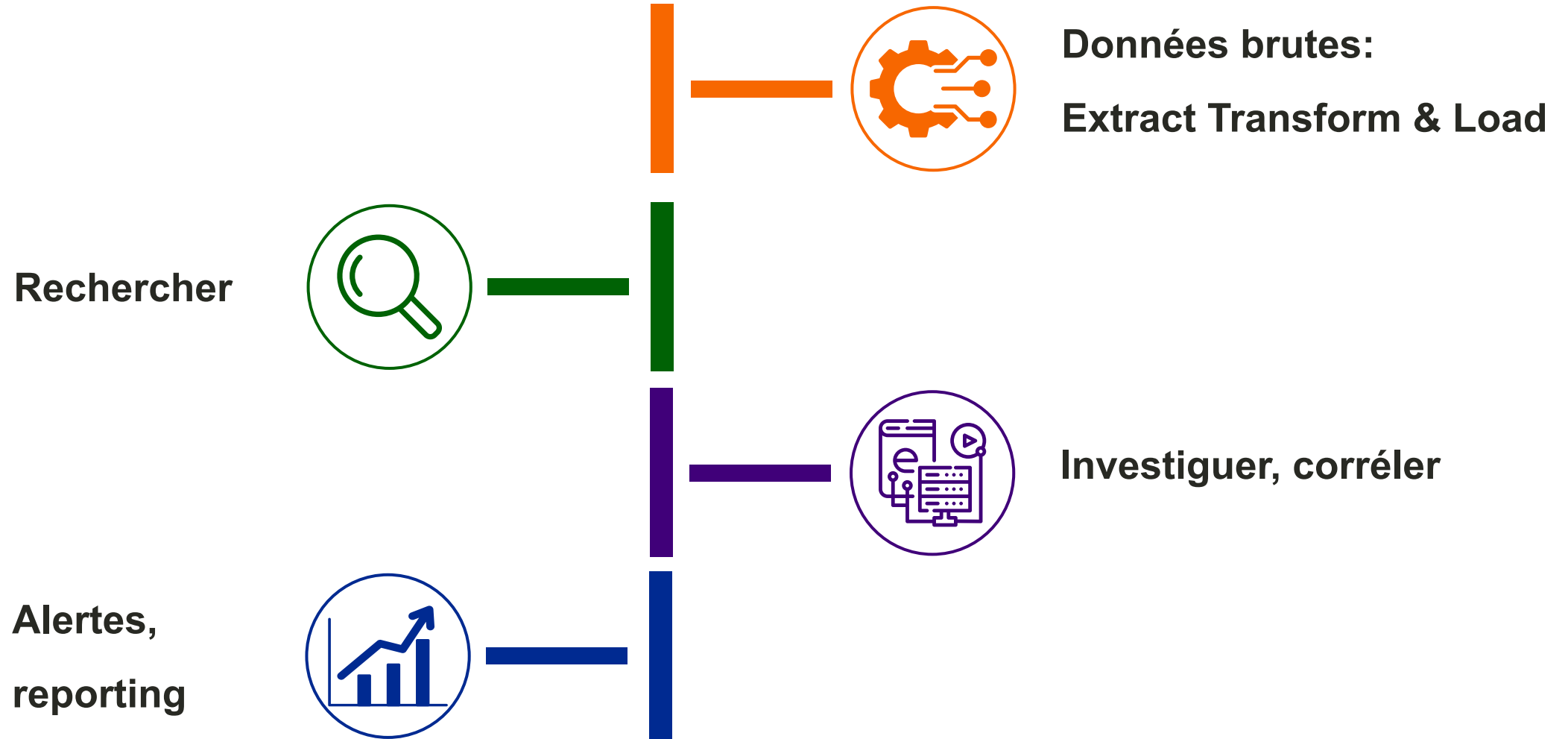




Décembre 2021



## Joe Hellerstein

Contributor



Joe Hellerstein is co-founder and chief strategy officer of [Trifacta](#) and the Jim Gray Chair of Computer Science at UC Berkeley.

In February 2010, The Economist published a report called “[Data, data everywhere](#).” Little did we know then just how simple the data landscape actually was. That is, comparatively speaking, when you consider the data realities we’re facing as we look to 2022.

