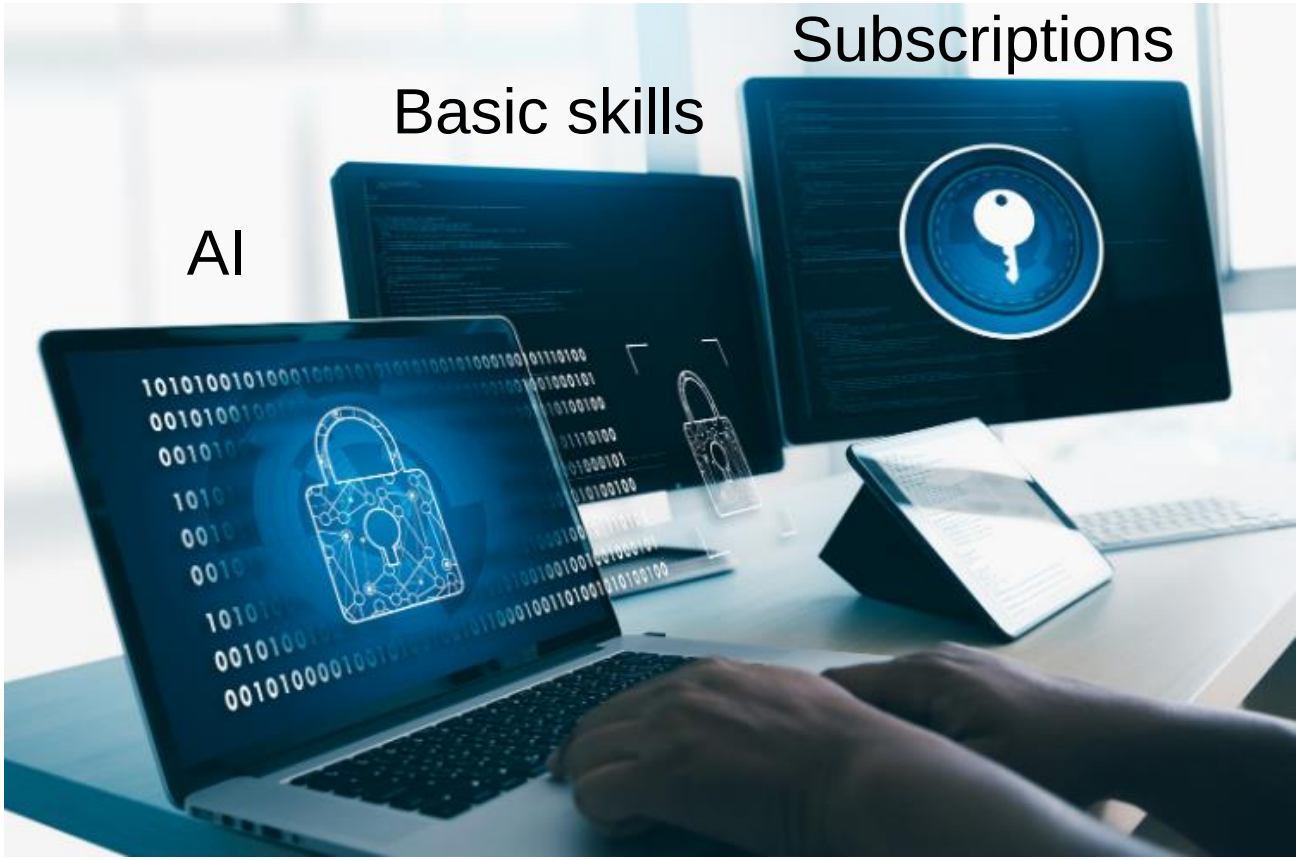
Veronica
Old school



Cedric
Subscriptions

Bernard
Spanish speaker
Urdu speaker
Korean speaker

# Force Multiplication



# Foreign Language OSINT Fundamentals & How to Beat AI with Human Expertise

# Methodology

- Fundamentals
  - Search formula:

Language identification and scope
Keywords
Search operators
\+ Source discovery / SOCMINT

---

**Results**

# Obstacles

即使这没有发生也是真相

مګر دا ریښتیا ده حتی که دا پیښ شوي نه وي

*Але це таки правда,*
*навіть якщо цього і не*
*відбулось.*

하지만 실제로 일어나지 않았더라도 진실입니다

- Common Obstacles
  - Different scripts
  - Text orientation
  - Text direction in the browser
  - Grammatical structure
  - How particles and verbs affect nouns
  - Available online translation software

- Language Identification Tools

- Language Identification Tools
  - One is none!
    - https://www.lexicool.com/language-identifier-guesser.asp
    - https://translate.google.com
    - https://translatedlabs.com/language-identifier
    - https://fanyi.baidu.com/
    - https://www.deepl.com/en/translator
    - https://www.tomedes.com/tools/language-detector
    - https://translate.yandex.com/

- Accurate Vocabulary Use
  - Beware of online dictionaries
  - Cross-check with several sources
  - If it’s taking long – you’re doing it right!

