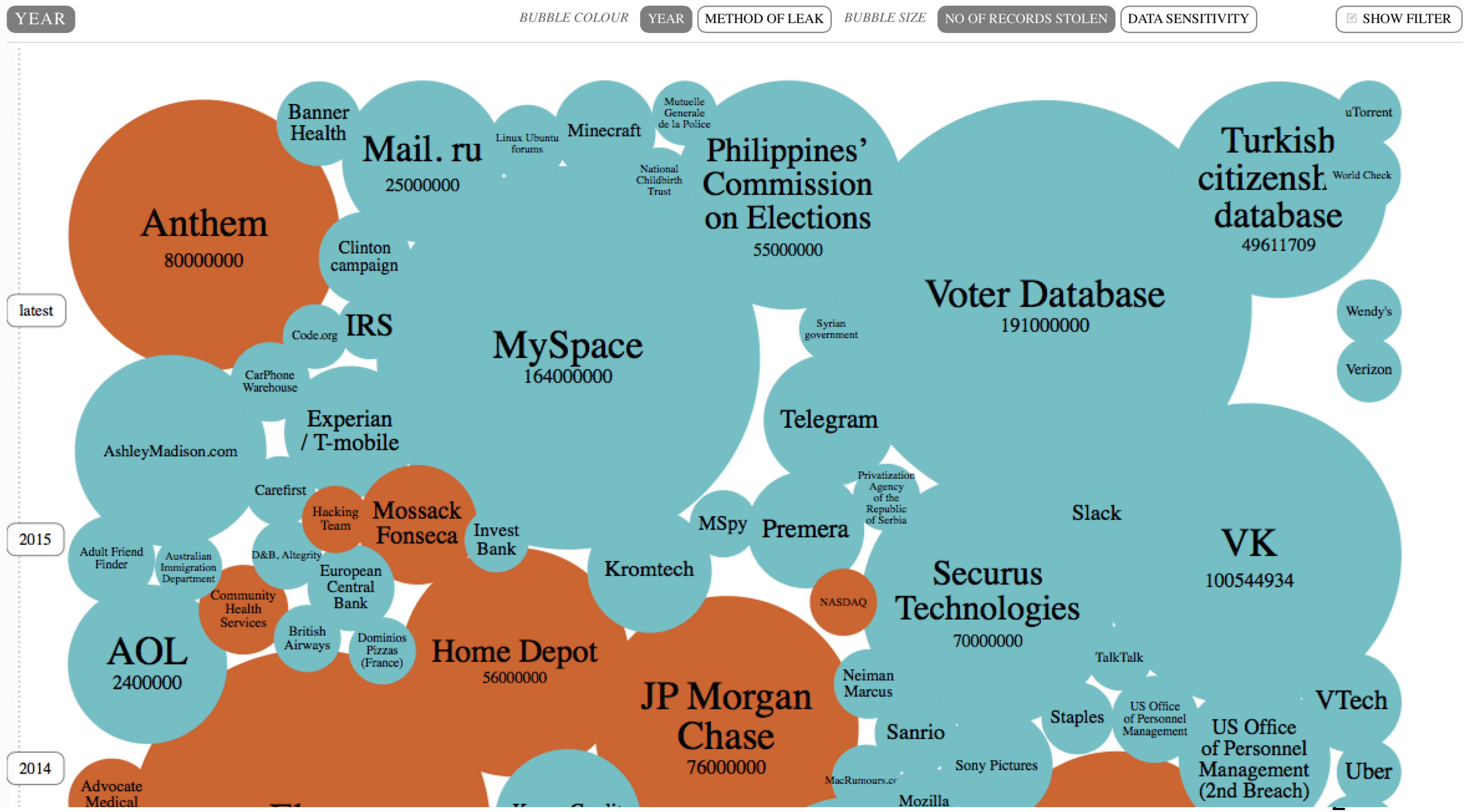
# inpher.io

search and operate on encrypted data

# Why?

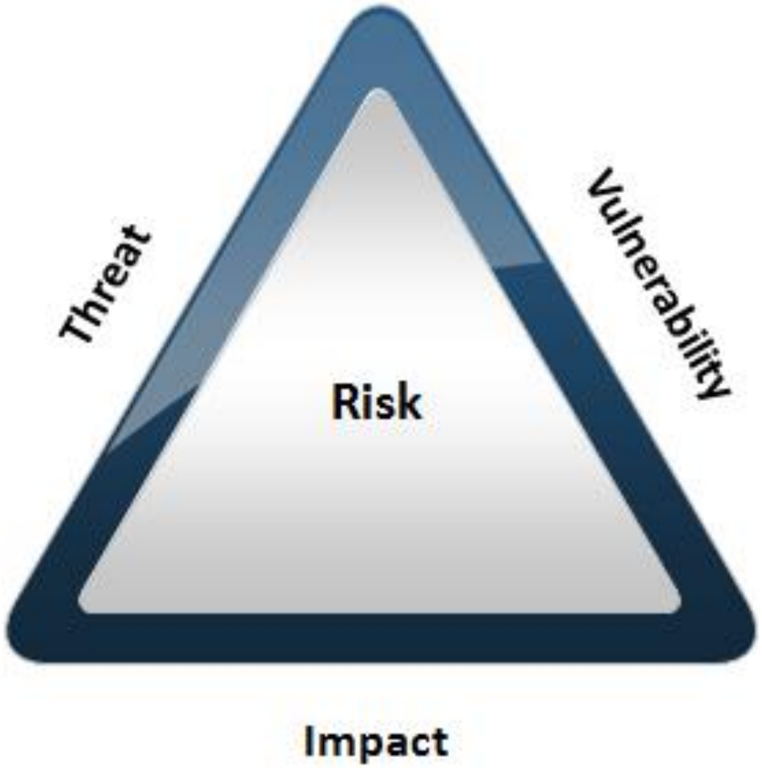


Inpher Confider

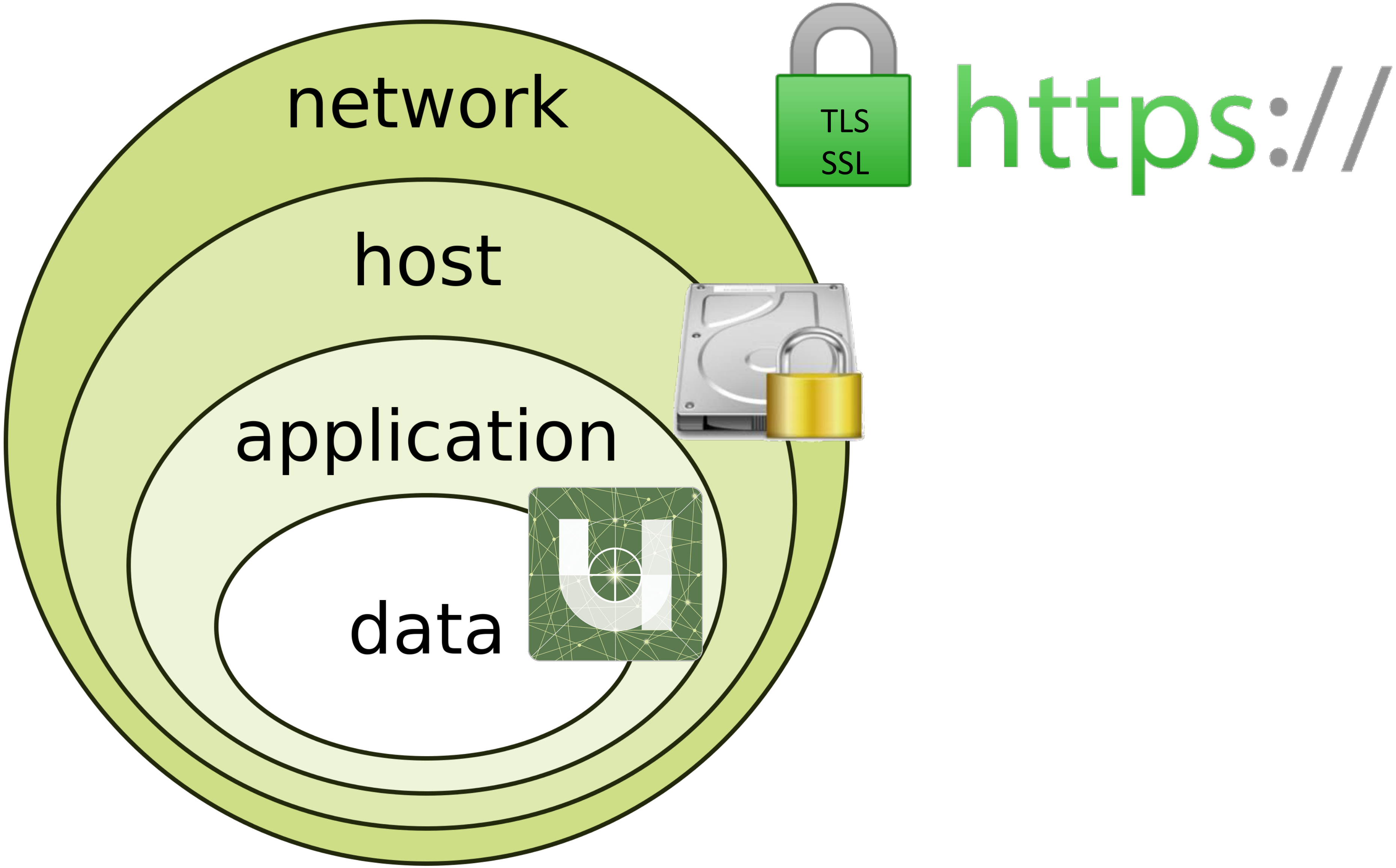
[http:// /]

# What?

- Company  vs. product considerations
- What are the assets?
- What is the associated risk?
- How can you protect?



