Siren

CATALOGUE

Siren Platform™ Version 13

Product overview

Deloitte. 50 Technology Fast 50 2022 ranked company

Gartner COOL VENDOR 2020

TABLE OF CONTENTS

1 Search-based investigative intelligence

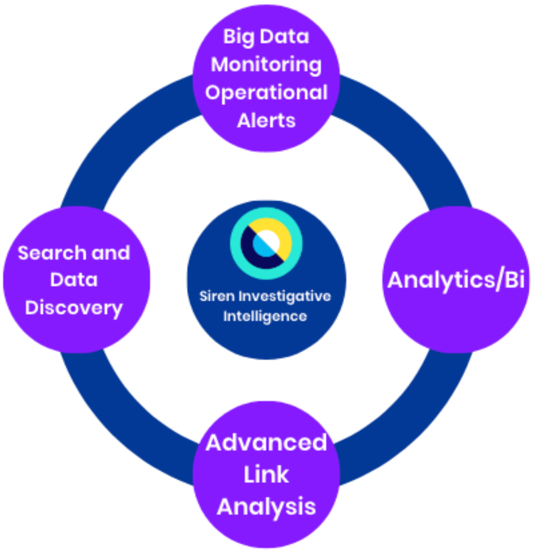
Investigative intelligence is a specialized area of data analytics that serves the needs of those hunting for bad actors to protect people, networks, and assets.

Such investigations are the primary focus of law enforcement and intelligence agencies but are also critical to uncovering financial crime activities and threat hunting in cybersecurity.

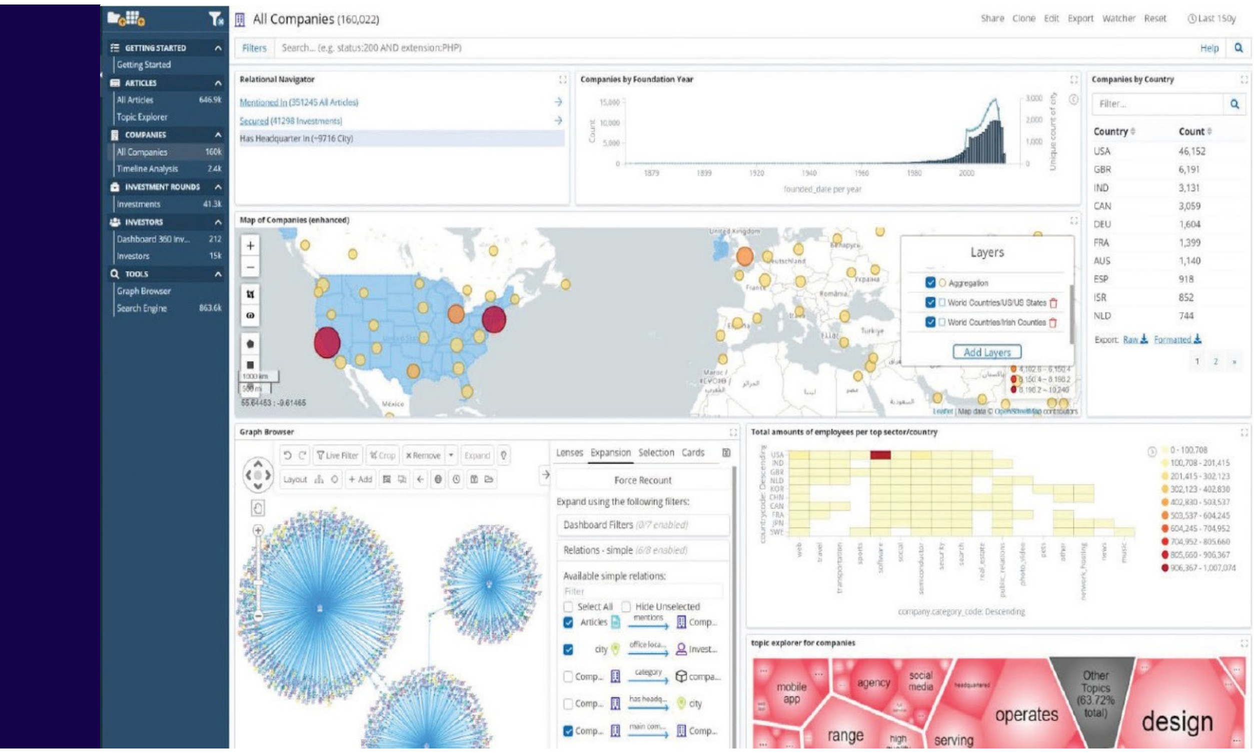
These investigations often involve connecting the dots on both structured (well-defined records) and unstructured data (textual and other media) across systems and schemas.

The Siren Platform fuses previously disconnected capabilities, such as advanced, big data search, link analysis, associative business intelligence dashboards, and big data stream monitoring and alerting capabilities.

Analysts are provided with investigation-grade, active dashboards that serve pre-built use cases or act as a starting point for the exploration of data.


Big Data Monitoring Operational Alerts
Search and Data Discovery
Siren Investigative Intelligence
Analytics/Bi
Advanced Link Analysis

Siren Platform's fusion of investigative capabilities.



All Companies (160,022)

Search... (e.g. status:200 AND extension:PHP)

Mentioned In (351245 All Articles)

Secured (41298 Investments)

Has Headquarter In (~9716 City)

Companies by Country

Country	Count
USA	46,152
GBR	6,191
IND	3,131
CAN	3,059
DEU	1,604
FRA	1,399
AUS	1,140
ESP	918
ISR	852
NLD	744

Export: Raw Formatted

Blended capabilities: Textual discovery - Link analysis - Geo spatial analysis - Time series - Associative navigation

Data model driven.

The investigative world is made up of disjointed data that needs to be connected. People, for example, are connected to vehicles they own, which are connected to locations where they have been, which may be connected to events, and so on.