- Easy Task

  - John Smith who studied at the Normal University (B.a. of Philately) in 2015 has been cheering racist quotes.

  - Task: Find other Social Media sites this person is active on and gather all possible identifiers and selectors.

- Medium - Easy Task

  - A person associated with telephone number + 264 555 23 68 has been harassing people on dating platforms.

  - Task: Who is behind this number?

- Hard Task
  - Task: Determine the general sentiment on social media in Malaysia vis-à-vis the August 2023 state elections.

# Keywords

- Encyclopedias
  - Excellent resources to confirm the accuracy of a translation
  - Regional happenings – regional sources – regional languages

# Politics of Malaysia

11 languages

Article Talk Read Edit View history Tools

From Wikipedia, the free encyclopedia

**Politics of Malaysia** takes place in the framework of a federal representative democratic constitutional monarchy, in which the Yang di-Pertuan Agong is head of state and the Prime Minister of Malaysia is the head of government. Executive power is exercised by the federal government and the 13 state governments. Legislative power is vested in the federal parliament and the 13 state assemblies. The judiciary is independent of the executive and the legislature, though the executive maintains a certain level of influence in the appointment of judges to the courts.

The Constitution of Malaysia is codified and the system of government is based on the Westminster system. The hierarchy of authority in Malaysia, in accordance to the Federal Constitution, stipulates the three branches (administrative components) of the Malaysian government as consisting of the Executive, Judiciary and Legislative branch. Whereas, the Parliament consists of the *Dewan Negara* (Upper House/Senate) and *Dewan Rakyat* (Lower House/House of Representatives).[1]

This article is part of a series on the

**Politics of Malaysia**

| | |
|---|---|
| Head of State | show |
| Legislature | show |
| Executive | show |
| Judiciary | show |
| Elections | show |

# Politik Malaysia

11 bahasa

Rencana Perbincangan Baca Sunting Sunting sumber Lihat sejarah Peralatan

Daripada Wikipedia, ensiklopedia bebas.

Malaysia mengamalkan sistem demokrasi berparlimen di bawah pentadbiran raja berperlembagaan. Malaysia diketuai oleh Seri Paduka Baginda Yang di-Pertuan Agong yang dipilih daripada sembilan sultan negeri Melayu untuk berkhidmat selama lima tahun sebagai Ketua Negara dan Pemerintah Tertinggi Angkatan Tentera.

Sistem ini adalah berdasarkan sistem Westminster kerana Malaysia merupakan bekas tanah jajahan British. Kuasa eksekutif ditetapkan oleh kabinet yang dipimpin oleh Perdana Menteri. Berdasarkan perlembagaan Malaysia, Perdana Menteri mestilah seorang anggota Dewan Rakyat, yang pada pendapat Yang di-Pertuan Agong, memimpin kumpulan majoriti dalam parlimen. Manakala kabinet merupakan ahli parlimen yang dipilih daripada Dewan Rakyat atau Dewan Negara.

**Politik Malaysia**

| Negara Anggota ASEAN | |
|---|---|
| Kerajaan Persekutuan | [Buka] |
| Pemerintahan beraja | [Buka] |
| Legislatif | [Buka] |

- Vocabulary Lists – Words are Key!
  - Build “on-the-go” vocabulary lists
  - Sharing is caring – group effort
  - Correct terms will give you actionable results

## CHINESE VOCABULARY LIST - CHINESE MARKETS

| | | | | |
|---|---|---|---|---|
| after-sale service | 售后服务 | | interest rate | 利率 |
| Alibaba cloud services | 阿里云服务 | | internal control | 内控 |
| Aluminum Industry | 铝业 | | International Economy & Trade | 国际经贸 |
| Asia-Pacific | 亚太区 | | Internet Technology Consulting | 互联网技术咨询 |
| Assc.of South East Asian Nations (ASEAN) | 东南亚国家联盟 | | investment | 投资 |
| Asset | 资产 | | investment banking | 投资银行 |
| Auto Parts | 汽车零部件 | | Limited Company | 有限公司 |
| banking finance; financial | 金融 | | Limited Liability Company | 有限责任公司 |
| Big Data | 大数据 | | Lithium Industry | 锂业 |
| Biochemical Pharmaceutical | 生化医药 | | luxury goods | 奢侈品 |
| Biotechnology | 生物科技 | | Macau | 澳门 |
| Bitcoin | 比特币 | | macro-economic | 宏观经济 |
| Blockchain | 区块链 | | manufacture | 制造；制造业 |
| Branch Company | 分公司 | | market research | 市场调查 |
| brand | 品牌 | | market share | 市场占有率 |
| Branding strategy | 品牌战略 | | market survey | 市场调查 |
| BRIC countries | 金砖四国 | | market value | 市值 |

- Finding Vocab Lists – Save Time!
  - Language enthusiasts often share their vocabulary lists
  - Find these using keywords and search operators

Google

vocabulary list English Farsi media

Images | Pdf | With meaning | Videos | News | Books | Maps | Flights | Finance

About 47,400,000 results (0.42 seconds)

PersianPod101
https://www.persianpod101.com › ... · Translate this page

## The Top Persian Words you Need to Know for different Media

Learn the best **Persian vocabulary** for different **Media**. Learn new **words** and translations in this FREE PersianPod101.com lesson.

https://www.persianpod101.com › persian-vocabulary...