# How?



network

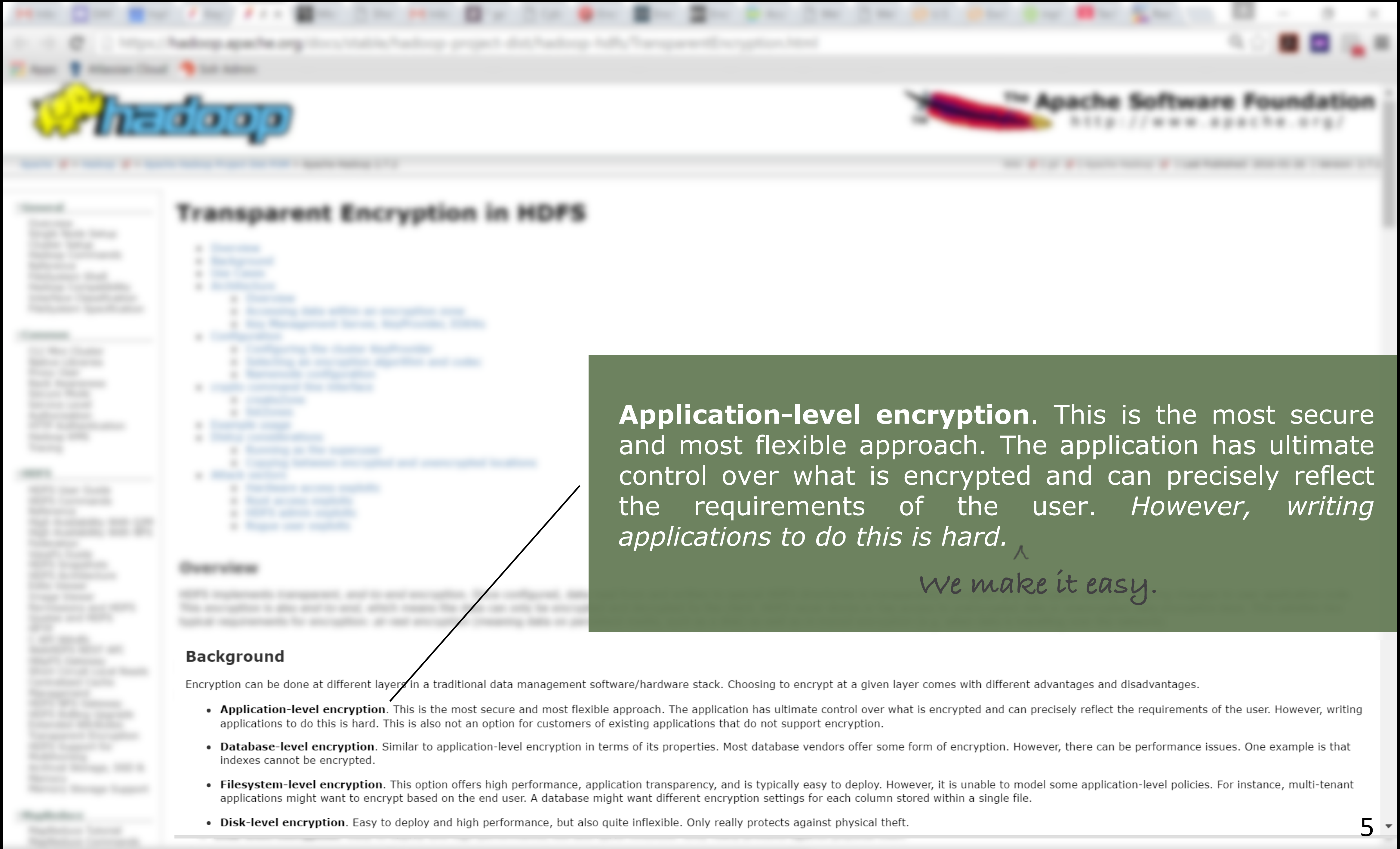
host

application

data

TLS
SSL

https://



The Apache Software Foundation
http://www.apache.org/

# Transparent Encryption in HDFS

## Overview

## Background

Encryption can be done at different layers in a traditional data management software/hardware stack. Choosing to encrypt at a given layer comes with different advantages and disadvantages.

- **Application-level encryption**. This is the most secure and most flexible approach. The application has ultimate control over what is encrypted and can precisely reflect the requirements of the user. However, writing applications to do this is hard. This is also not an option for customers of existing applications that do not support encryption.
- **Database-level encryption**. Similar to application-level encryption in terms of its properties. Most database vendors offer some form of encryption. However, there can be performance issues. One example is that indexes cannot be encrypted.
- **Filesystem-level encryption**. This option offers high performance, application transparency, and is typically easy to deploy. However, it is unable to model some application-level policies. For instance, multi-tenant applications might want to encrypt based on the end user. A database might want different encryption settings for each column stored within a single file.
- **Disk-level encryption**. Easy to deploy and high performance, but also quite inflexible. Only really protects against physical theft.

**Application-level encryption**. This is the most secure and most flexible approach. The application has ultimate control over what is encrypted and can precisely reflect the requirements of the user. *However, writing applications to do this is hard.*

We make it easy.

# SSE - very naïve approach

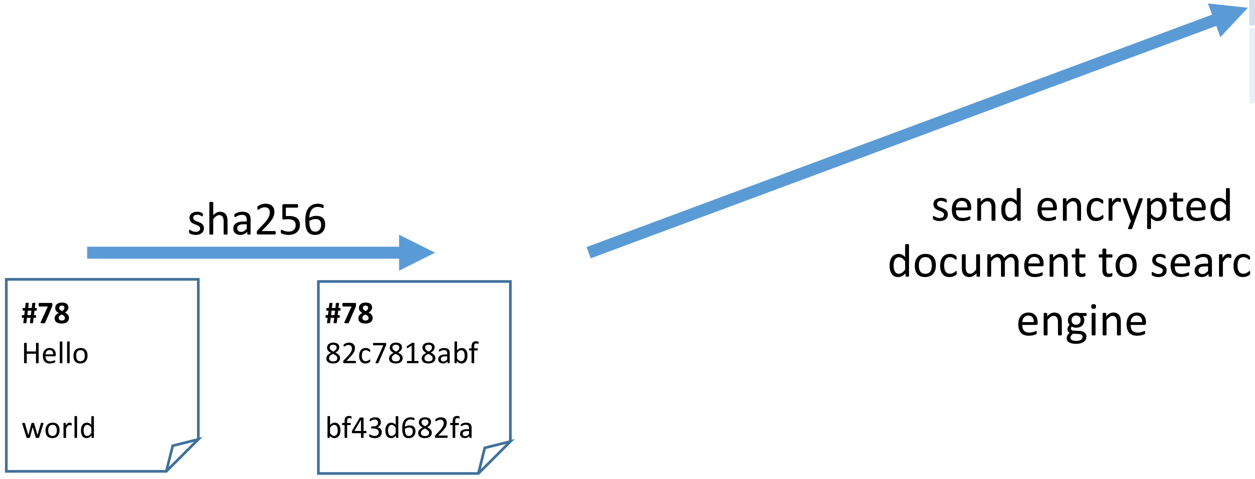
plaintext inverted index

| keyword | documents:position |
|---|---|
| dolomiti | [1:3,25],[4:2],[77:14] |
| trento | [4:4,16,25,67] |
| crypto | [3:2],[5:23] |
| inpher | [1:2,13],[3:54],[5:12] |

sha256

encrypted inverted index

| keyword | documents:position |
|---|---|
| 2d7b45b490 | [1:3,25],[4:2],[77:14] |
| 5d34a4c561 | [4:4,16,25,67] |
| da2f073e06 | [3:2],[5:23] |
| eac41ea006 | [1:2,13],[3:54],[5:12] |
| 82c7818abf | [78:1] |
| bf43d682fa | [78:2] |



**#78**
Hello

world

**#78**
82c7818abf

bf43d682fa

send encrypted document to search engine

# SSE – less naïve approach

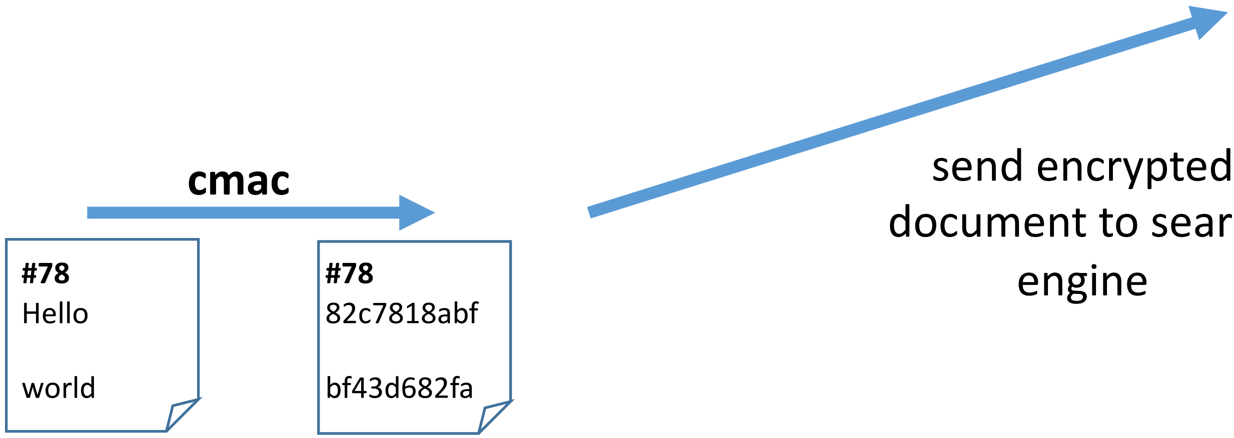
plaintext inverted index

| keyword | documents:position |
|---|---|
| dolomiti | [1:3,25],[4:2],[77:14] |
| trento | [4:4,16,25,67] |
| crypto | [3:2],[5:23] |
| inpher | [1:2,13],[3:54],[5:12] |

**cmac**

encrypted inverted index

| keyword | documents:position |
|---|---|
| 2d7b45b490 | [1:3,25],[4:2],[77:14] |
| 5d34a4c561 | [4:4,16,25,67] |
| da2f073e06 | [3:2],[5:23] |
| eac41ea006 | [1:2,13],[3:54],[5:12] |
| 82c7818abf | [78:1] |
| bf43d682fa | [78:2] |