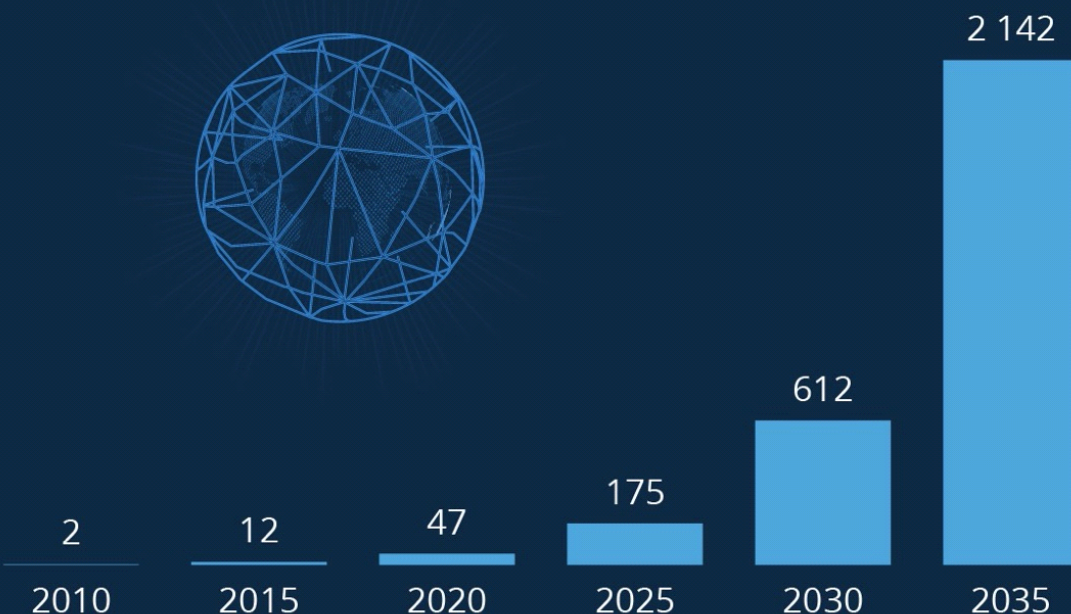
In that Economist report, I spoke about society entering an “Industrial Revolution of Data,” which kicked off with the excitement around Big Data and continues into our current era of data-driven AI. Many in the field expected this revolution to bring standardization, with more signal and less noise. Instead, we have *more noise, but a more powerful signal*. That is to say, we have harder data problems with bigger potential business outcomes.

Source: <https://techcrunch.com/2021/11/14/the-industrial-data-revolution-what-founders-got-wrong/>

CONFIDENTIEL - NE PAS DISTRIBUER

## Le big bang du big data

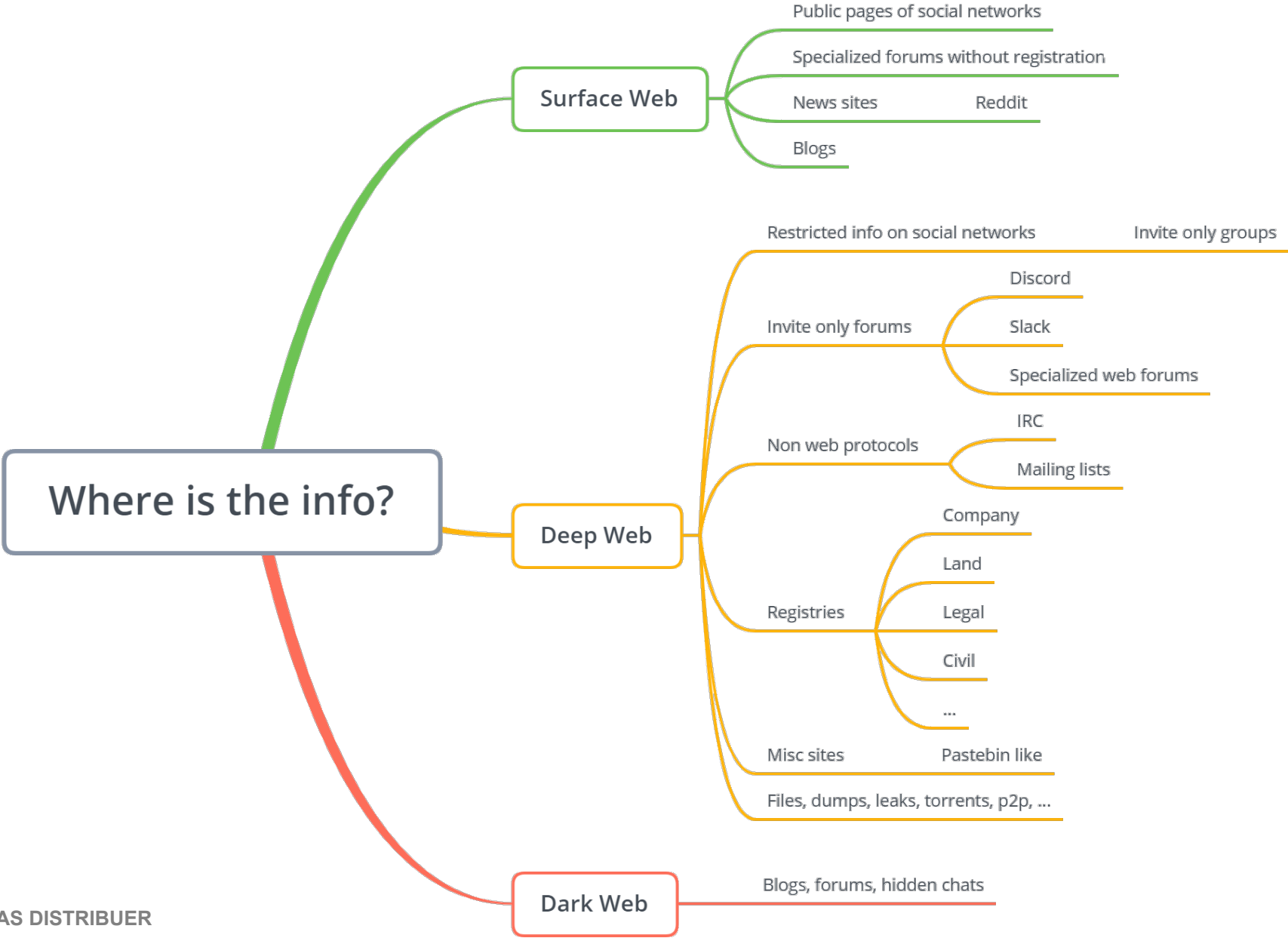
Volume annuel de données numériques créées à l'échelle mondiale depuis 2010, en zettaoctets \*