## Learn Persian with Free Vocabulary Lists

Learn **Persian vocabulary**, phrases and **words** FAST with TONS of FREE lessons! Always Updated. You also get BONUS Audio Lessons here at PersianPod101.

Amazon UK
https://www.amazon.co.uk › Persian-Essential-Eastern-...

## Media Persian (Essential Middle Eastern Vocabularies) ...

Buy **Media Persian** (Essential Middle Eastern **Vocabularies**) (Essential Middle ... Terms grouped in thematic sections * Easy-to-learn **lists** to test **translation** ...

★★★★☆ Rating: 4.1 · 19 reviews · £17.22

- Search Strategies
  - Preliminary search of the subject in English or another language you master.
  - Scope – 30’000ft view
  - Read - patience
  - Force search using a specific language

- Search Strategies
  - Narrowing down results
  - Identifying sources that will strengthen our information position

Google

ما هي عواقب إزالة الغابات في البرازيل؟

About 2,470,000 results (0.72 seconds)

نتيجة إزالة الغابات البرازيل

About 2,310,000 results (0.54 seconds)

"عاقبة" "إزالة الغابات" "البرازيل"

About 11,100 results (0.34 seconds)

filetype:pdf site:eg "عاقبة" "إزالة الغابات" "البرازيل"

About 2 results (0.40 seconds)

# Googling vs Online Research

- Advanced Search Operators
  - Find relevant content quicker
  - Foreign languages and alphabets
  - Often interchangeable with other browsers

- Phrasal Search
  - “Russian Internet”
  - “русский Интернет”

- The NOT (-) Operator
  - “BRICS” -”South America”
  - “金砖国家” -“南美”

- Filetype: Operator
  - Natural gas” filetype:pdf
  - “غاز طَبيعيّ” filetype:pdf

- Other browsers – other operators

- Search Operators per Browser
  - Google: https://shorturl.at/oCFQR
  - Bing: https://shorturl.at/fjQT4
  - Yandex: https://shorturl.at/jrsN2
  - DuckDuckGo: https://rb.gy/cb01l
  - Baidu: https://shorturl.at/qsTX5
  - Yahoo: https://rb.gy/gd29h

# The Importance of Diversifying Sources

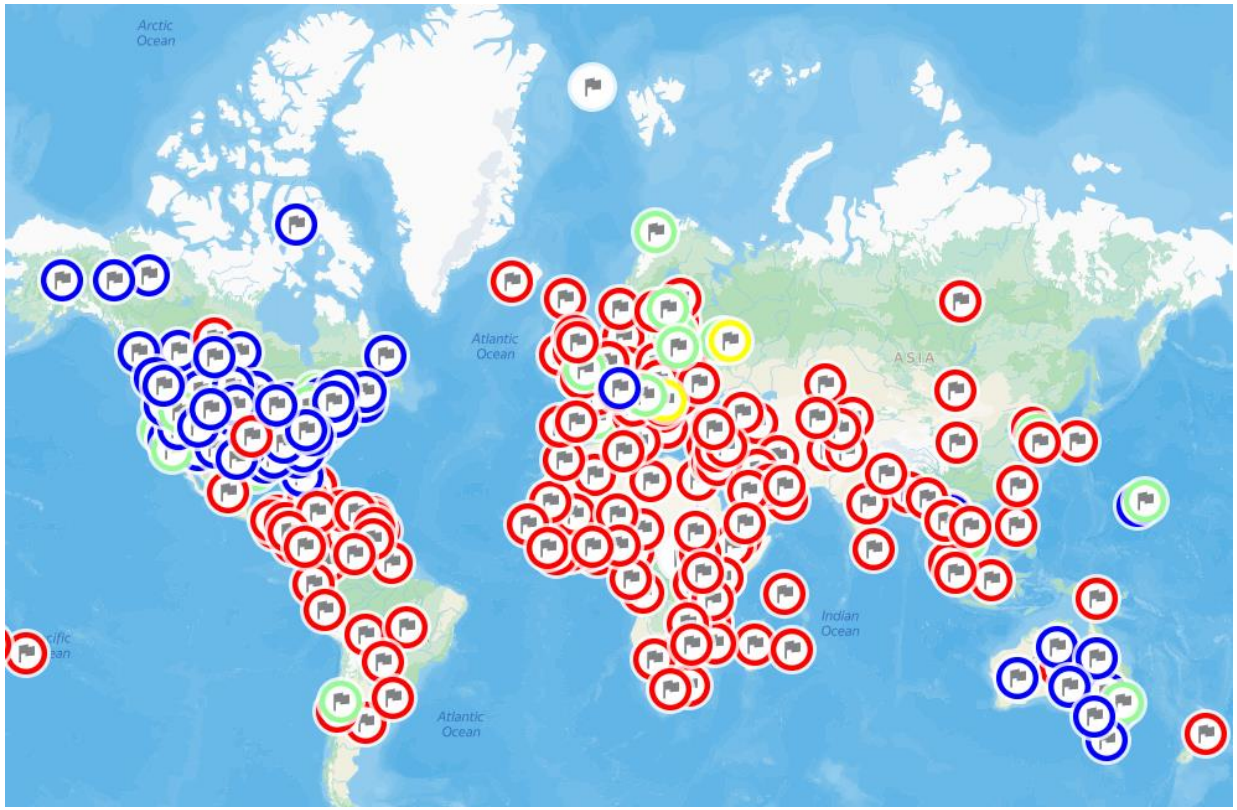
- Other Than English or “Western News”
  - Increase in other than English content
  - More perspectives - best information
  - Different opinions – they matter
  - News Directories