**#78**
Hello

world

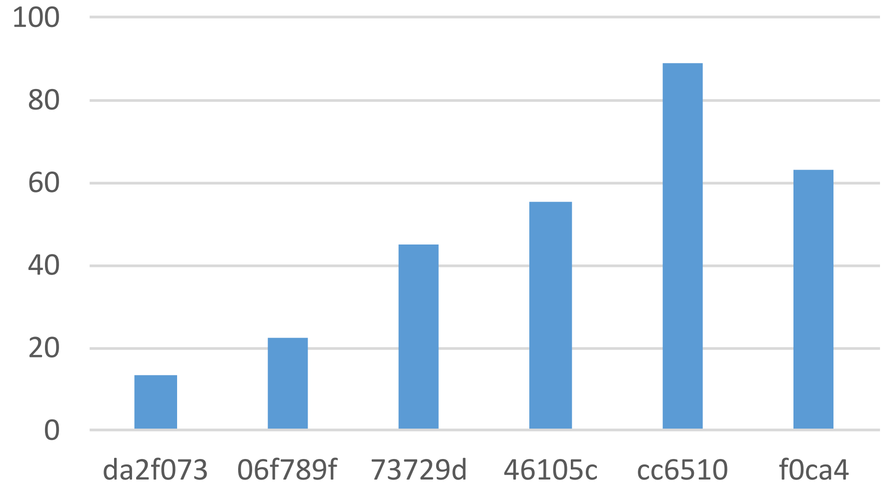
**#78**
82c7818abf

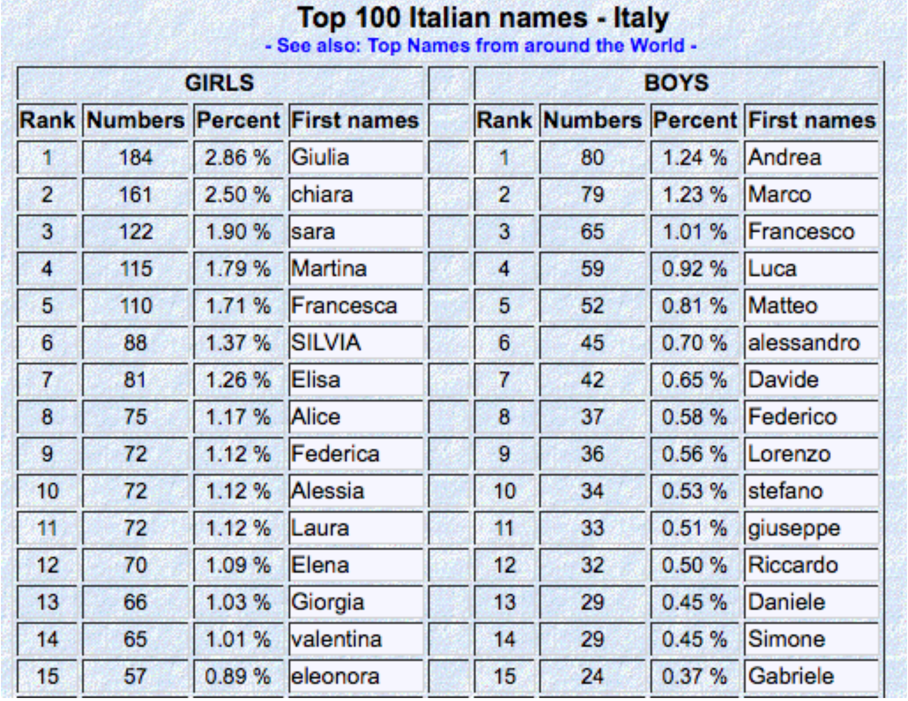
bf43d682fa

send encrypted document to search engine

# SSE – Frequency Attacks

distribution of encrypted index





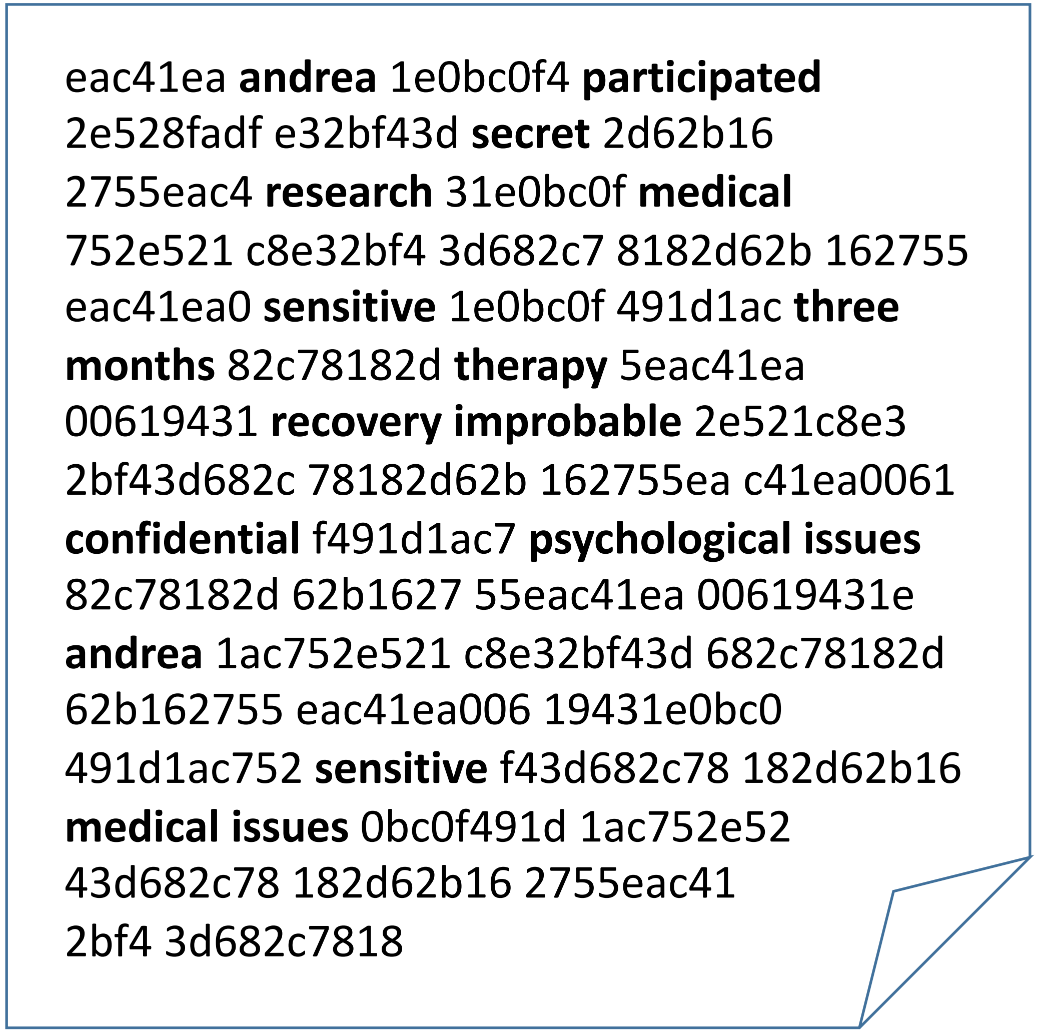
**Top 100 Italian names - Italy**

- See also: Top Names from around the World -

| GIRLS | | | | | BOYS | | | |
|---|---|---|---|---|---|---|---|---|
| Rank | Numbers | Percent | First names | | Rank | Numbers | Percent | First names |
| 1 | 184 | 2.86 % | Giulia | | 1 | 80 | 1.24 % | Andrea |
| 2 | 161 | 2.50 % | chiara | | 2 | 79 | 1.23 % | Marco |
| 3 | 122 | 1.90 % | sara | | 3 | 65 | 1.01 % | Francesco |
| 4 | 115 | 1.79 % | Martina | | 4 | 59 | 0.92 % | Luca |
| 5 | 110 | 1.71 % | Francesca | | 5 | 52 | 0.81 % | Matteo |
| 6 | 88 | 1.37 % | SILVIA | | 6 | 45 | 0.70 % | alessandro |
| 7 | 81 | 1.26 % | Elisa | | 7 | 42 | 0.65 % | Davide |
| 8 | 75 | 1.17 % | Alice | | 8 | 37 | 0.58 % | Federico |
| 9 | 72 | 1.12 % | Federica | | 9 | 36 | 0.56 % | Lorenzo |
| 10 | 72 | 1.12 % | Alessia | | 10 | 34 | 0.53 % | stefano |
| 11 | 72 | 1.12 % | Laura | | 11 | 33 | 0.51 % | giuseppe |
| 12 | 70 | 1.09 % | Elena | | 12 | 32 | 0.50 % | Riccardo |
| 13 | 66 | 1.03 % | Giorgia | | 13 | 29 | 0.45 % | Daniele |
| 14 | 65 | 1.01 % | valentina | | 14 | 29 | 0.45 % | Simone |
| 15 | 57 | 0.89 % | eleonora | | 15 | 24 | 0.37 % | Gabriele |

compare distribution

# SSE – Frequency Attacks



eac41ea **andrea** 1e0bc0f4 **participated**
2e528fadf e32bf43d **secret** 2d62b16
2755eac4 **research** 31e0bc0f **medical**
752e521 c8e32bf4 3d682c7 8182d62b 162755
eac41ea0 **sensitive** 1e0bc0f 491d1ac **three**
**months** 82c78182d **therapy** 5eac41ea
00619431 **recovery improbable** 2e521c8e3
2bf43d682c 78182d62b 162755ea c41ea0061
**confidential** f491d1ac7 **psychological issues**
82c78182d 62b1627 55eac41ea 00619431e
**andrea** 1ac752e521 c8e32bf43d 682c78182d
62b162755 eac41ea006 19431e0bc0
491d1ac752 **sensitive** f43d682c78 182d62b16
**medical issues** 0bc0f491d 1ac752e52
43d682c78 182d62b16 2755eac41
2bf4 3d682c7818

# SSE – encrypt rows

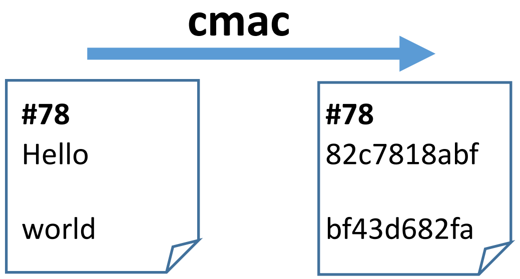
plaintext inverted index

| keyword | documents:position |
|---|---|
| dolomiti | [1:3,25],[4:2],[77:14] |
| trento | [4:4,16,25,67] |
| crypto | [3:2],[5:23] |
| inpher | [1:2,13],[3:54],[5:12] |

**cmac**

encrypted inverted index
per row encryption key

| keyword | enc$_{Ki}$(documents:position) |
|---|---|
| 2d7b45b490 | 🟦 |
| 5d34a4c561 | 🟧 |
| da2f073e06 | 🟨 |
| eac41ea006 | 🟩 |
| 82c7818abf | 🟥 |
| bf43d682fa | 🟪 |



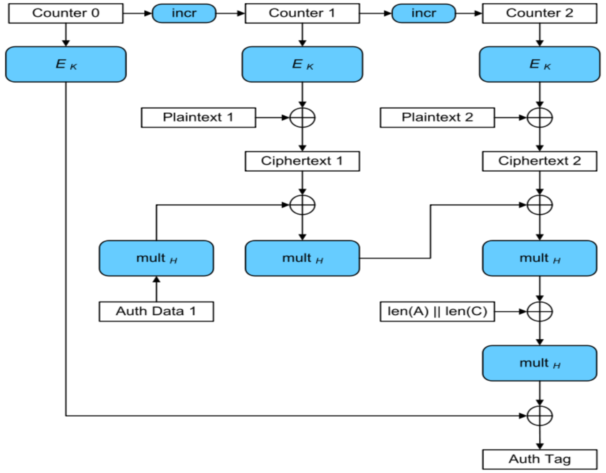
send encrypted
document to search
engine

# Inpher Encrypt Module

| VALUE | CIPHERTEXT |
|---|---|
| 99.21 | 0x9231231bfh23131 |
| 100.56 | 0x1132bbbfdh45243 |
| 101.78 | 0xbb2313fg1233700 |
| 110.34 | 0xccbaa3431325321 |