In Siren Investigate, administrators or advanced analysts define an **associative data model** to describe how records in data tables are meant to be interconnected to form a **knowledge graph.**

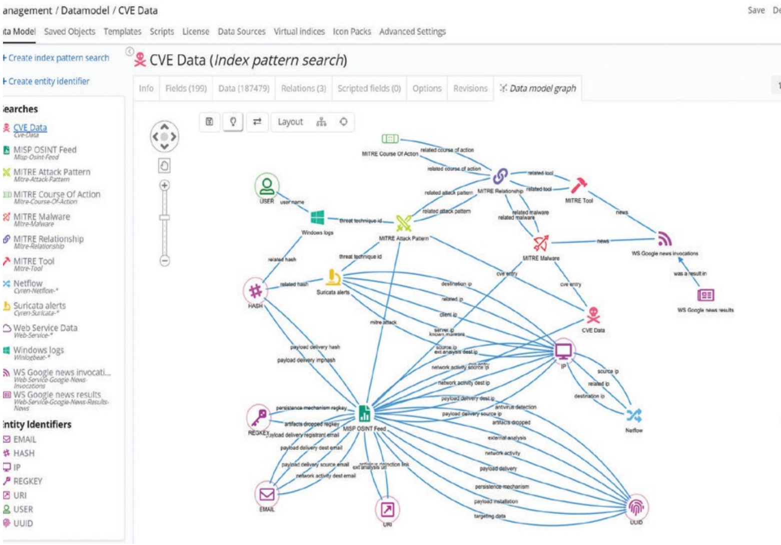
This model then drives all of the analytical operations.

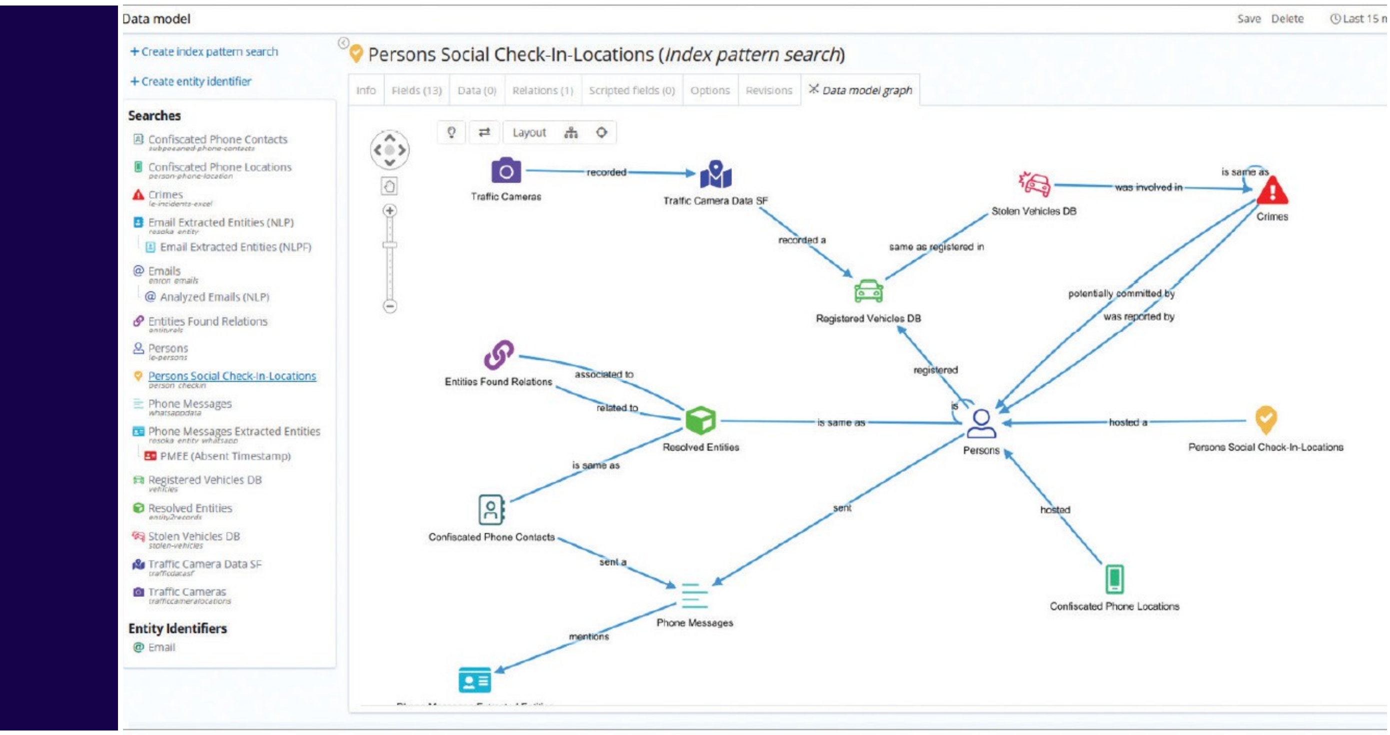
Data model example

The following screenshot shows an example data model for a Law Enforcement scenario. It connects existing tables of vehicles, cases, and traffic camera license plate readings.

The model can be automatically discovered or created and refined manually.

In a **Cybersecurity** scenario, it is common to use concepts such as IPs, MD5 hash values, or user IDs to tie together security logs.





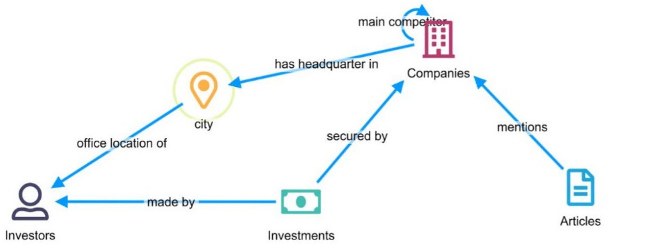
Associative drill downs: A unique investigative superpower

Thanks to the data model, Siren Investigate's analytic dashboards go beyond classic BI drill downs and gain unique associative capabilities.

For example, in a financial investigation scenario, we have data on companies that have received investments
by investors but are also mentioned by articles and have headquarters in cities. This is represented by the data model to the right.

With this data model, Siren Investigate can move from a set of companies to the set of records connected to it - for example, the investments received by those companies. This operation is called set-to-set navigation or associative drill down.