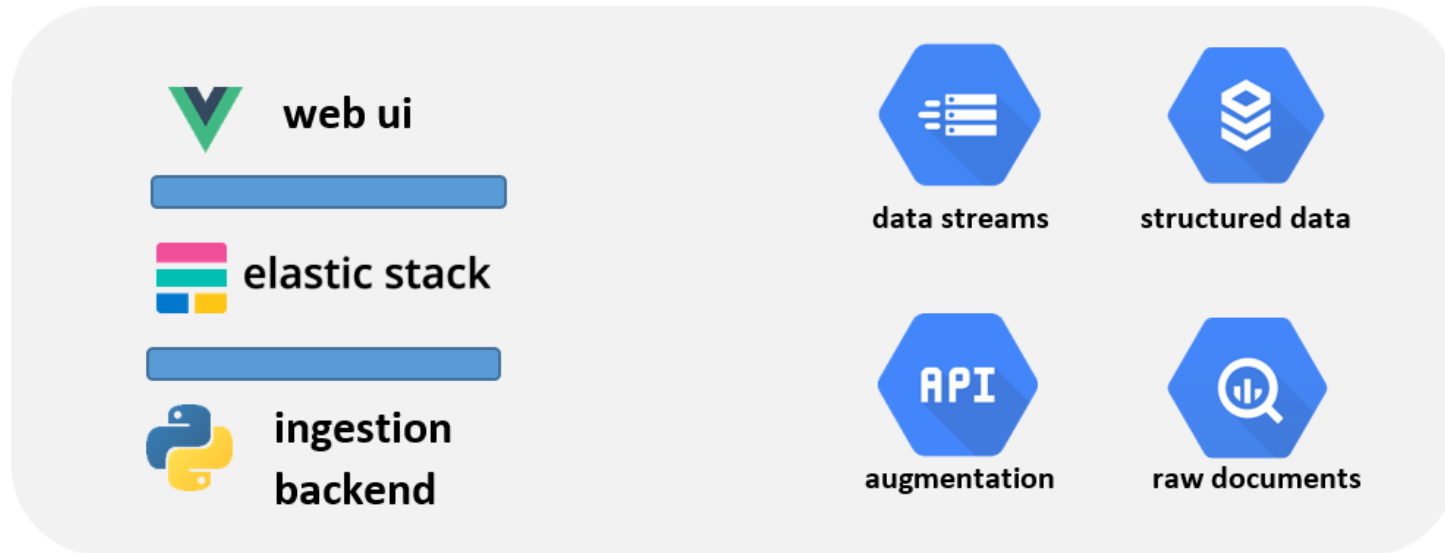
\* Prévisions de 2020 à 2035. Un zettaoctet équivaut à mille milliards de gigaoctets.  
Source : Statista Digital Economy Compass 2019