**ORE.compare(Enc(x), Enc(y)) < 0 => x < y**

**Order Revealing Encryption**

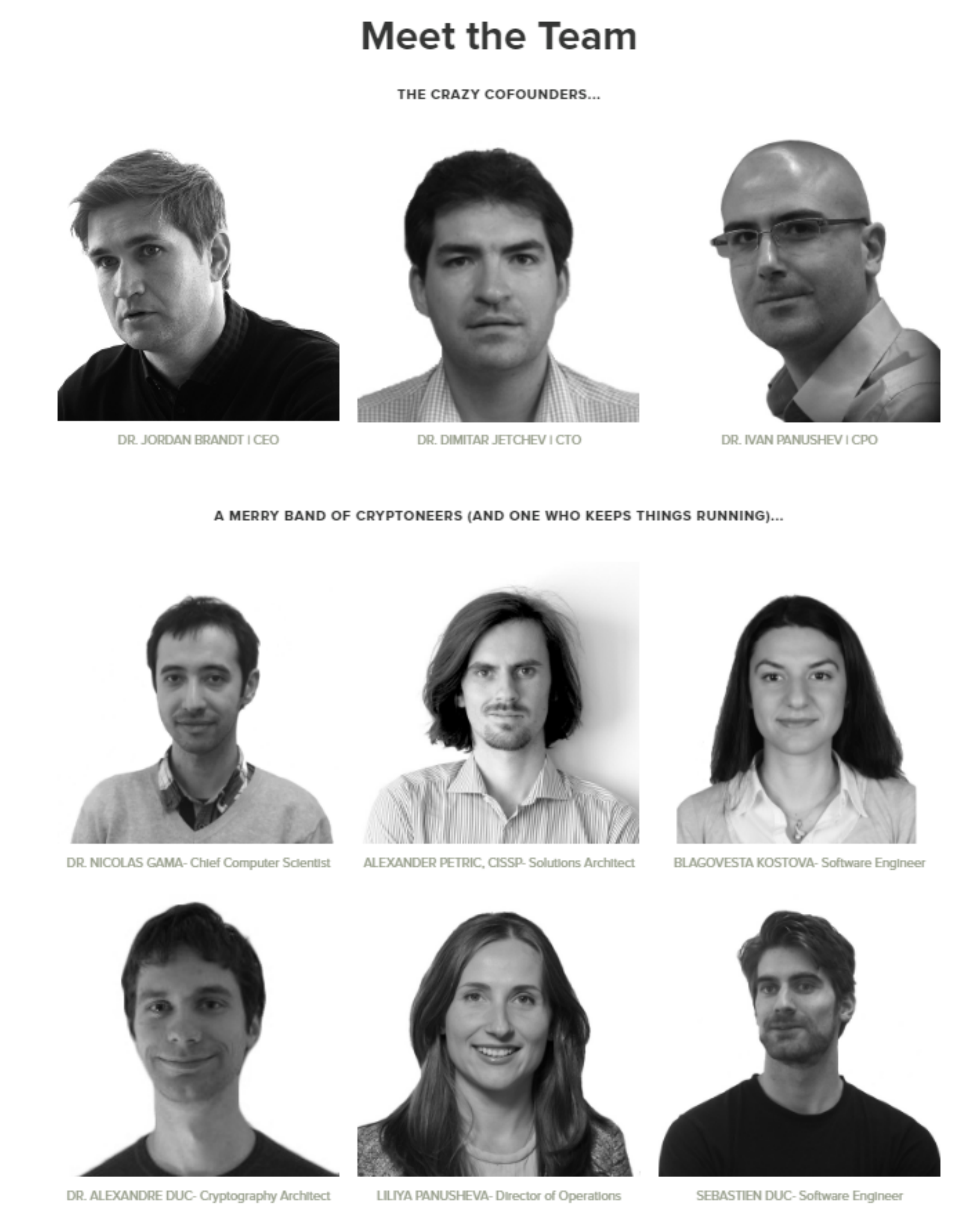


**Authenticated Symmetric Encryption**

**(AES-GCM)**

inpher.io

Inpher is a team of veteran founders, cryptographers and software engineers who believe that encryption is foundational to the future of computing.



## Meet the Team

THE CRAZY COFOUNDERS...

DR. JORDAN BRANDT I CEO

DR. DIMITAR JETCHEV I CTO

DR. IVAN PANUSHEV I CPO

A MERRY BAND OF CRYPTONEERS (AND ONE WHO KEEPS THINGS RUNNING)...

DR. NICOLAS GAMA- Chief Computer Scientist

ALEXANDER PETRIC, CISSP- Solutions Architect

BLAGOVESTA KOSTOVA- Software Engineer

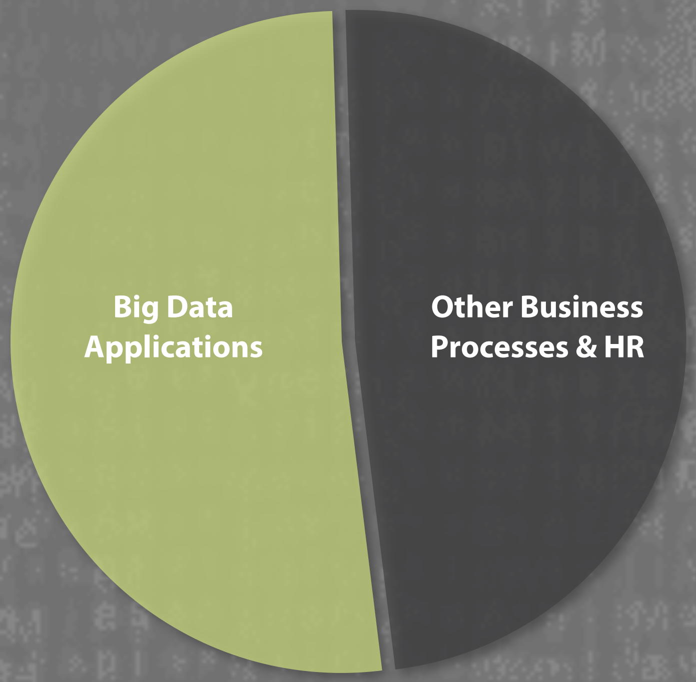
DR. ALEXANDRE DUC- Cryptography Architect

LILIYA PANUSHEVA- Director of Operations

SEBASTIEN DUC- Software Engineer

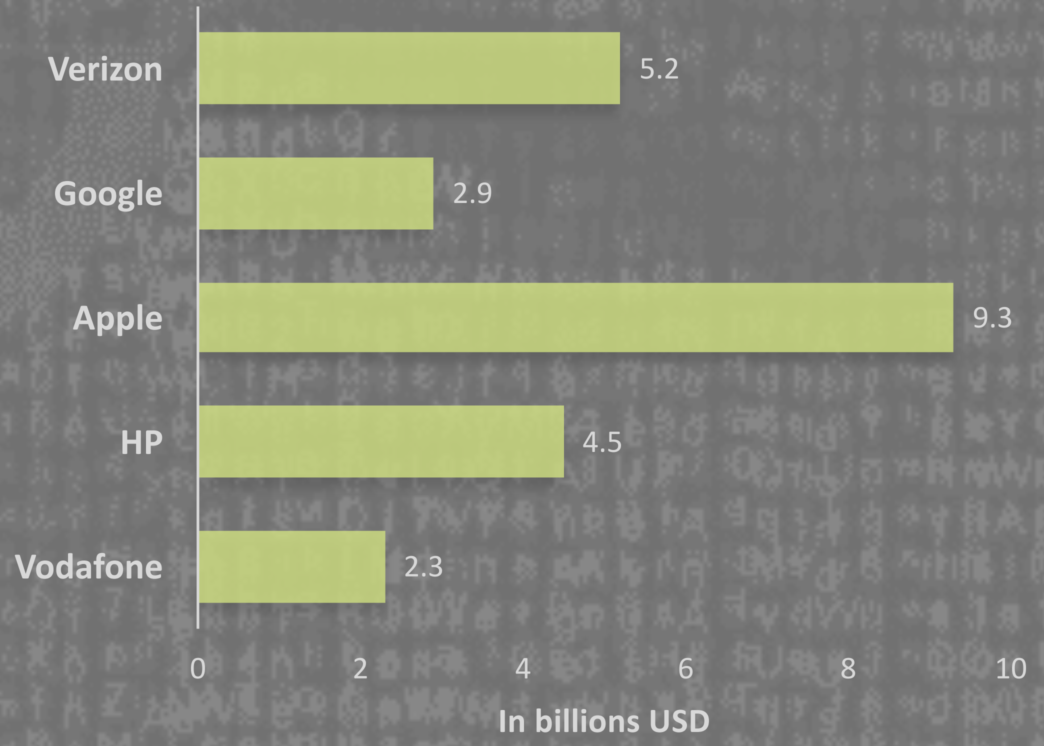
# The price of not securing data before using the cloud.

## Source of Ethics and Privacy Violations by 2018



Gartner- "Seven Best Practices for Your Big Data Analytics Projects", Oct. 2015

## Fees for failure to comply with General Data Protection Regulation

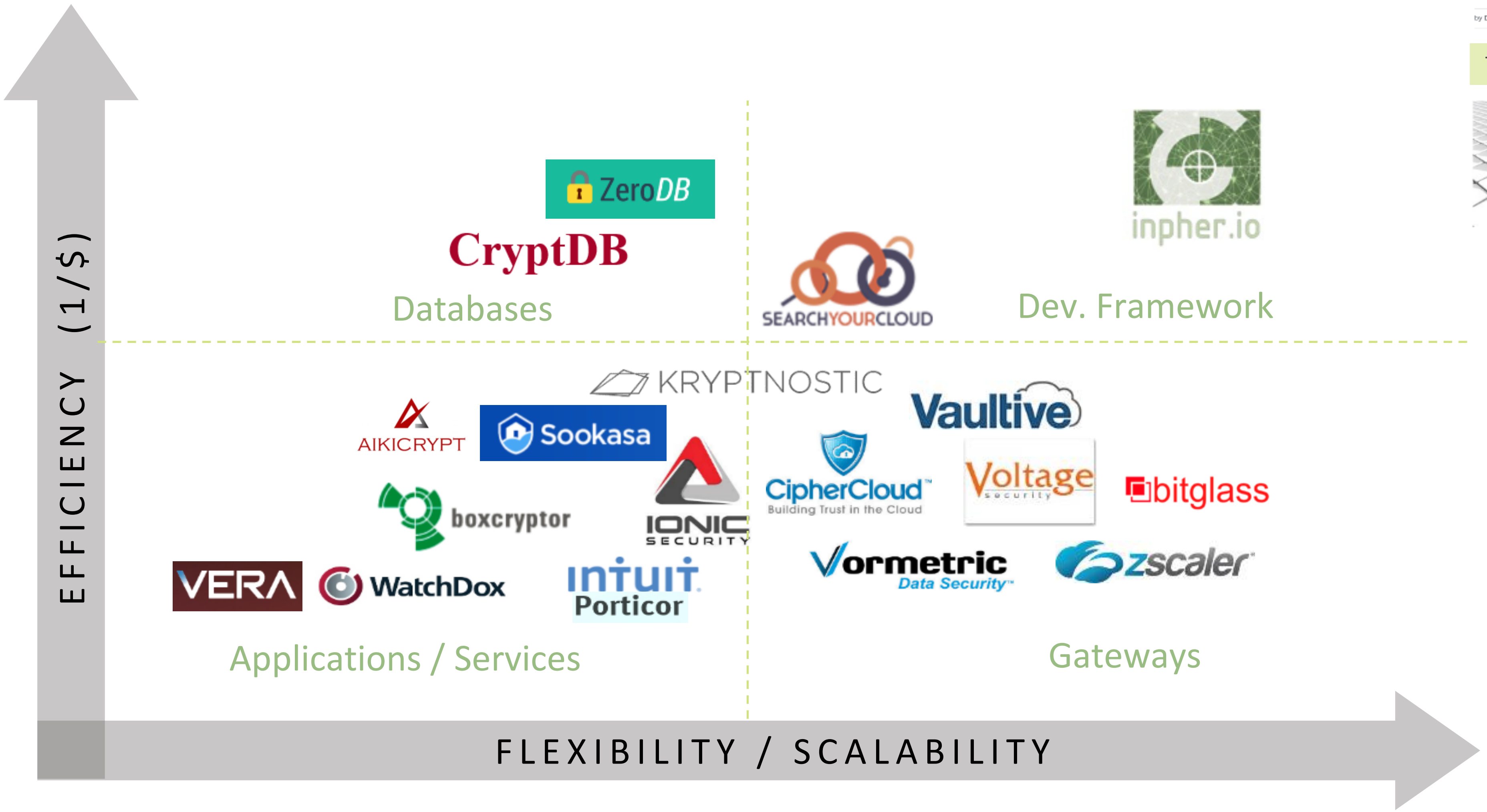


# So what is impeding adoption?

+ Complexity: application-level encryption is hard*
+ Functionality: preserving search and collaboration

* From Hadoop Wiki https://hadoop.apache.org/docs/hadoop-hdfs/TransparentEncryption.html

# Competitive Landscape



EFFICIENCY (1/$)

FLEXIBILITY / SCALABILITY

Databases

Dev. Framework

Applications / Services

Gateways

**Researchers poke hole in custom crypto built for Amazon Web Services**

Even when engineers do everything by the book, secure crypto is still hard.