- International News Directories

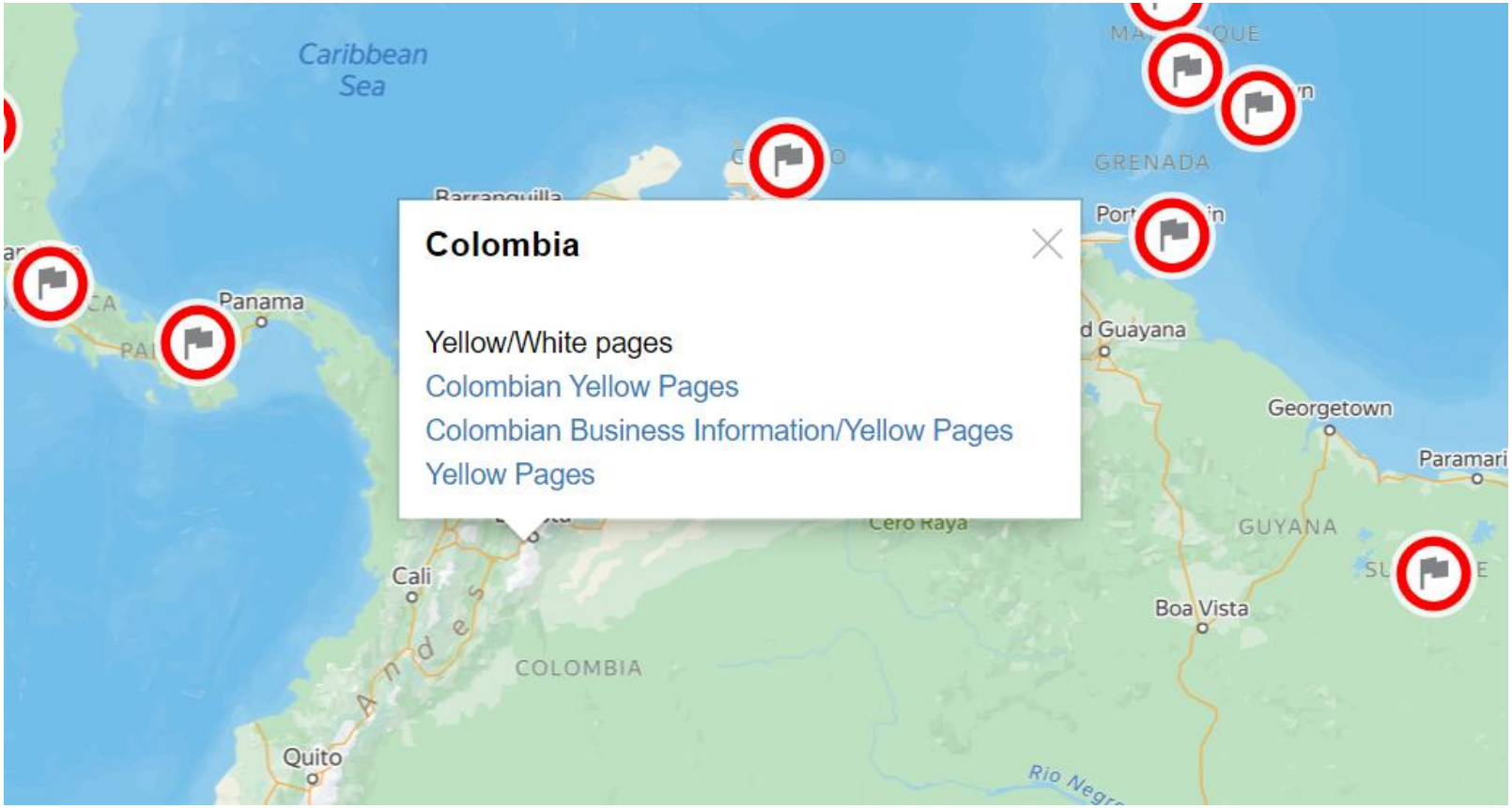
  - All You Can Read: www.allyoucanread.com
  - Online Newspapers www.onlinenewspapers.com
  - Newspaper Map: https://newspapermap.com
  - W3 Newspapers www.w3newspapers.com
  - ABYZ News Links: www.abyznewslinks.com
  - Newspaper Index: www.newspaperindex.com
  - World News: www.world-newspapers.com/countries