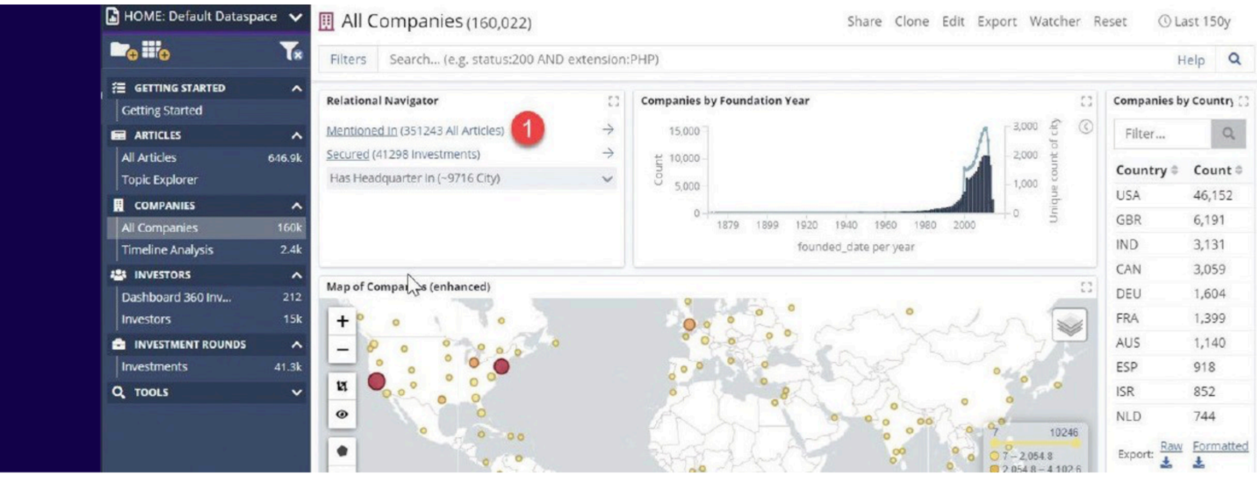
You can create unlimited sub-searches, which inherit the filter of the parent search, comparable to Extended Entity Relation systems. Relationships can be created between sub-searches.



The data model graph in a financial investigation

» Siren Platform operates at a **billion record+ scale** and in real time on data loaded in Elasticsearch.

» **Unique offering**: No other product can perform similar correlations in Elasticsearch – powered by the Siren Federate plugin.



All Companies (160,022)

Search... (e.g. status:200 AND extension:PHP)

Relational Navigator

Mentioned In (351243 All Articles)

Secured (41298 Investments)

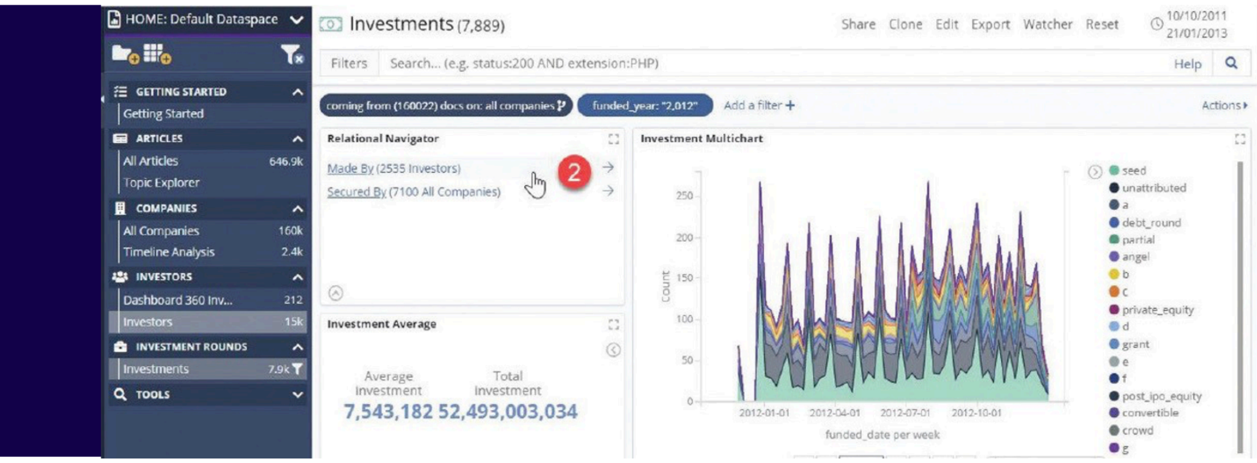
Has Headquarter In (~9716 City)

Companies by Country

Filter...

Country	Count
USA	46,152
GBR	6,191
IND	3,131
CAN	3,059
DEU	1,604
FRA	1,399
AUS	1,140
ESP	918
ISR	852
NLD	744

Export: Raw Formatted



Investments (7,889)

10/10/2011 21/01/2013

Search... (e.g. status:200 AND extension:PHP)

coming from (160022) docs on: all companies

funded_year: "2,012"