**statista** 

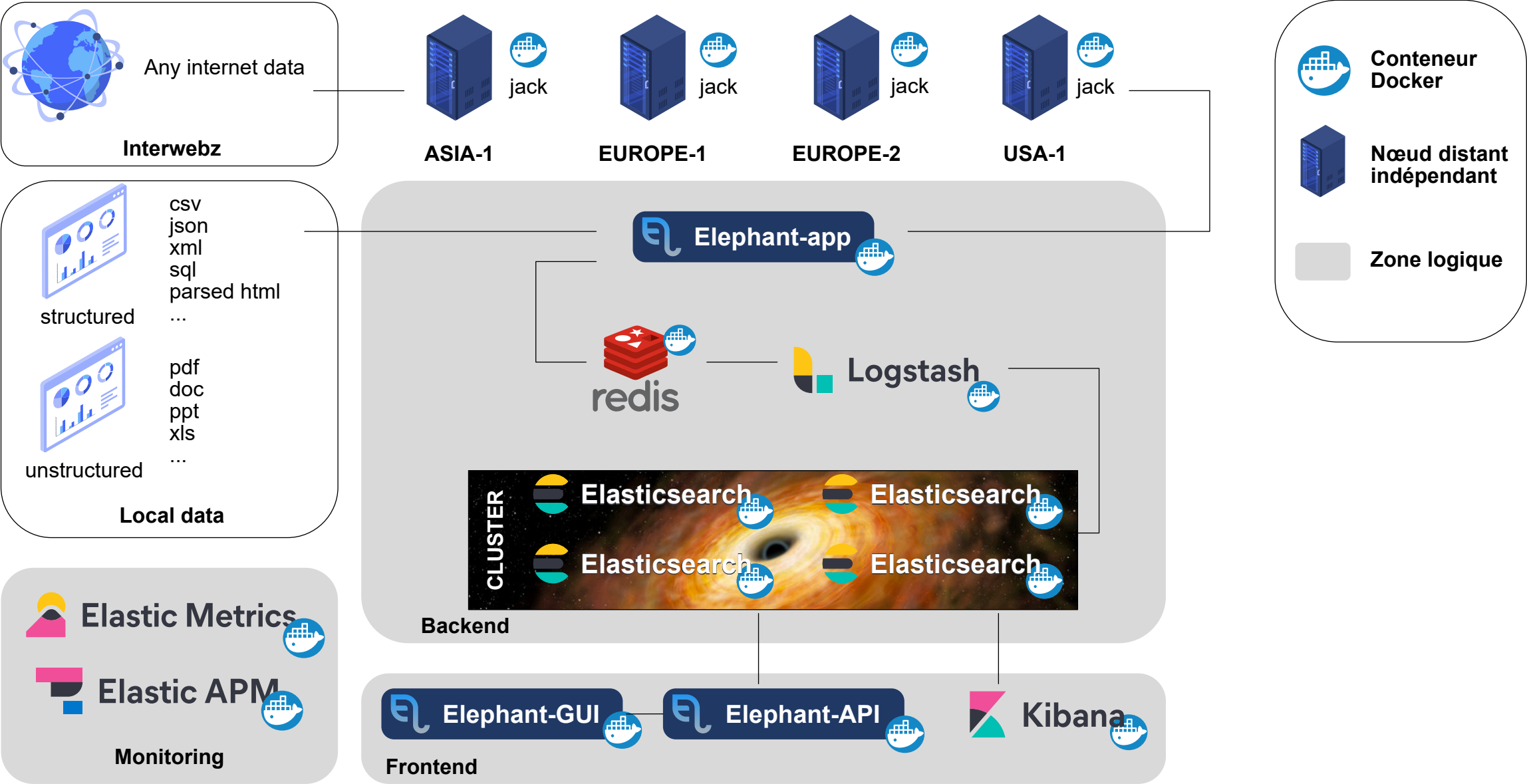


**Elephantastic** est une plateforme d'investigation et de traitement de **données hétérogènes** développée et opérée par Axis&Co. Fondée sur un *backend sur mesure* et associé à un *stack Elasticsearch*. Les données ne sont pas seulement indexées en *full text* mais aussi **normalisées** afin d'extraire des **données significatives** comme les noms d'utilisateurs, dates, adresses, informations de géolocalisation, emails, adresses IP, etc. Cette approche permet de **corréler les entités** dans l'application de façon simple et efficace.



ELEPHANTASTIC

“ From *information* to  
*intelligence* ”



Jack est facile à étendre et personnaliser grâce à du code python sans *boilerplate*. Il est modulaire par nature et instrumente des conteneurs dockers.

Ses principaux atouts:

- Pipelines en YAML dans un format intelligible par un humain
- Résultats immédiats
- Piles incluses:
  - Téléchargement http, ftp, webdav, ...
  - Parsing html, xml, json, ...
  - Extraction zip, rar, 7z, ...
  - Ingestion de fichiers et OCR
  - Manipulation de chaines de caractères et de nombres
  - Boucles et conditions
- Journalisation et gestion des erreurs
- Surveillance et planification via la GUI
- Agents distants



# JACK

*“Un moteur de scripts YAML pour gérer des séquences de tâches inter-reliées”*

COLLECTIONS

MAPPINGS

PIPELINES

SCHEDULED

JOBS

Filter

Label 

1

Last

Scheduled

Jobs

Action

Certificate Transparency Logs

log\_url

block\_index

Certificate logs are simple network services that maintain cryptographically assured, publicly auditable, append-only records of certificates. Includes hostnames and timestamps of certificates.

0

0

0

critical.io Service Fingerprints [Sonar]

The Critical.IO project was designed to uncover large-scale vulnerabilities on the global IPv4 internet and scanned a number of ports across between May 2012 and March 2013. The current dataset contain a monthly snapshot of each unique service response found on each responsive host. In the cases of services where a unique response is received for each request (HTTP, due to Date headers, etc), each of these will be represented as a unique record in one of the monthly JSON exports.