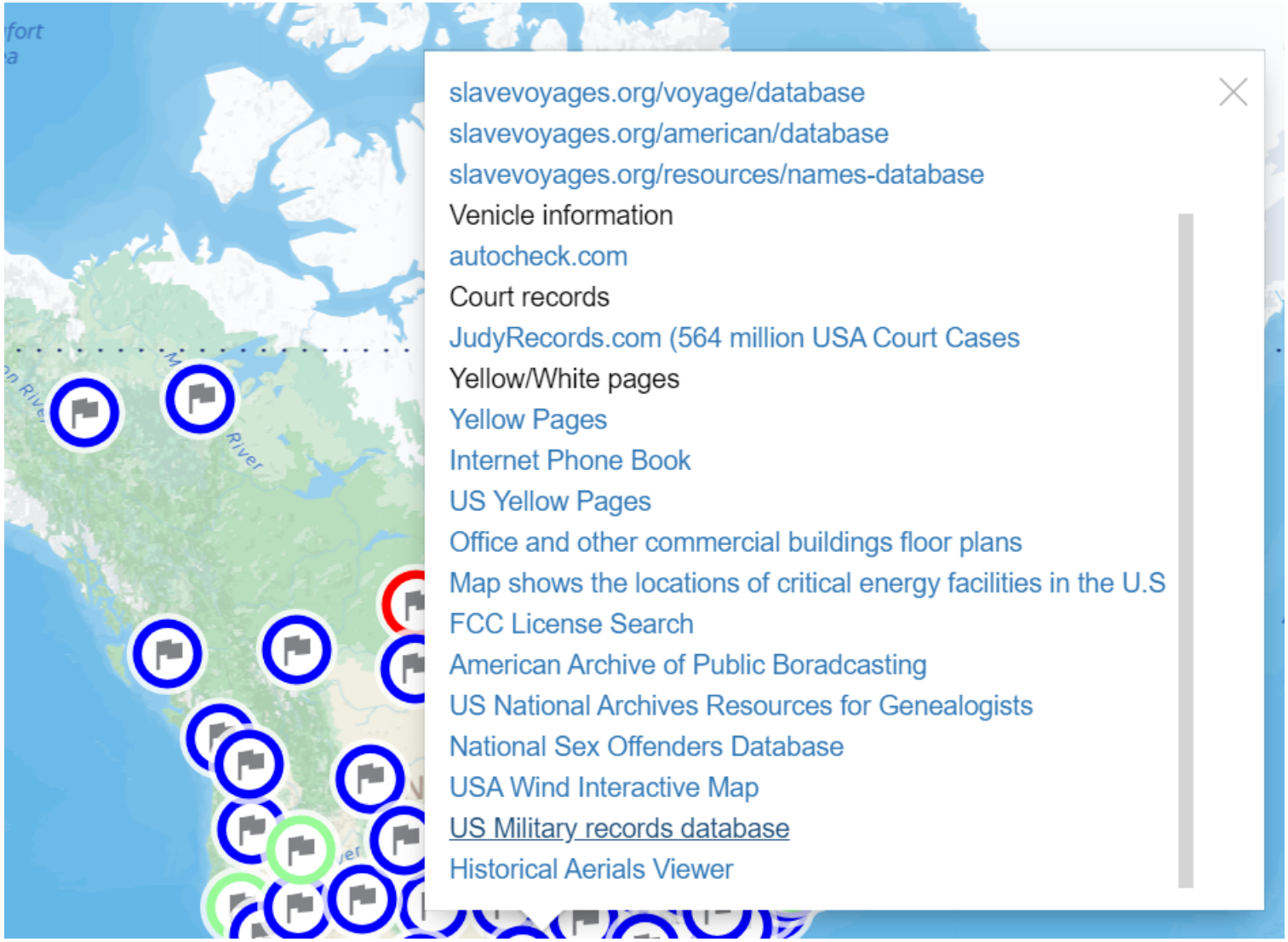
# Country Specific OSINT

- Worldwide OSINT tools map (GitHub)