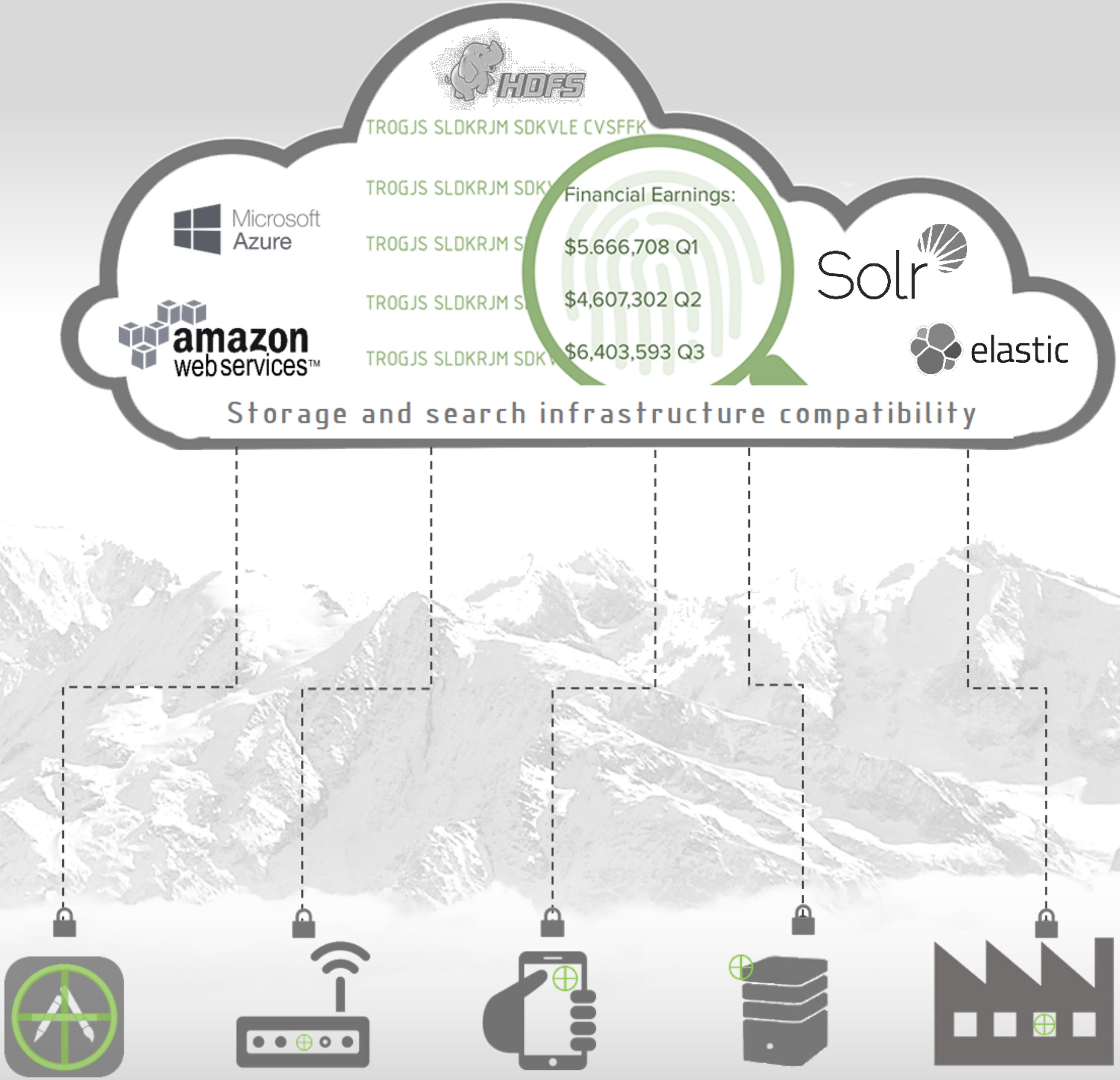
by Dan Goodin - Nov 21, 2015 8:40am PST

Two reasons why Inpher uses standard crypto.

The Dark Side of Cryptography: Kleptography in Black-Box Implementations

infosecurity-magazine.com

# inpher.io



TROGJS SLDKRJM SDKVLE CVSFFK

TROGJS SLDKRJM SDK

TROGJS SLDKRJM S

TROGJS SLDKRJM S

TROGJS SLDKRJM SDK

Financial Earnings:

$5.666,708 Q1

$4,607,302 Q2

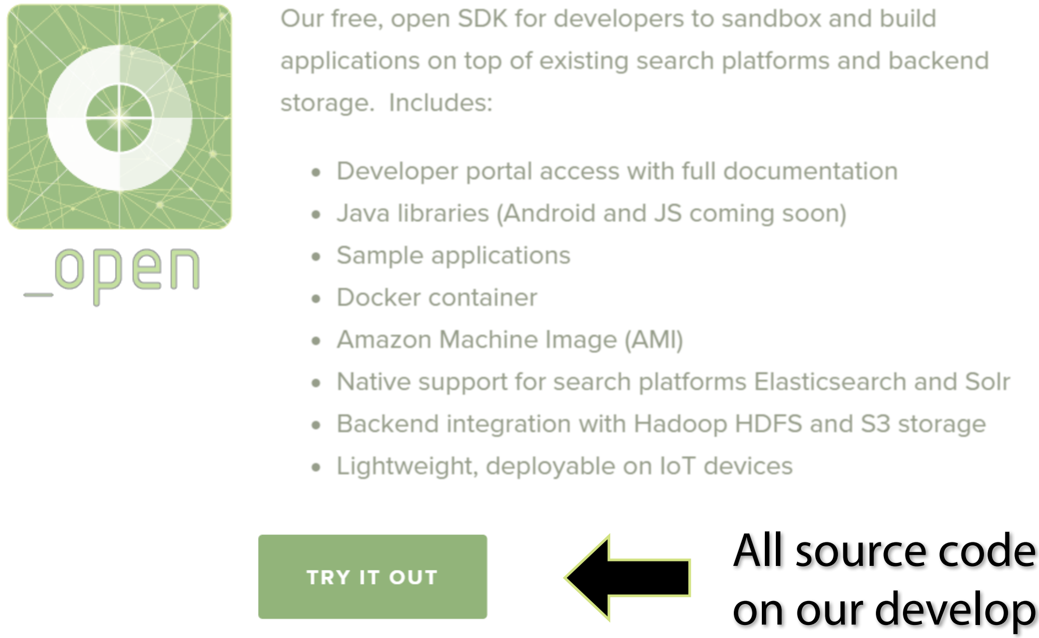
$6,403,593 Q3

Storage and search infrastructure compatibility

## Client-side SDK:

# Search and Share Ciphertext Like Plaintext.

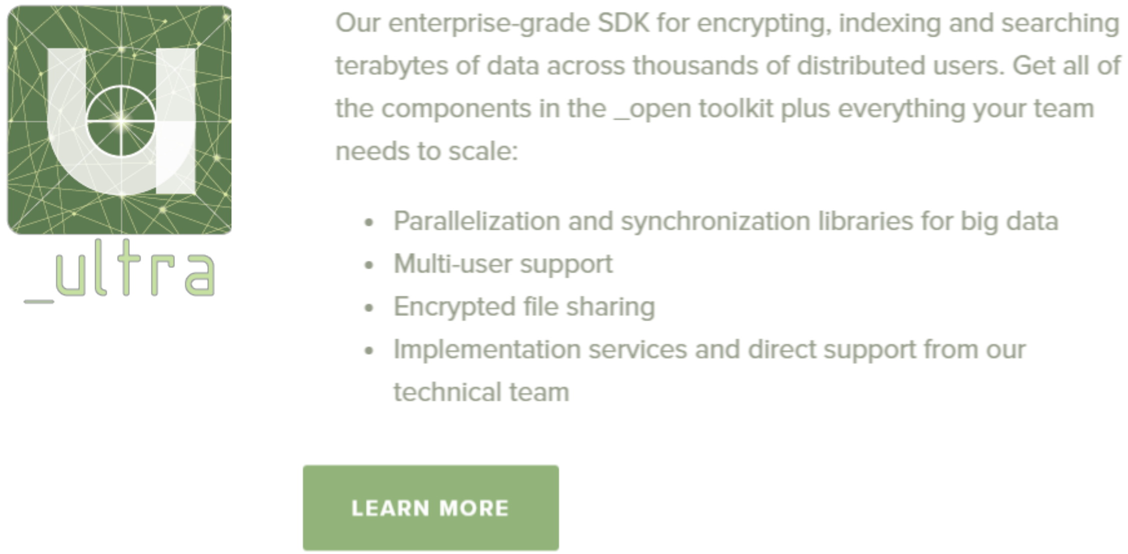
- Empowers developers to *quickly* create secure applications without being crypto experts
- Applications can search and share encrypted data without decryption on existing infrastructure



Our free, open SDK for developers to sandbox and build applications on top of existing search platforms and backend storage. Includes:

- Developer portal access with full documentation
- Java libraries (Android and JS coming soon)
- Sample applications
- Docker container
- Amazon Machine Image (AMI)
- Native support for search platforms Elasticsearch and Solr
- Backend integration with Hadoop HDFS and S3 storage
- Lightweight, deployable on IoT devices

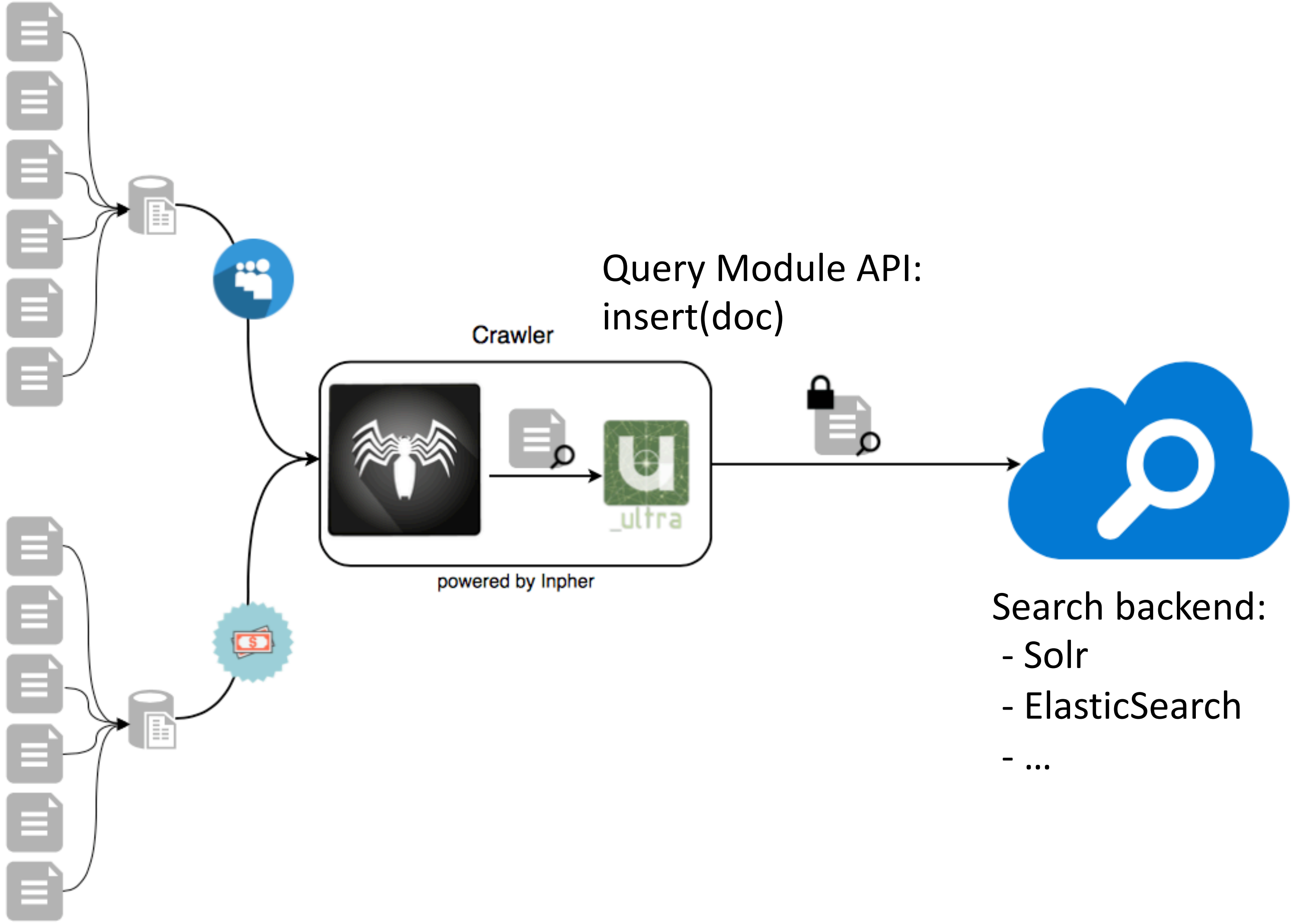
All source code published on our developer portal !