0

0

0

dnsrecon.py Domain Information - Get Domain Information [dnsrecon.py]

domain

Perform common DNS queries on a domain (SOA, NS, A, AAAA, MX and SRV).

0

0

0

Domain Name System (DNS) Queries - Get Hostname DNS Records

hostname

Any DNS queries performed

0

0

0

Domain Name System (DNS) Queries - Get IP DNS Records

ip

Any DNS queries performed

0

0

0

Find Domains From Email [domaintools.com]

email

List of partially obscured domain names associated with a given email obtained via the "reversewhois" API of domaintools.com

0

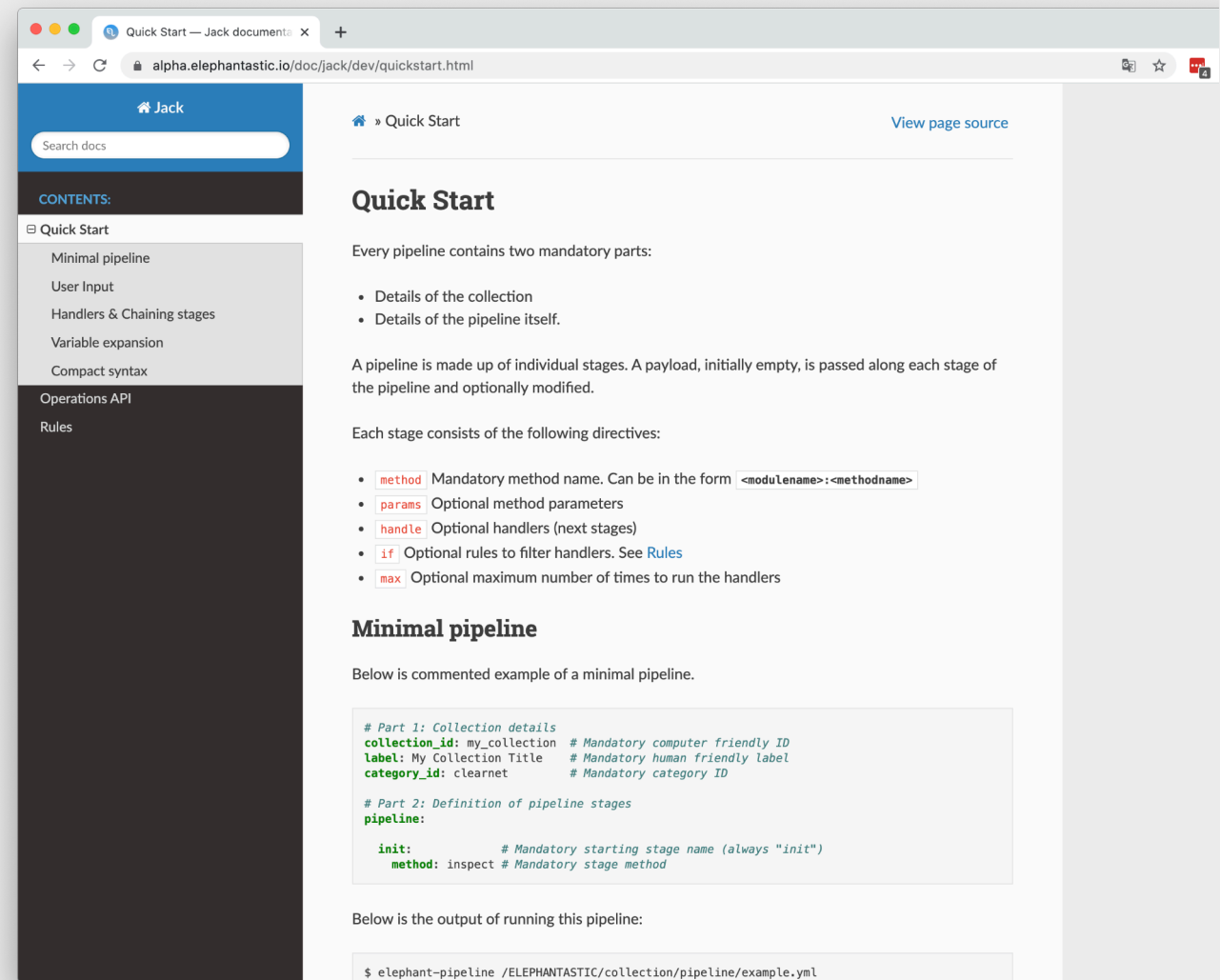
0

0



Notre philosophie: **ne pas réinventer la roue**. Une pléthore de solutions et d'outils existe déjà, la plupart pouvant être directement instrumentalisés et isolés dans un conteneur docker. Nous ne sommes dépendants d'aucun protocole, langage, librairie et pouvons lancer n'importe quel outil, par exemple:

- Binaires (nmap, dig, whois, ...)
- Programme Java (JBoss twiddle, JMX console, ...)
- Scripts dans n'importe quel langage (Perl, Ruby, Python, Bash, ...)
- Navigateur *headless* (Chrome, Firefox)
- Suite d'outils de Kali Linux
- Metasploit
- Bots IM (Telegram, Whatsapp, etc)
- ...