Relational Navigator

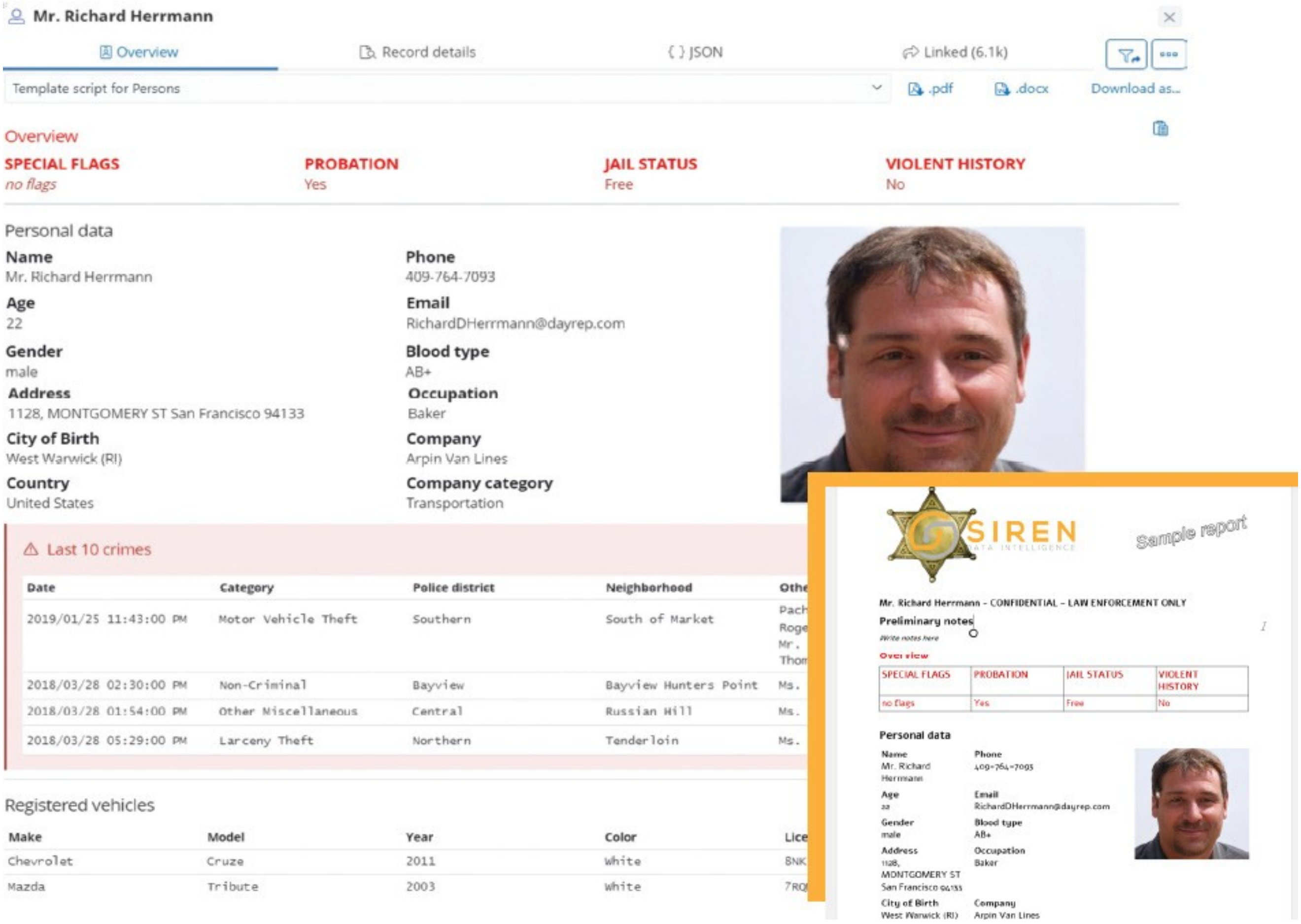
Made By (2535 Investors)

Secured By (7100 All Companies)

Investment Average

Average Investment	Total Investment
7,543,182	52,493,003,034

Search and get full reports in a single click



Mr. Richard Herrmann

Overview

SPECIAL FLAGS	PROBATION	JAIL STATUS	VIOLENT HISTORY
no flags	Yes	Free	No

Personal data

Name
Mr. Richard Herrmann

Age
22

Gender
male

Address
1128, MONTGOMERY ST San Francisco 94133

City of Birth
West Warwick (RI)

Country
United States

Phone
409-764-7093

Email
RichardDHerrmann@dayrep.com

Blood type
AB+

Occupation
Baker

Company
Arpin Van Lines

Company category
Transportation

Last 10 crimes

Date	Category	Police district	Neighborhood	Othe
2019/01/25 11:43:00 PM	Motor Vehicle Theft	Southern	South of Market	Pach Roge Mr. Thom
2018/03/28 02:30:00 PM	Non-Criminal	Bayview	Bayview Hunters Point	Ms.
2018/03/28 01:54:00 PM	Other Miscellaneous	Central	Russian Hill	Ms.
2018/03/28 05:29:00 PM	Larceny Theft	Northern	Tenderloin	Ms.

Registered vehicles

Make	Model	Year	Color	Lice
Chevrolet	Cruze	2011	White	8NK
Mazda	Tribute	2003	White	7RQ

Mr. Richard Herrmann - CONFIDENTIAL - LAW ENFORCEMENT ONLY

Preliminary notes

Write notes here

Overview

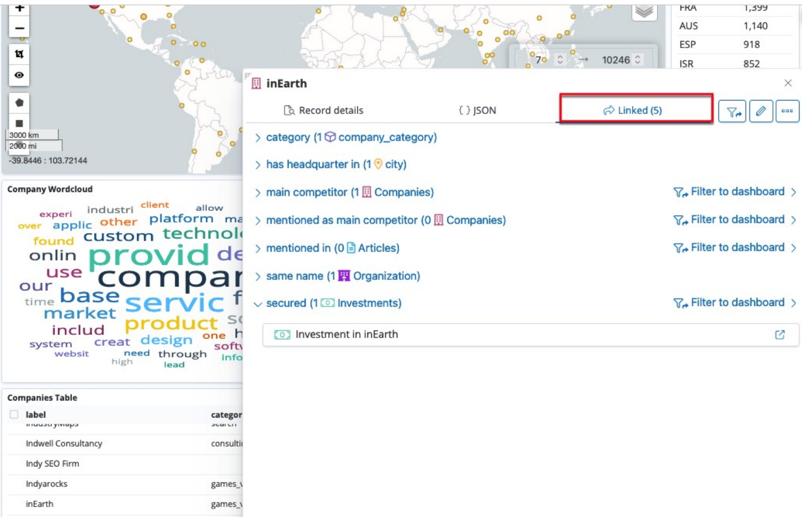
SPECIAL FLAGS	PROBATION	JAIL STATUS	VIOLENT HISTORY
no flags	Yes	Free	No

Personal data

Name Mr. Richard Herrmann	**Phone** 409-764-7093
Age 22	**Email** RichardDHerrmann@dayrep.com
Gender male	**Blood type** AB+
Address 1128, MONTGOMERY ST San Francisco 94133	**Occupation** Baker
City of Birth West Warwick (RI)	**Company** Arpin Van Lines

Searching data is one of the most fundamental actions of an investigator. Siren Platform provides this intuitively through the Global Search interface, where analysts can search all of the data accessible to Siren - according to their current security permissions.

From the search results, analysts can get rich reports in PDF or DOCX or use the results as a starting point in dashboard or graph investigations.