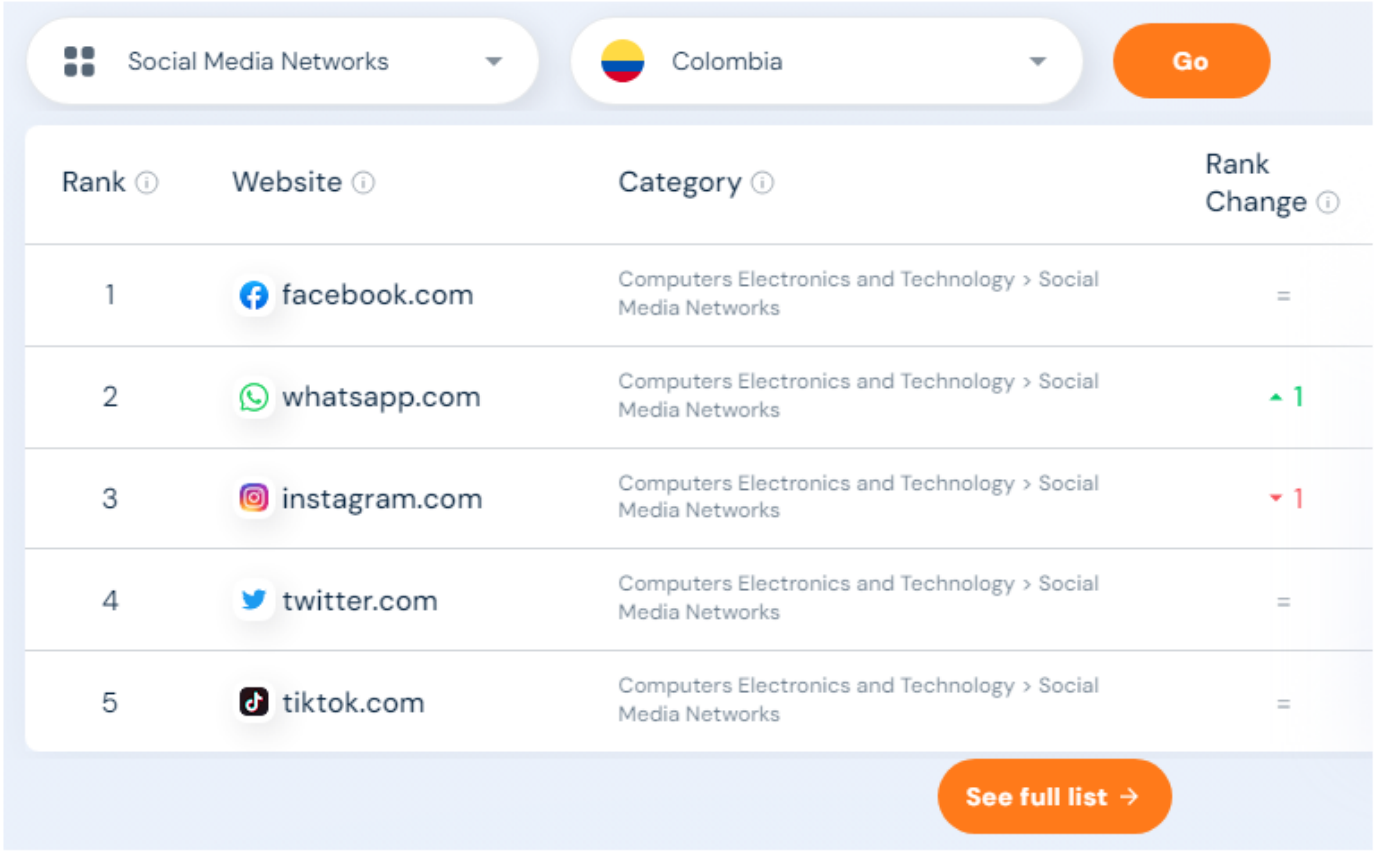



Source: http://cybdetective.com/osintmap/


Caribbean Sea
GRENADA
Colombia
Yellow/White pages
Colombian Yellow Pages
Colombian Business Information/Yellow Pages
Yellow Pages
Panama
Guayana
Georgetown
Paramari
Cero Raya
GUYANA
Cali
Boa Vista
COLOMBIA
Andes
Quito


slavevoyages.org/voyage/database
slavevoyages.org/american/database
slavevoyages.org/resources/names-database
Venicle information
autocheck.com
Court records
JudyRecords.com (564 million USA Court Cases
Yellow/White pages
Yellow Pages
Internet Phone Book
US Yellow Pages
Office and other commercial buildings floor plans
Map shows the locations of critical energy facilities in the U.S
FCC License Search
American Archive of Public Boradcasting
US National Archives Resources for Genealogists
National Sex Offenders Database
USA Wind Interactive Map
US Military records database
Historical Aerials Viewer

similarweb




Source: https://rb.gy/4qa6e

# Most popular social media in Colombia: these are the 8 most used social networks by Colombians

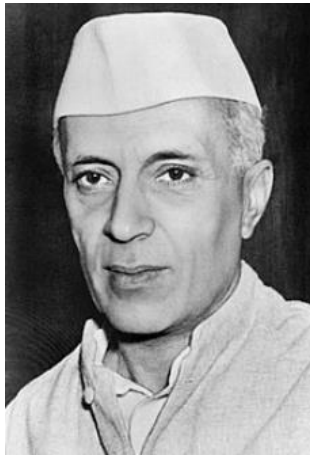
DANIELA CARRASCO x 7 DECEMBER, 2022

Source: https://rb.gy/z78ae

# International Naming Conventions

- Example
  - Johnathan Smith
  - John SMITH
  - John Smith
  - John S.
  - J. Smith
  - Smith, John
  - Mr. Smith
  - Dr. Smith
  - Smith, J
  - Johnny Smith
  - Smithy
  - “JS”
  - “The Wrench”