The screenshot shows a web browser displaying the 'Quick Start' page for Jack. The page has a dark sidebar with a search bar and a table of contents. The main content area is white and contains the following text:

## Quick Start

Every pipeline contains two mandatory parts:

- Details of the collection
- Details of the pipeline itself.

A pipeline is made up of individual stages. A payload, initially empty, is passed along each stage of the pipeline and optionally modified.

Each stage consists of the following directives:

- **method**: Mandatory method name. Can be in the form `<moduleName>:<methodName>`
- **params**: Optional method parameters
- **handle**: Optional handlers (next stages)
- **if**: Optional rules to filter handlers. See [Rules](#)
- **max**: Optional maximum number of times to run the handlers

### Minimal pipeline

Below is commented example of a minimal pipeline.

```
# Part 1: Collection details
collection_id: my_collection # Mandatory computer friendly ID
label: My Collection Title # Mandatory human friendly label
category_id: clearnet # Mandatory category ID

# Part 2: Definition of pipeline stages
pipeline:
  init: # Mandatory starting stage name (always "init")
    method: inspect # Mandatory stage method
```

Below is the output of running this pipeline:

```
$ elephant-pipeline /ELEPHANTASTIC/collection/pipeline/example.yml
```



# Cyber Threat Intelligence

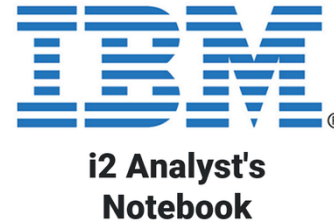


**CONFIDENTIEL - NE PAS DISTRIBUER**

# Historique et concurrence

---

- R&D interne
- Plateforme créée autour d'une méthodologie
- Marché de l'OSINT concurrentiel mais peu de solutions globales (plateformes, briques logicielles, fournisseurs de services)
- ... **Un marché énorme**



ASIRIS  
MONITORING  
PLATFORM





MS

## Architecte Elephantastic et développeur backend

Guru Python et Docker

Sysadmin, expert sécurité webapp

A passé trop de temps dans le code source Java de Maltego

Qualifié 5 ans consécutifs pour le CTF Defcon à Las Vegas



NDR

## Architecte Elephantastic

Pratique l'OSINT depuis 1997

Consultant Intelligence Economique & Sécurité des Système d'Information

A dirigé à la fois des investigations anti-contrefaçon et des projets de sécurité IT