3000 km
2000 mi
-39.8446 : 103.72144
7
10246
FRA 1,399
AUS 1,140
ESP 918
ISR 852
Company Wordcloud
industri client allow experi over applic other platform found custom technol onlin provid use compar our time base servic market product includ system creat design one websit need through high lead softw info

inEarth

Record details
{ } JSON
Linked (5)

- category (1 company_category)
- has headquarter in (1 city)
- main competitor (1 Companies) — Filter to dashboard
- mentioned as main competitor (0 Companies) — Filter to dashboard
- mentioned in (0 Articles) — Filter to dashboard
- same name (1 Organization)
- secured (1 Investments) — Filter to dashboard
  - Investment in inEarth

Companies Table

label	categor
Indwell Consultancy	consulti
Indy SEO Firm	
Indyarocks	games_
inEarth	games_

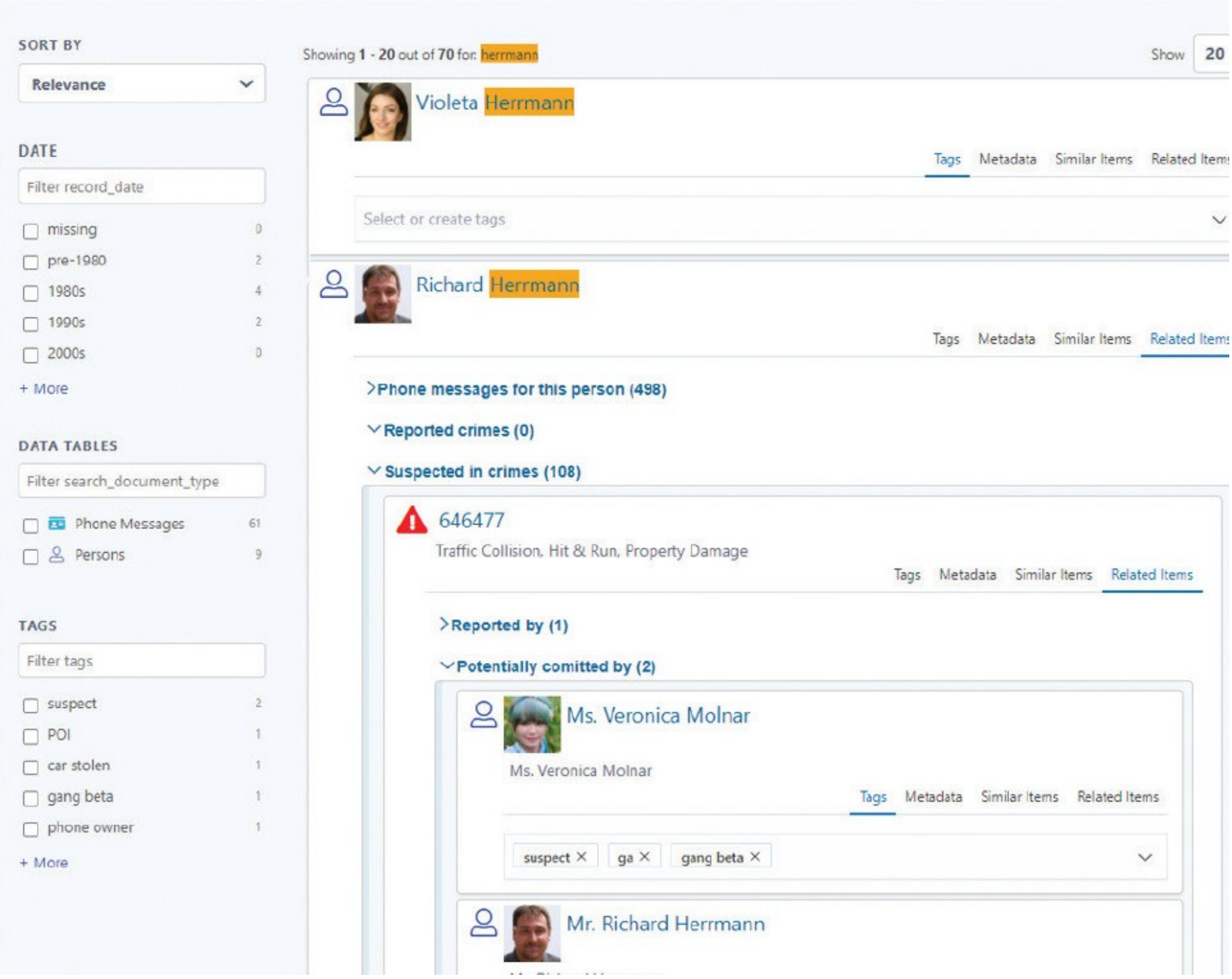
Advanced search for faces, signals, vector content and more

Siren leverages the backend capability to perform instantaneous searches in vector spaces. This allows you to search for advanced entities, such as faces, and media and images, but also specialized domain specific entities for example, in life science searching molecules by structure or content by meaning not by keyword. See the section on Image and Video processing for more details.

Data on the move with Siren Search UI

With Siren Investigate, analyzing an entity, or a set of entities, is a straightforward task for a qualified analyst. With the Siren Search UI, we can also support less experienced users on mobile devices.

In the example search below, you can see a record that was searched for by a mobile user who entered the family name of the suspect in the search field. The user can simply expand on the details of Suspected in crimes and navigate to view the associated records.



SORT BY

Relevance

DATE

Filter record_date

- □ missing 0
- □ pre-1980 2
- □ 1980s 4
- □ 1990s 2
- □ 2000s 0

+ More

DATA TABLES

Filter search_document_type

- □ Phone Messages 61
- □ Persons 9

TAGS

Filter tags

- □ suspect 2
- □ POI 1
- □ car stolen 1
- □ gang beta 1
- □ phone owner 1

+ More

Showing **1** - **20** out of **70** for: herrmann

Show 20

Violeta Herrmann

Tags Metadata Similar Items Related Items

Select or create tags

Richard Herrmann

Tags Metadata Similar Items Related Items

>Phone messages for this person (498)

∨Reported crimes (0)

∨Suspected in crimes (108)

646477

Traffic Collision, Hit & Run, Property Damage

Tags Metadata Similar Items Related Items

>Reported by (1)

∨Potentially comitted by (2)