Our enterprise-grade SDK for encrypting, indexing and searching terabytes of data across thousands of distributed users. Get all of the components in the _open toolkit plus everything your team needs to scale:

- Parallelization and synchronization libraries for big data
- Multi-user support
- Encrypted file sharing
- Implementation services and direct support from our technical team

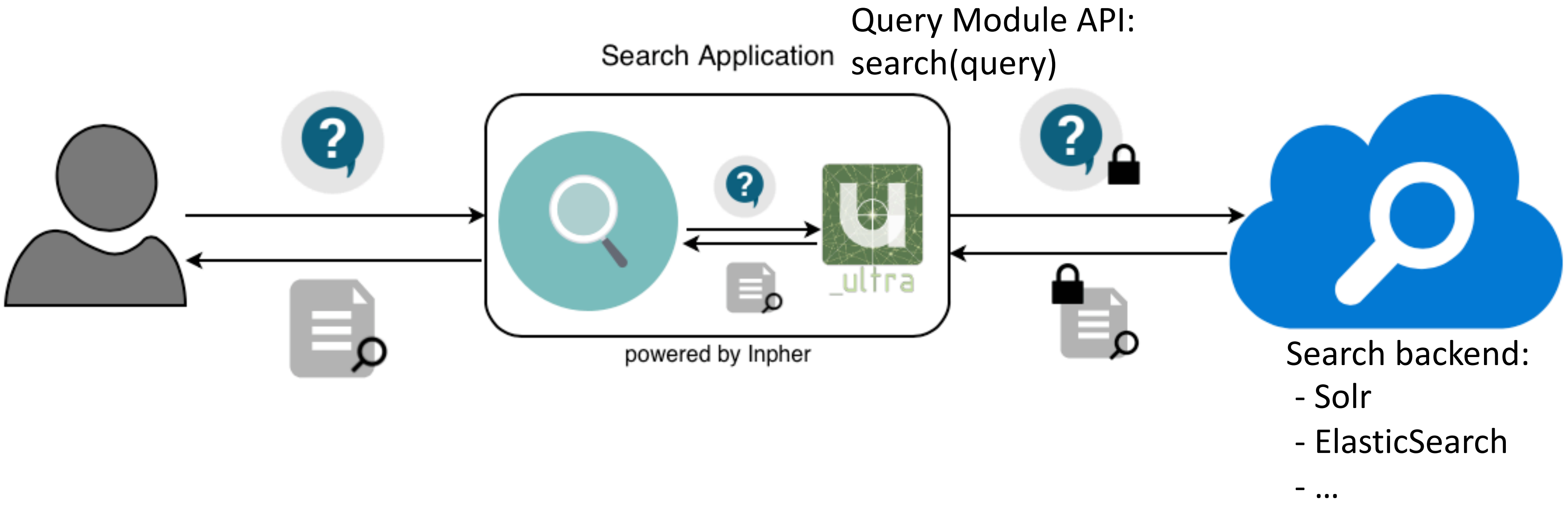
# Inpher Query Module Indexing



Query Module API:
insert(doc)

Search backend:
- Solr
- ElasticSearch
- …

# Inpher Query Module Search

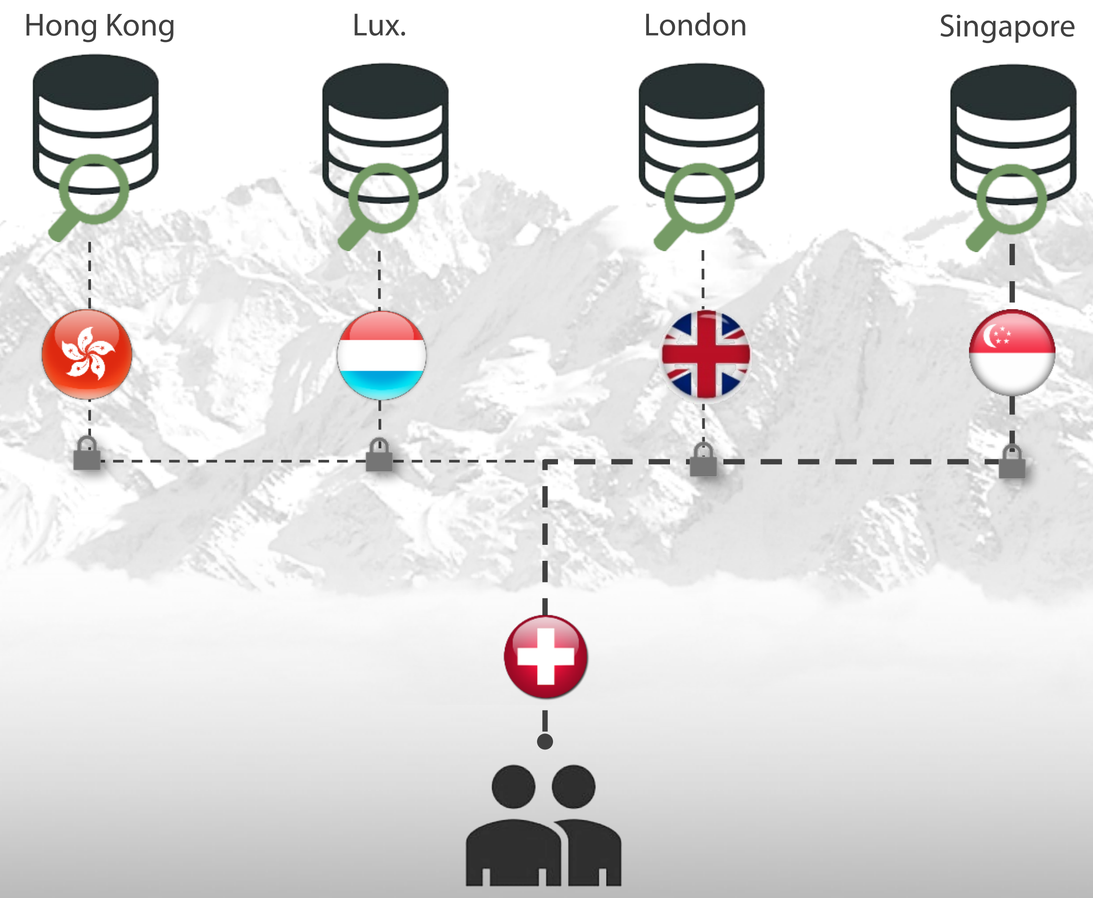


# Searchable Encryption Other approaches, future development

- Bloom Filters
- ORAM
- Multilinear Maps
- ORE / OPE
- (FHE)
- …

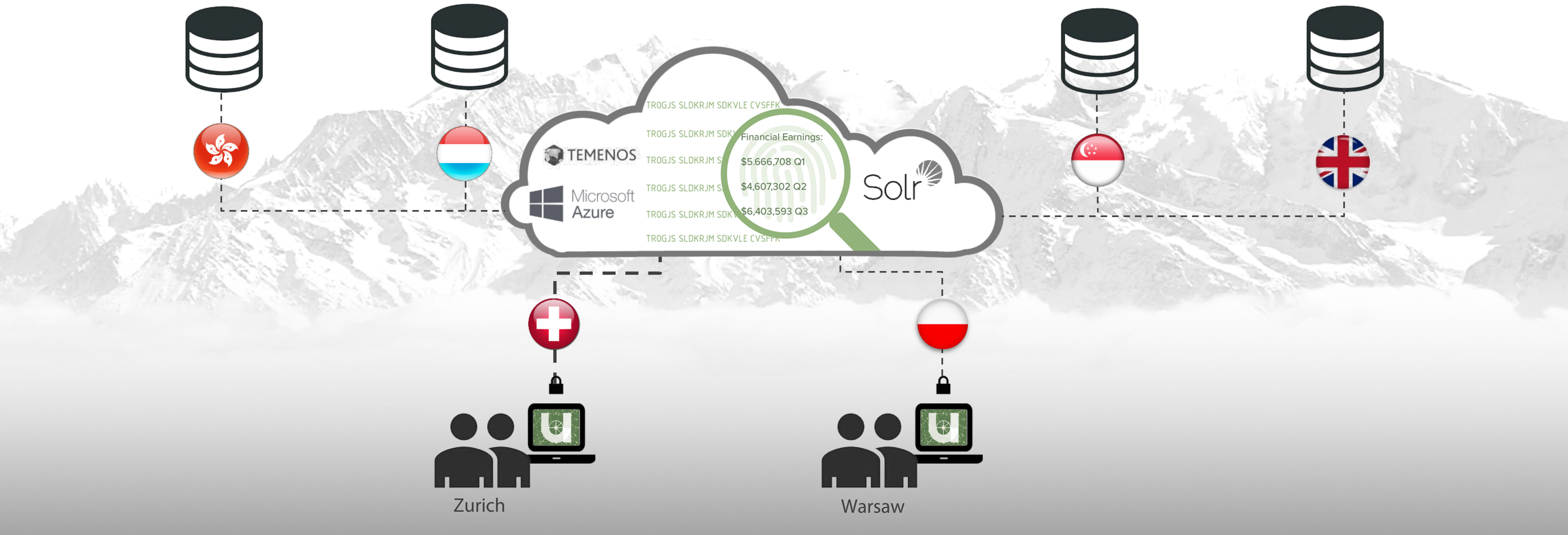
# PoC Challenge: Centralized Data Discovery