Compétences Python, VueJS et Linux



RR

## Responsable des pipelines techniques Elephantastic

Reverse Engineering: x86, x86-64, arm, aarch64, CELL

Guru C

Pentester, chercheur de vulnérabilités

Expert Unix / Linux

Ceinture noire zmap

Qualifié 5 ans consécutifs pour le CTF Defcon à Las Vegas



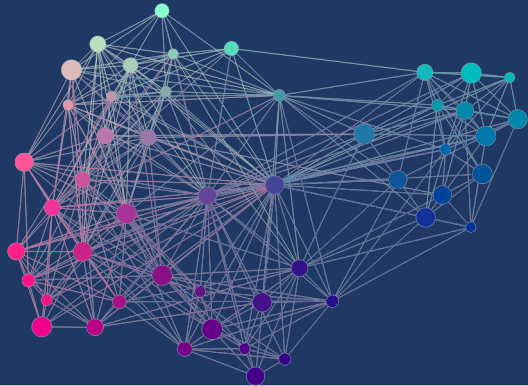
MK

## Développeur front Elephantastic

Expert VueJS expert

Magicien UX et UI

## Analyse visuelle



## Machine Learning / NLP

- Named Entity Recognition
- Categorisation
- Traduction, résumé

## RGPD

## Case Management

- Méthodologie Axis&Co
- Collaboration
- Ecrans orientés métiers

## Pipelines

- Nouveaux pipelines
- Assistant création pipeline
- Veille OSINT

Text

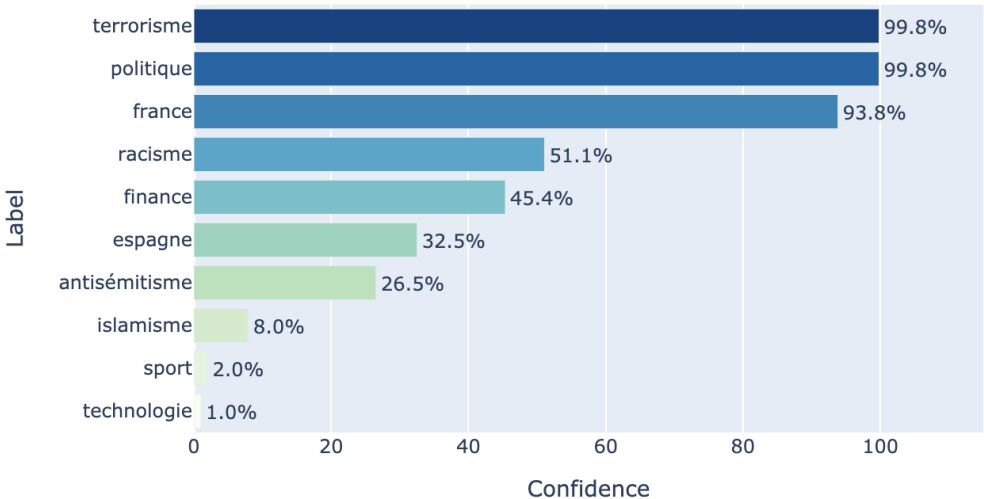
Je vais tuer le président ce soir

Possible topics (separated by ``,` `)

religion, terrorisme, technologie, aviation, football, sport, islamisme, politique, france, espagn163/1000

☒ Allow multiple correct topics

Top Predictions



Text

Je vais tuer tout le monde dans Call of Duty ce soir

Possible topics (separated by ``,` `)

religion, terrorisme, technologie, aviation, football, sport, islamisme, politique, france, espagn163/1000

☒ Allow multiple correct topics

Top Predictions



brandporter

Global Overview

Projects

53

Saved Queries

70

Last 7 days

989

Yesterday

151

Today

111

Jobs in the Last 14 days

numlookup - Phone number lookup

✓ 3

▲ 0

📄 3

Whatsapp API - Get Whatsapp account from phone

✓ 3

▲ 0

📄 2

Telegram Messenger - Get Telegram account from phone

✓ 3

▲ 0

📄 0

Skype - Search by phone

✓ 3

▲ 0

📄 0

Find Emails From Domain [email-format.com]

✓ 3

▲ 0

📄 0

DNS Records - Get Hostname DNS Records

✓ 5

▲ 0

📄 20

archive.org - Take screenshot of URL

✓ 2

▲ 1

📄 15

Website Screenshots - Take screenshot of URL

✓ 2

▲ 1

📄 1

Whois (viewdns)

✓ 3

▲ 0

📄 3

Reverse IP (viewdns)

✓ 3

▲ 0

📄 10.0k

Whois history (securitytrails)

✓ 3

▲ 0

📄 0

Subdomains (securitytrails)

✓ 2

▲ 1

📄 68

Quick actions

Browse investigations

Search documents

Graph

Browse collections

Notifications

brandporter

Updated 2 months ago

🔍 2

📄 0

📄 0

📄 0

ciceron

Updated 8 months ago

🔍 11

📄 2

📄 1

📄 0

CLIPPER

Updated a month ago

🔍 0

📄 0

📄 0

📄 0

Crypto Scam

Updated a month ago

🔍 0

📄 0

📄 0

📄 0

CUIVRE

Updated a month ago

🔍 2

📄 0

📄 0

📄 0

Demo

Updated a year ago

🔍 11

📄 1

📄 0

📄 0

DI

Updated 5 months ago

🔍 0

📄 0

📄 0

📄 0

ERATO

Updated 2 months ago

🔍 7

📄 0

📄 0

📄 0

FAKIR 1

🔍 1

📄 0

📄 0

📄 0

brandporter

BRANDPORTER

SEARCHES (2)

SEARCH HISTORY (12)

TAGS (0)

UPLOADS (0)

axis-a.com \* (95)

Advanced axis-a.com

FiltersTimeframeRaw (JSON)

FIELD	OPERATOR	VALUE(S)
-------	----------	----------

OVERVIEWRESULTSGEO (9)IMAGESANALYTICS

22 Whois Records

COLLECTION	LABEL	DATE
Who.is whois data from 2013	axis-a.com	01/12/2013
Who.is whois data from 2013	axis-a.net	01/12/2013
Who.is whois data from 2013	axis-a.org	01/12/2013
Whois history (securitytrails)	2012-11-29: whois axis-a.com	29/11/2012
Whois history (securitytrails)	2013-11-29: whois axis-a.com	29/11/2013
Whois history (securitytrails)	2011-11-29: whois axis-a.com	29/11/2011
Whois history (securitytrails)	2013-11-29: whois axis-a.com	29/11/2013
Whois history (securitytrails)	2015-01-20: whois axis-a.com	20/01/2015
Whois history (securitytrails)	2015-09-14: whois axis-a.com	14/09/2015
Whois history (securitytrails)	2015-09-14: whois axis-a.com	14/09/2015