Ms. Veronica Molnar

Ms. Veronica Molnar

Tags Metadata Similar Items Related Items

suspect × ga × gang beta ×

Mr. Richard Herrmann

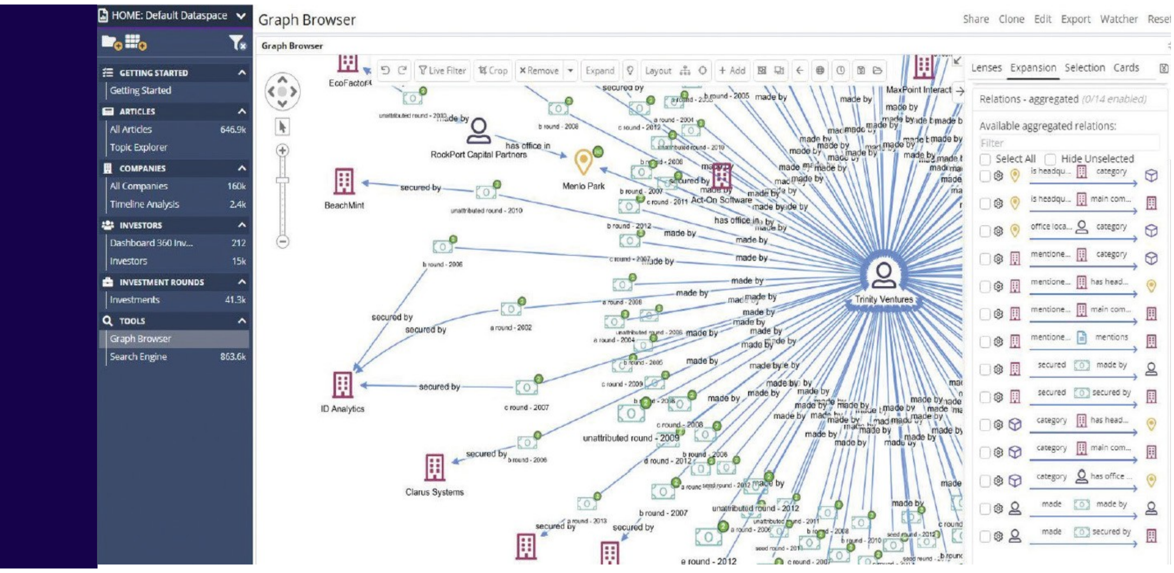
Record search and associative navigation as seen by an officer in the field on a mobile device.

Seamless link analysis at any stage of the investigation

Associative dashboards are extremely powerful, yet there are questions that no dashboard can answer. For example, in our scenario:

» Which investors invested in which companies?
» Are they investing in pairs or in groups?
» Are there groups that appear to be investing in competing companies?

To answer these questions, you can simply drag and drop any dashboard selection into the Graph Browser to start a link analysis.



Performing Link Analysis in the Graph Browser

Work live, on distributed Big Data

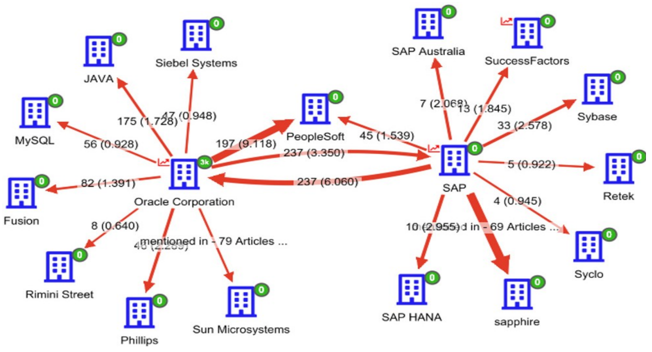
Link analysis operates by accessing a virtual graph that is created by the data model and reads data directly from Elasticsearch with no need for graph databases.
The graph allows the exploration of every record; and easily scales to over a billion records.

Important links first with graph aggregations

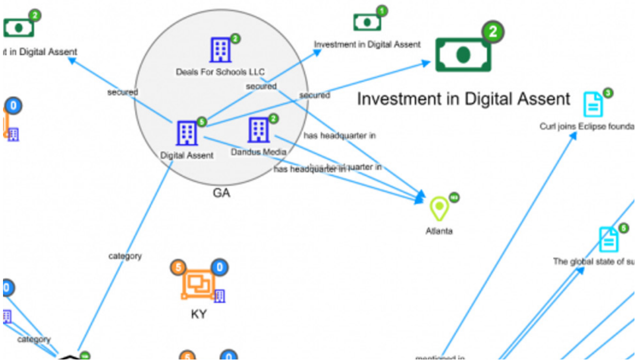
Arguably, the greatest limiting factor of link analysis is the clutter that can take over large graph explorations.
Siren Platform addresses this in powerful and unique ways. Big data aggregations can be activated at any time. In this modality, thousands or millions of intermediate nodes (messages, emails, or low-level data communications) are automatically summarized as metrics on the edges between connecting nodes.

In the following screenshot, we do not see the thousands of articles that co-mention the given companies, we just see a count for them (and in parentheses a relevance metric).

Likewise, it is always possible to group nodes based on a common data value, such as nationality, or manually expand groups as needed.



Seamless link analysis at any stage of the investigation



Grouping nodes in the Graph Browser

Graph expansion control

In a scenario where a node is connected to millions and millions of other nodes, how can expanding the graph be still useful?

Siren Platform is the only system that allows associative filters to restrict graph expansion. In the screenshot below, we have applied an associative filter "only articles mentioning French companies" to our Articles dashboard and we use this to control expansion on the graph.