- Jawaharlal Nehru
  - Bengali জওহরলাল নেহেরু
  - Hindi जवाहरलाल नेहरू
  - Punjabi ਜਵਾਹਰ ਲਾਲ ਨਹਿਰੂ
  - Sindhi جواھر لال نھرو
  - Urdu جواہر لعل نہرو
  - Gujarati જવાહરલાલ નેહરુ
  - Sinhala ජවහර්ලාල් නේරු
  - Oriya ଜବାହରଲାଲ ନେହେରୁ
  - Malayalam ജവഹർലാൽ നെഹ്രു
  - Telugu జవాహర్ లాల్ నెహ్రూ
  - Tamil ஜவகர்லால் நேரு
  - Santali ᱡᱚᱣᱦᱟᱨᱞᱟᱞ ᱱᱮᱦᱨᱩ



Source: https://bit.ly/3YEJBTZ

# Law Enforcement Guide to INTERNATIONAL NAMES

Introduction ..... 3
Hispanic Names ..... 4
Arabic Names ..... 6
Indian Names ..... 8
Russian Names ..... 10
Polish Names ..... 11
South Asian Muslim Names ..... 12
Chinese Names ..... 13
Korean Names ..... 14
Vietnamese Names ..... 15
Japanese Names ..... 16


Source: https://shorturl.at/qJN01

WIKIPEDIA
The Free Encyclopedia

# Personal name

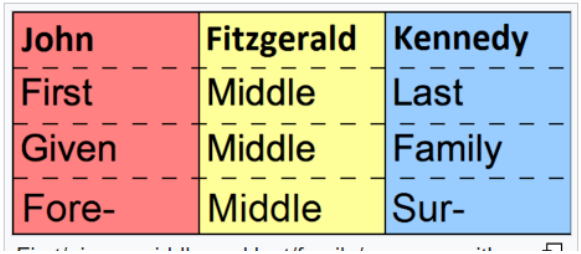
47 languages

Article Talk | Read Edit View history Tools

From Wikipedia, the free encyclopedia

This article **needs additional citations for verification**. Please help improve this article by adding citations to reliable sources. Unsourced material may be challenged and removed.
*Find sources:* "Personal name" – news · newspapers · books · scholar · JSTOR *(April 2010)* *(Learn how and when to remove this template message)*

A **personal name**, **full name**, or **prosoponym** (from Ancient Greek *prósōpon* - person, and *onoma* - name)[1] is the set of names by which an individual person is known, and that can be recited as a word-group, with the understanding that, taken together, they all relate to that one individual.[2] In many cultures, the term is synonymous with the *birth name* or *legal name* of the individual. In linguistic classification, personal names are studied within a specific onomastic discipline, called anthroponymy.[3]




Source: https://shorturl.at/eqJR6

Jason Greenberg Motamedi
Zafreen Jaffery
Allyson Hagen
Sun Young Yoon
Education Northwest

ies NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE
Institute of Education Sciences