9 Certificates

DETAILORIGINALJSON

axis-a.com

Tags: + New Tag

WhoisRecord

Who.is Whois Data From 2013

Internet Infrastructure

Created: 01/12/2013 01:00

Last Stored: 14/05/2021 19:13

Last Updated: 14/05/2021 19:13

Fingerprints

Emails

Hostnames

Keywords

Names

Usernames

DOMAIN@AXIS-A.COM 11

axis-a.com 95

AXIS-A.COM 95

axis-a 7.3k

BERTRAND DE GUERMINER 11

DOMAIN 16.2M

Document

Do Not Translate Highlight Keywords

Domain

axis-a.com

Email

DOMAIN@AXIS-A.COM

Hostname

axis-a.com

Name

BERTRAND DE GUERMINER



brandporter

BRANDPORTER

SEARCHES (2)

brandporter.ru - dns

certificates

SEARCH HISTORY (12)

tooke2007@gmail.com

loopooyt

+85295810373

brandporter.ru

199.33.123.227

yufang1990789@163.com

7x24customer@care@gmail.com

+85255160582

+85295810373

+85266473671

199.33.123.226

brandporter.ru

TAGS (0)

UPLOADS (0)

axis-a.com \* (95)

+85295810373 (8)

199.33.123.227 (13)

ip 199.33.123.227

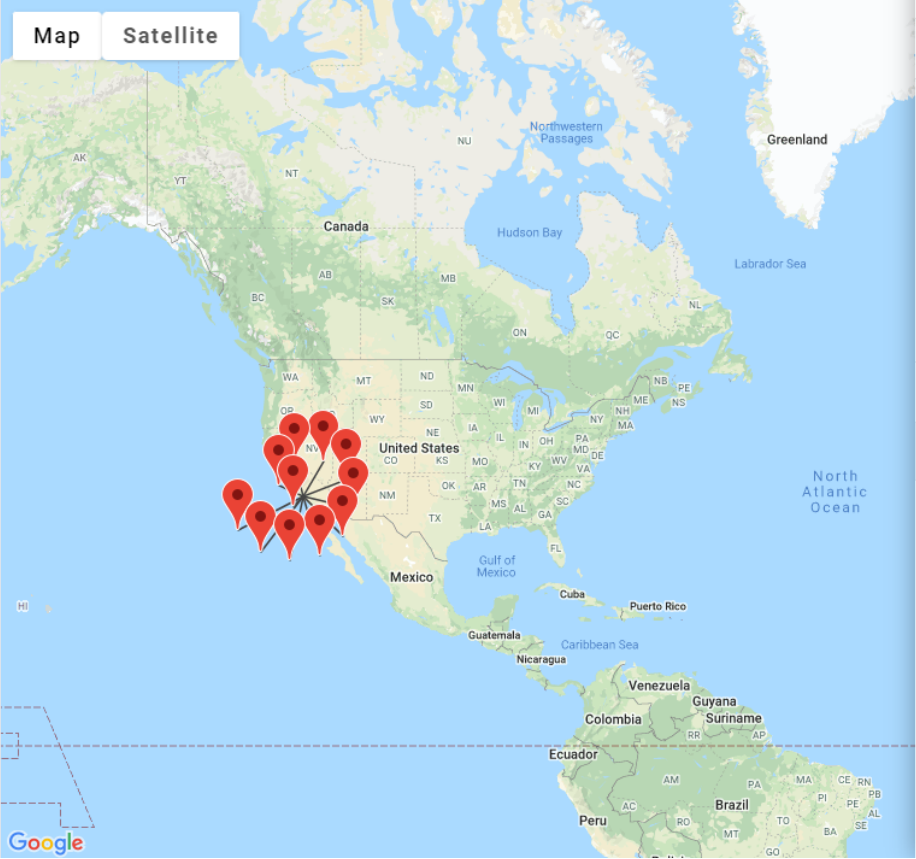
OVERVIEWRESULTS

GEO (11)

IMAGES

ANALYTICS

MapSatellite



DETAILORIGINALJSON

AbuseIPdb report: 199.33.123.227

Tags

LogRecordCreated: 06/10/2021 12:21

AbuseIPDBLast Stored: 06/10/2021 12:21

Internet InfrastructureLast Updated: 06/10/2021 12:21

Fingerprints

Coordinates34.106,-118.360743.8k

Countriesus

Hostnamesrebelhosting.net111

199.33.124.227.rebelhosting.net2

Ips199.33.123.22713

NamesRebel Hosting101

Document

Do Not TranslateShow EditorHighlight Keywords

AbuseIPdb report: 199.33.123.227No report for this IP



AXIS&CO

33 Avenue du Maine - 75015 Paris  
+33 (0)1 80 96 86 00

[www.axis-a.com](http://www.axis-a.com)  
[paris@axis-a.com](mailto:paris@axis-a.com)

[elephantastic.io](http://elephantastic.io)