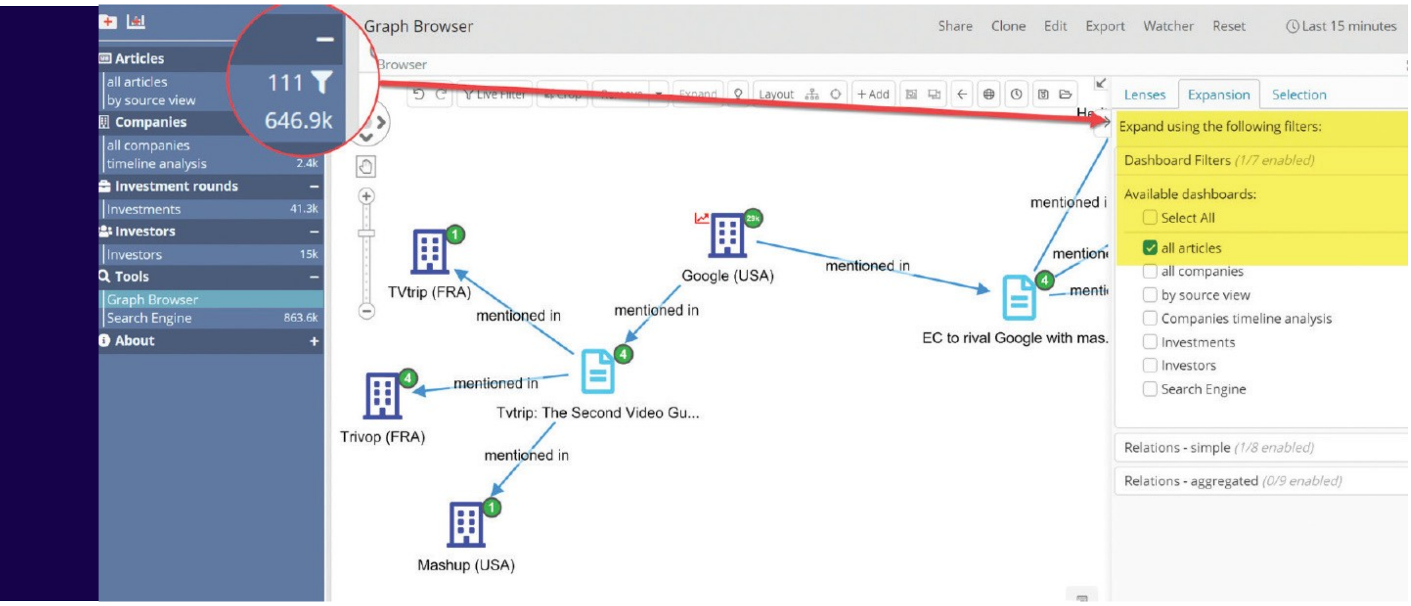
Whereas the Google node is normally connected to 20,000+ other nodes, in this example we see it connect only to the few articles that have that specific condition applied.

Conditional formatting and scripts

The Siren Platform graph supports powerful, scriptable, conditional formatting. This controls node and edge size, color, icon, halos, temporal and map position, tooltips, labels, additional glyphs visibility, and more.

You can install powerful scripts on the graph browser. These enable automatic expansion under certain conditions that facilitates specific use cases.

Likewise, administrators can install custom functions on buttons and contextual menus that can, for example, perform calls to Web services or return results from custom analytics scripts.



Expand using the following filters:

Dashboard Filters *(1/7 enabled)*

Available dashboards:

- □ Select All
- ☑ all articles
- □ all companies
- □ by source view
- □ Companies timeline analysis
- □ Investments
- □ Investors
- □ Search Engine

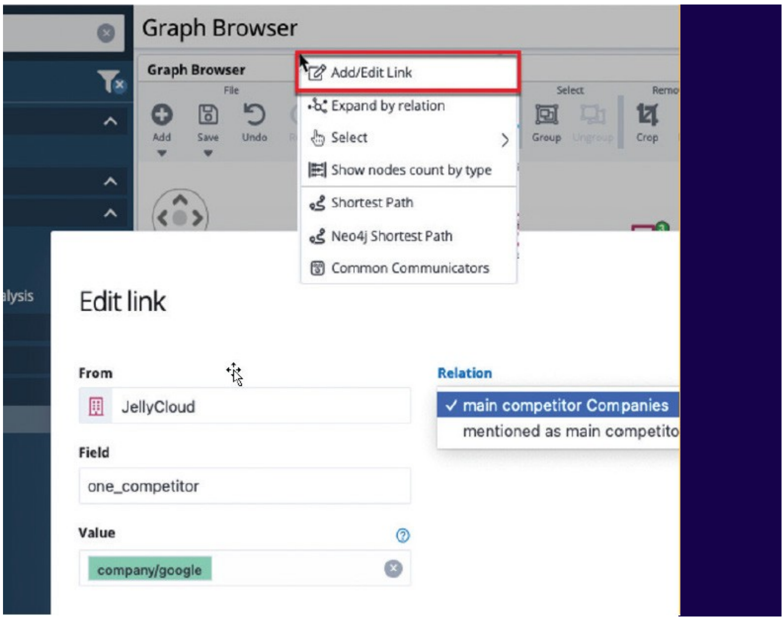
Relations - simple *(1/8 enabled)*

Relations - aggregated *(0/9 enabled)*

Graph expansion control using associative filters.

Graph authoring in the Siren Platform

Create and edit data as a graph



Edit link

From: JellyCloud

Field: one_competitor

Value: company/google

Relation: ✓ main competitor Companies
mentioned as main competito

Data curation and editing can also occur in the graph, and authored graphs can then be saved as part of the investigation.

This capability is typically used by analysts to either fix incorrect data or to create new records (nodes), links and compose new graphs.

The screenshot on the right shows the interface of how to add new links between nodes in a graph.

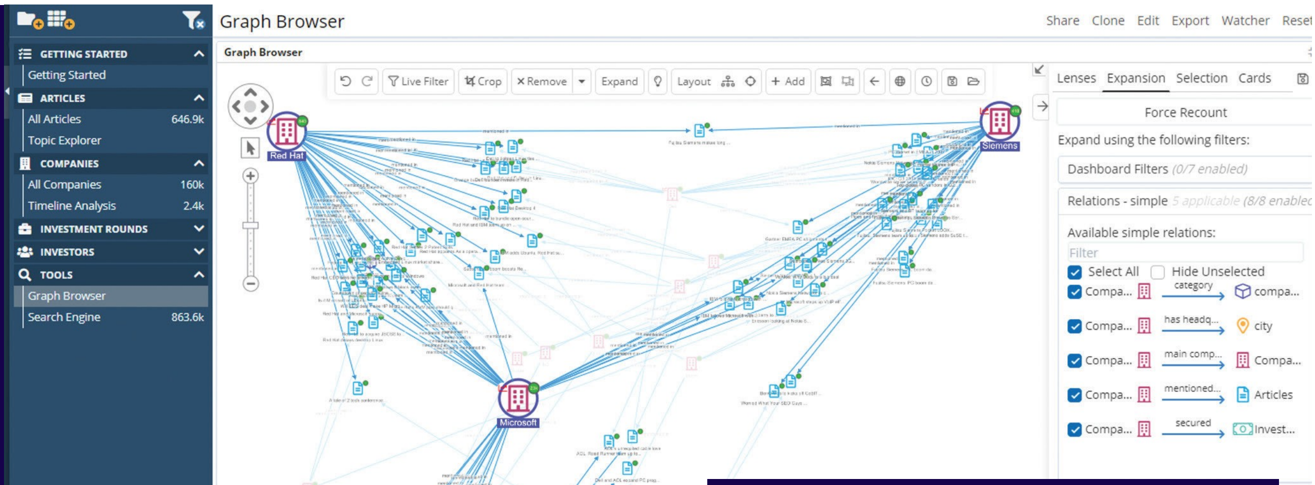
Export to PNG, PDF or IBM® i2® Analyst's Notebook®