REL NORTHWEST
Regional Educational Laboratory at Education Northwest

Source: https://shorturl.at/nwP15

# Researching it Before Running it

- Names on Social Media Using Arabic

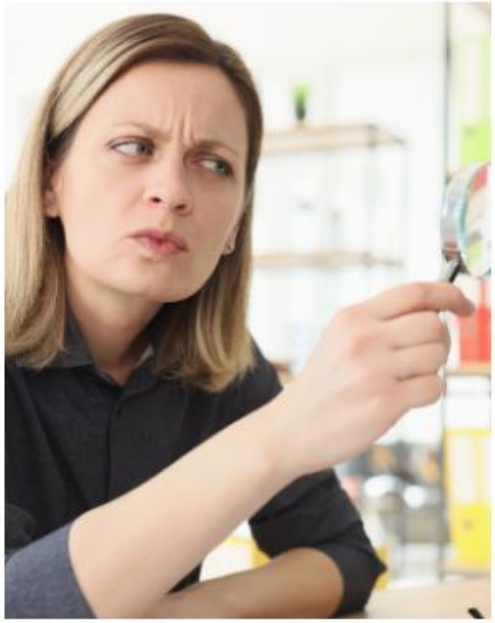
- Transliteration Methods - Chinese

Example: Wang Xuezhi

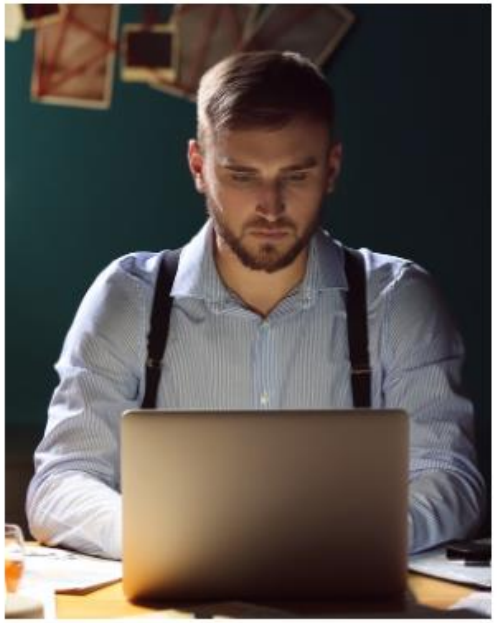
| | |
|---|---|
| 王学智 | Wong Hokzi |
| 王學智 | Huang Hokzi |
| 黄学智 | Xuezhi Wang |
| 黃學智 | XZ, Wang |
| 汪学智 | Wang, XZ |
| 汪學智 | Wang Xue Zhi |
| | Wang Xue-Zhi |
| | Etc… |

# Remember The Team

Jane



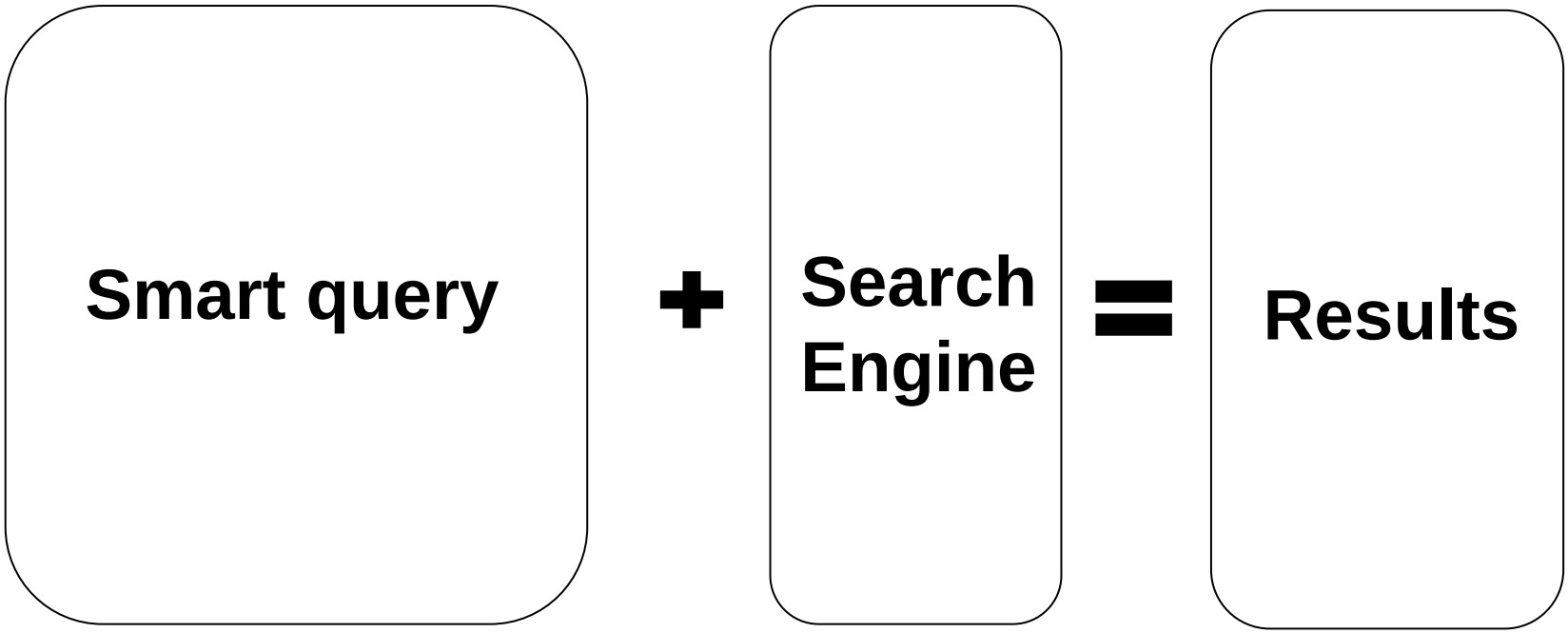
Veronica



Cedric

- Keep in Mind
  - Successful online research:
    - Identification of relevant **keywords** and **sources**
    - Knowledge of the **search engine** and its **operators**

- Keep in Mind



- I-Intelligence Language OSINT courses (five 3.5-hour sessions)

**Chinese OSINT**

**Arabic OSINT**

**Russian OSINT**

All I-Intelligence courses: https://i-intelligence.eu/training/course-index
**Promocode** for 10% discount: **IINT102023**

- Thank you Skopenow!

Skiposint#9557

@SchiphorstSkip

skip-schiphorst-721bb6200

www.i-intelligence.com

I-INTELLIGENCE