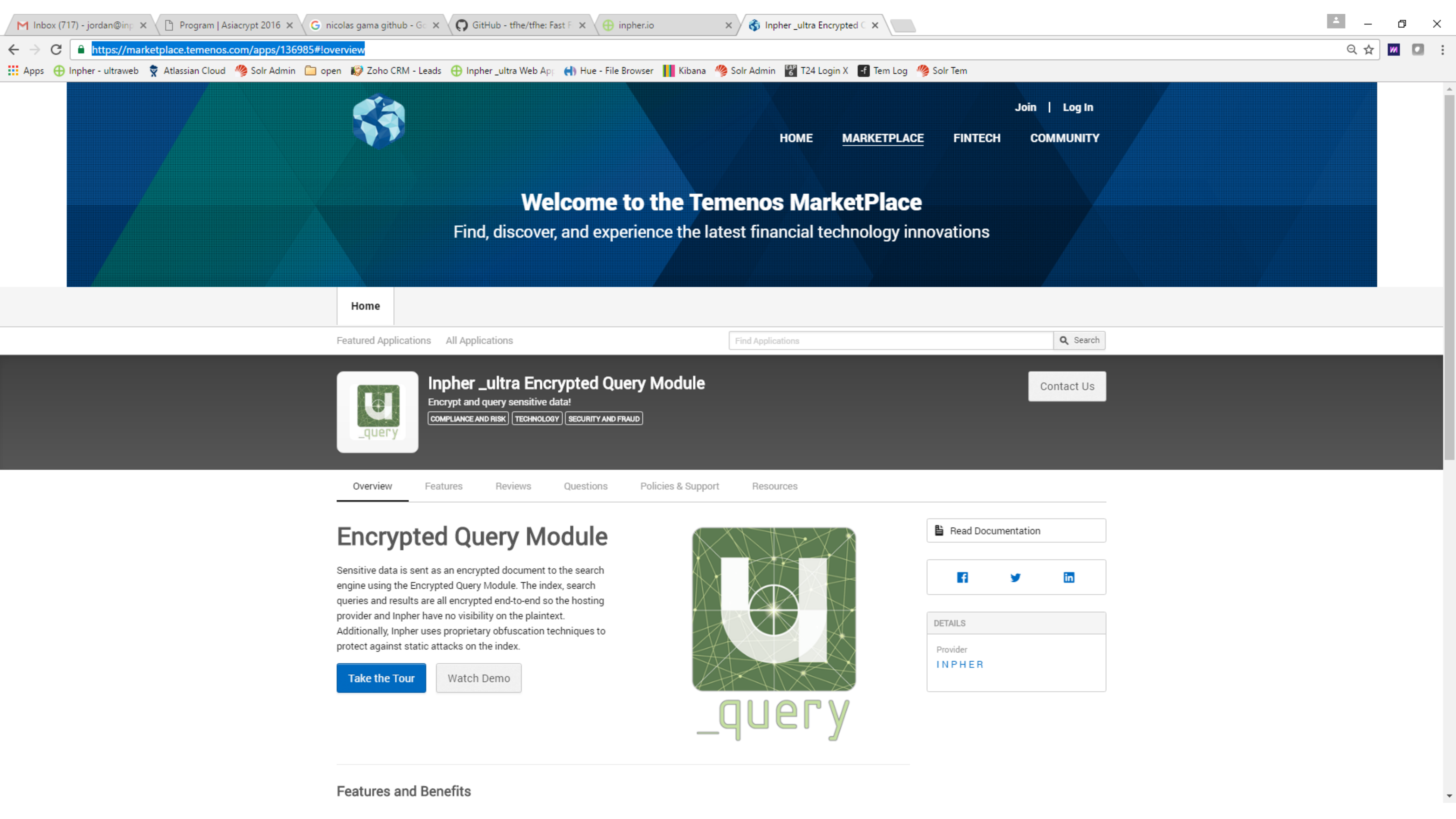
- Challenged with reconciling separate instances of customer data at each international location.
- Desire to centralize data in cloud, but hindered by security and cross-border data regulations.



# Solution: _ultra SDK integrated with Temenos/Azure cloud

- Deploying in cloud with core banking software Temenos using *_ultra* SDK for master data search.
- End-to-end encryption eases compliance with sovereign data regulations (zero-knowledge cloud).
- Can outsource more banking operations (account reconciliation, etc.) to lower-cost locations.



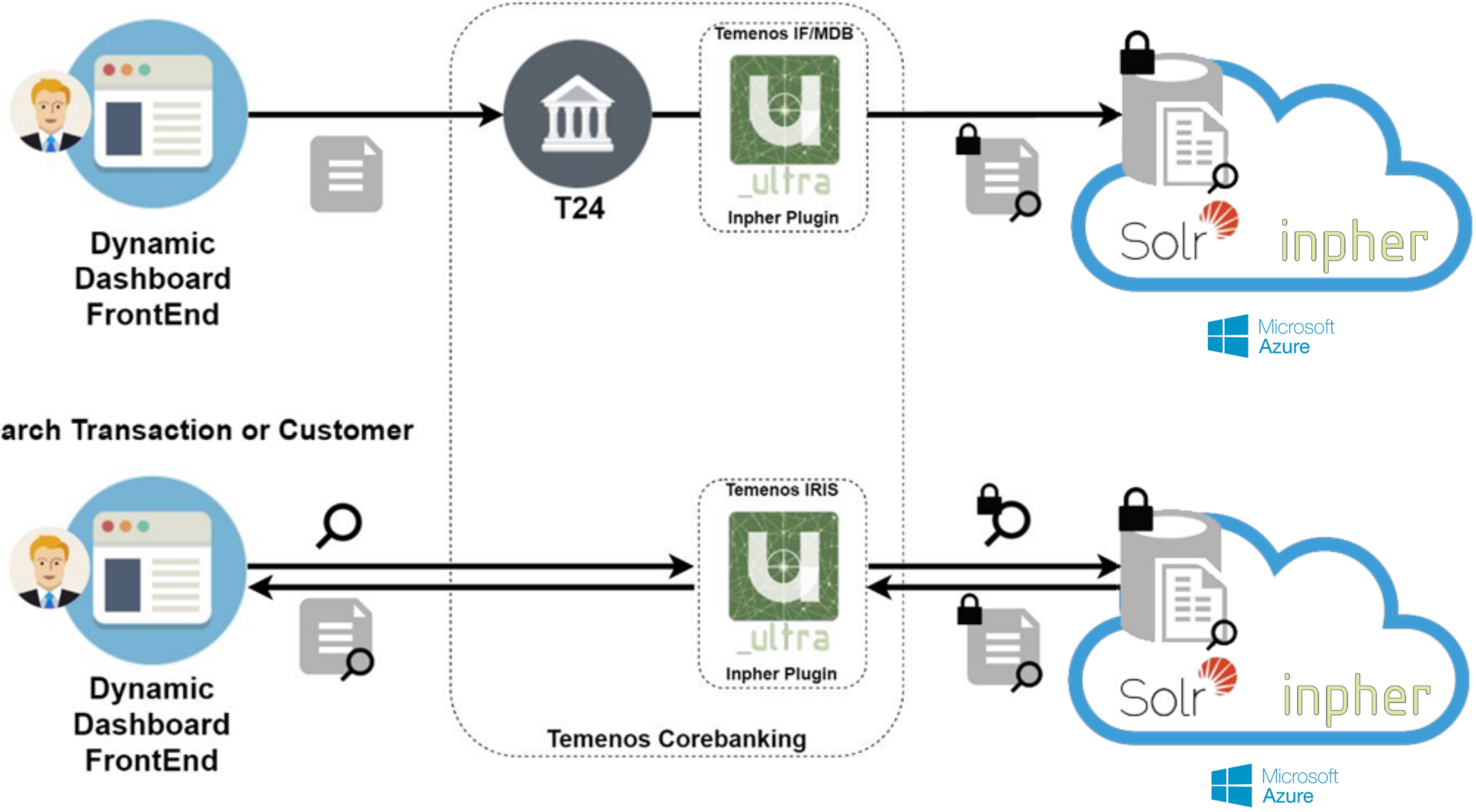


# Encrypted Query Module

Sensitive data is sent as an encrypted document to the search engine using the Encrypted Query Module. The index, search queries and results are all encrypted end-to-end so the hosting provider and Inpher have no visibility on the plaintext. Additionally, Inpher uses proprietary obfuscation techniques to protect against static attacks on the index.

## Features and Benefits

Create or Modify Transaction, Customer and others



Dynamic Dashboard FrontEnd

T24

Temenos IF/MDB

_ultra

Inpher Plugin

Solr

inpher

Microsoft Azure

Search Transaction or Customer

Dynamic Dashboard FrontEnd

Temenos IRIS

_ultra

Inpher Plugin

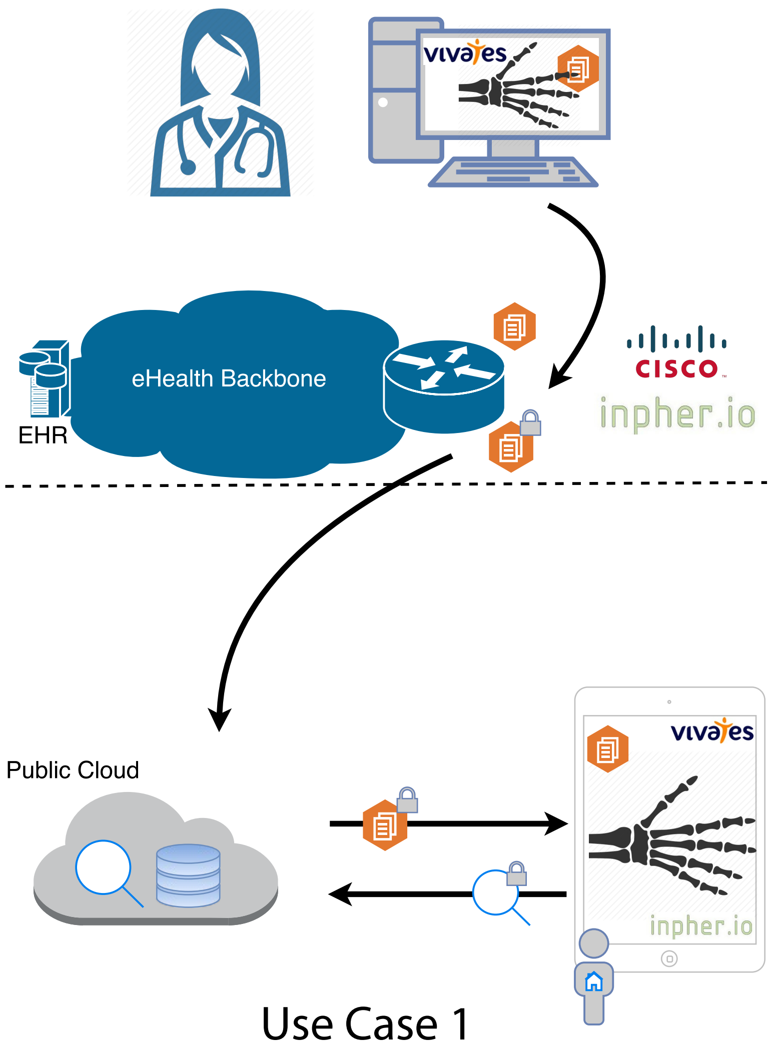
Temenos Corebanking

Solr

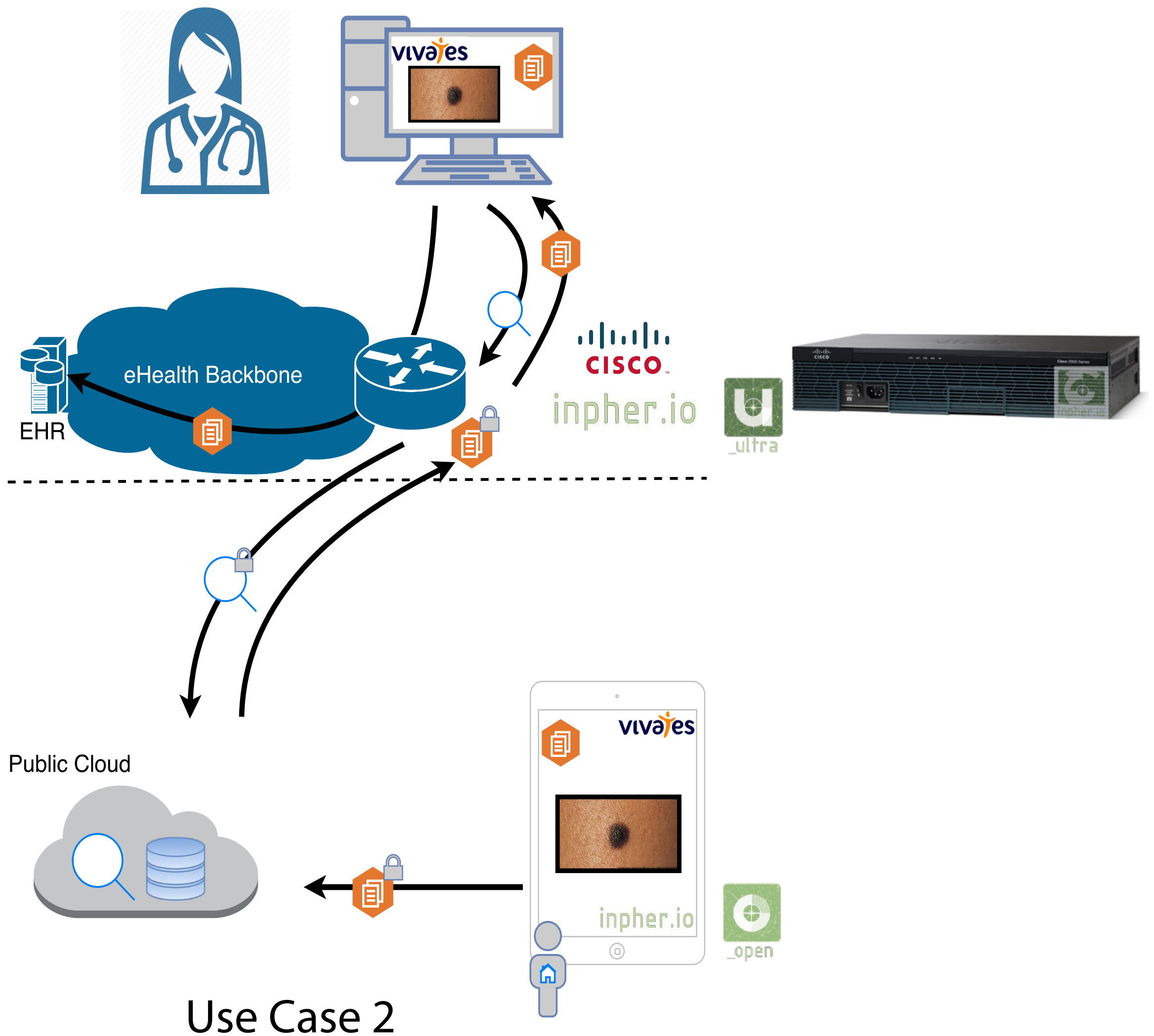
inpher

Microsoft Azure

# Cisco Medical Data Exchange

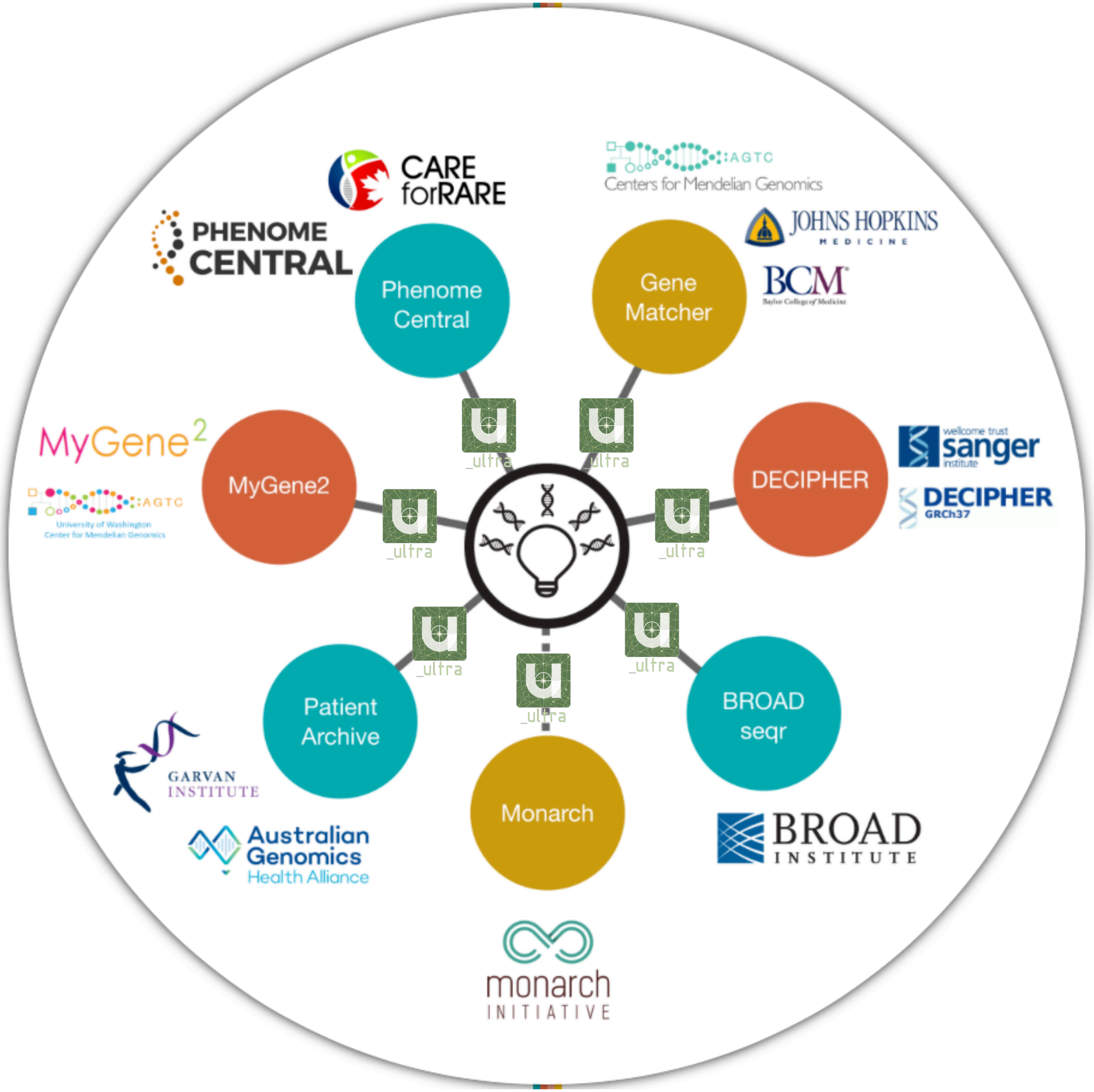


Use Case 1



Use Case 2

# Genetic Data Exchange



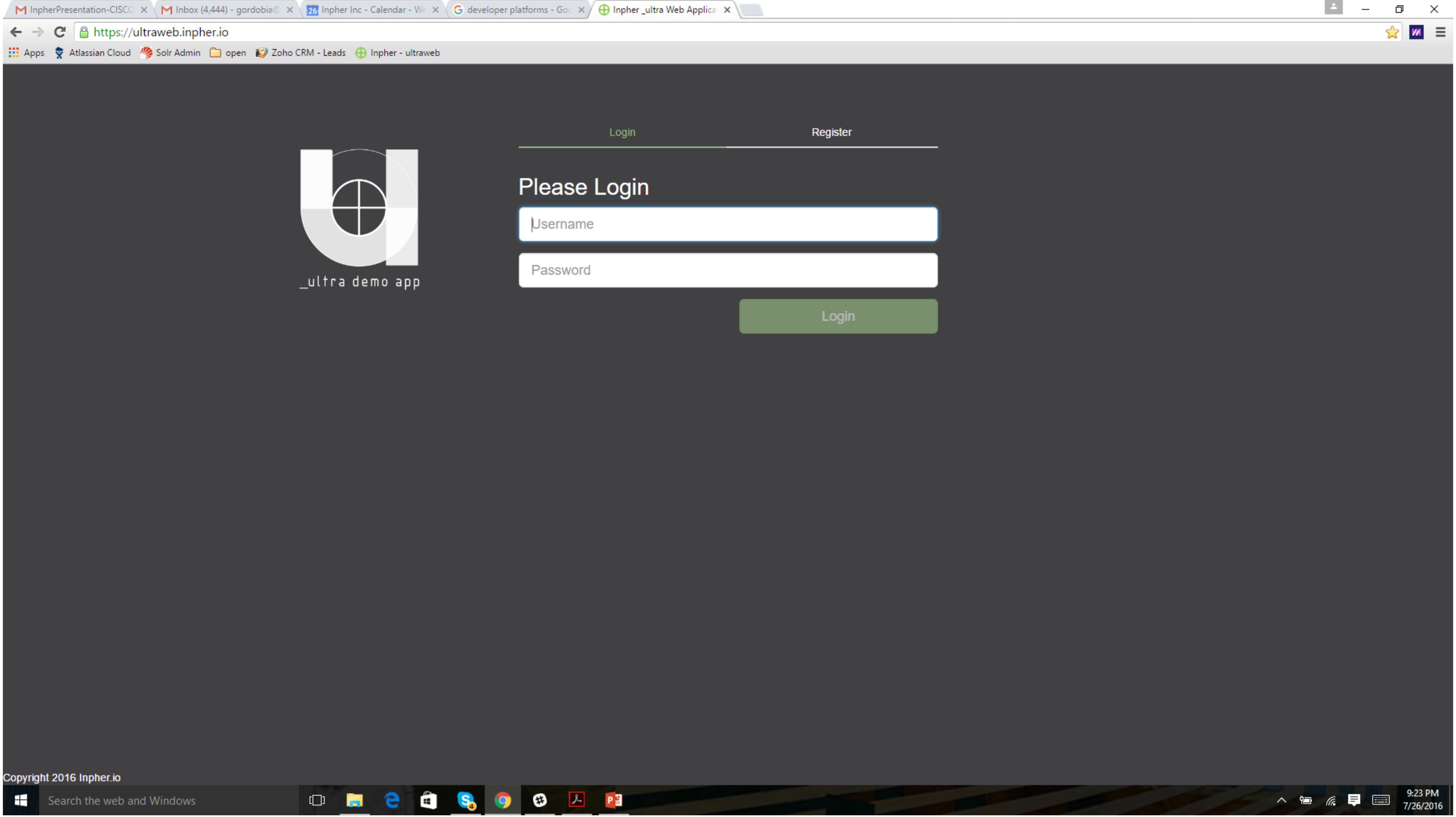
# Beyond Search

Putting the pieces together:

- Encrypted File System
- Encrypted Search Engine
- Group Key Management Sharing

Analytics:

Nicolas Paper



InpherPresentation-CISCO
Inbox (4,444) - gordobia@
Inpher Inc - Calendar - W
developer platforms - Goo
Inpher _ultra Web Applica
https://ultraweb.inpher.io
Apps
Atlassian Cloud
Solr Admin
open
Zoho CRM - Leads
Inpher - ultraweb
Login
Register
_ultra demo app

## Please Login

Copyright 2016 Inpher.io

# _*ultra* SDK Modules

## Encrypt.

Authenticated Randomized Encryption

Order Revealing Encryption (ORE)

Deterministic Encryption

## Collaborate.

Key Management

Secure Data Sharing

Encrypted Data Storage

## Query.

Parsing, Indexing and Encryption

Search

# THANK YOU!

Inpher Demo available at: https://www.youtube.com/watch?v=rSSoidc8XCM