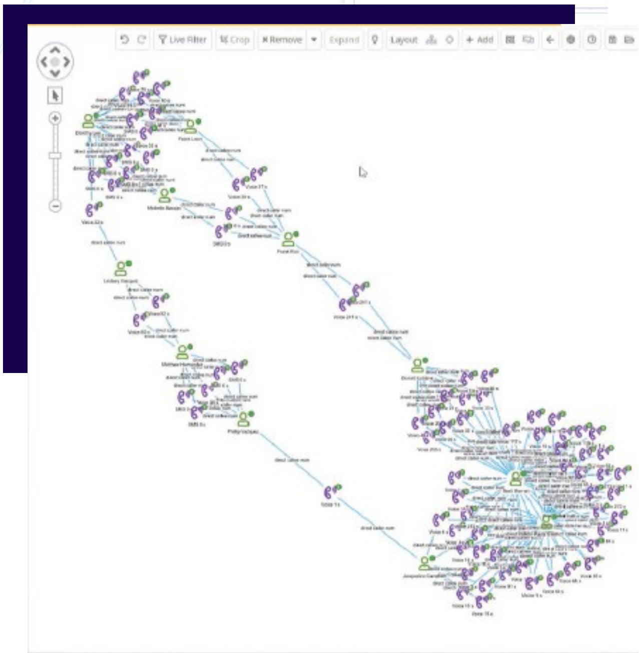
As with any other artifact generated in Siren Platform, graphs can be exported as PNG or PDF that you can print in high resolution up to 14,000 px X 10,000 px.

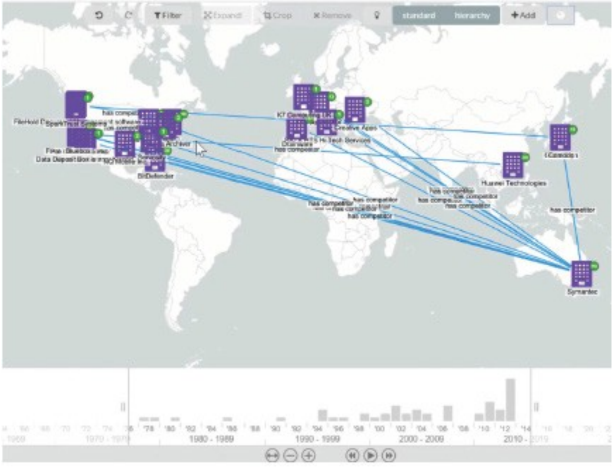
Alternatively, graphs can be exported in a format compatible with IBM i2 Analyst's Notebook, making Siren-to-i2 workflows a breeze.

Investigative graph algorithms at full data scale

Siren Platform offers a set of algorithms for graph analytics that operate on the visible graph – typically, with a size of up to 50 thousand nodes – or at full backend scale (over a billion nodes). Among the full backend scale algorithms, Siren Platform offers efficient shortest path and common communicator detection, which performs at full Elasticsearch scale, thanks to the capabilities of the Siren Federate plugin.


GETTING STARTED
Getting Started
ARTICLES
All Articles 646.9k
Topic Explorer
COMPANIES
All Companies 160k
Timeline Analysis 2.4k
INVESTMENT ROUNDS
INVESTORS
TOOLS
Graph Browser
Search Engine 863.6k
Graph Browser
Share Clone Edit Export Watcher Reset
Graph Browser
Live Filter
Crop
Remove
Expand
Layout
Add
Red Hat
Siemens
Microsoft
Lenses Expansion Selection Cards
Force Recount
Expand using the following filters:
Dashboard Filters (0/7 enabled)
Relations - simple 5 applicable (8/8 enabled
Available simple relations:
Filter
Select All
Hide Unselected
Compa... category compa...
Compa... has headq... city
Compa... main comp... Compa...
Compa... mentioned... Articles
Compa... secured Invest...





Siren Investigate Graph Capabilities

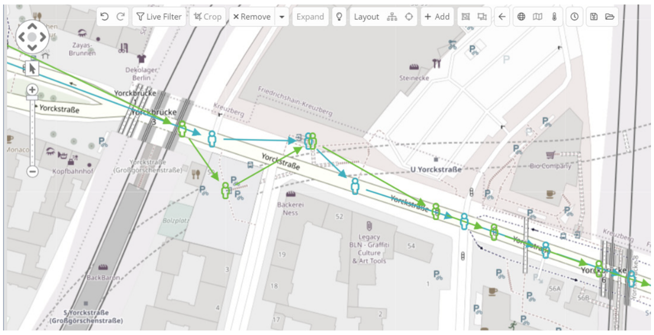
- **Round trip:** To and from dashboards using the same associative data model.
- **Operates at a billion record+** scale and in real time on data loaded in Elasticsearch.
- **Big data shortest path and common communicator** are native to Elasticsearch and other backend systems.
- **Graph metrics** include page rank, centrality, betweenness, closeness, eigenvector centrality, and more.
- **Scriptable**: Automatic behaviors and visualization rules can be scripted.
- **Web service integration**: Expand the graph by invoking Web services on the fly.
- **Spatial and temporal analysis**: See graph data on a map and as it evolves over time.

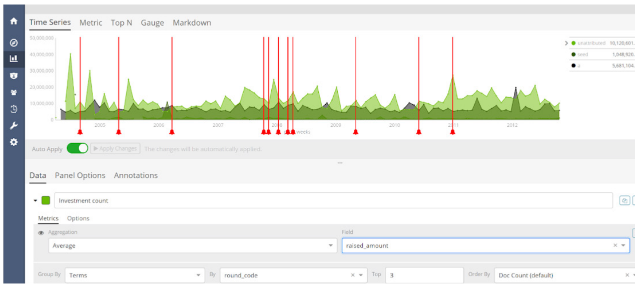
Spatial and temporal analysis

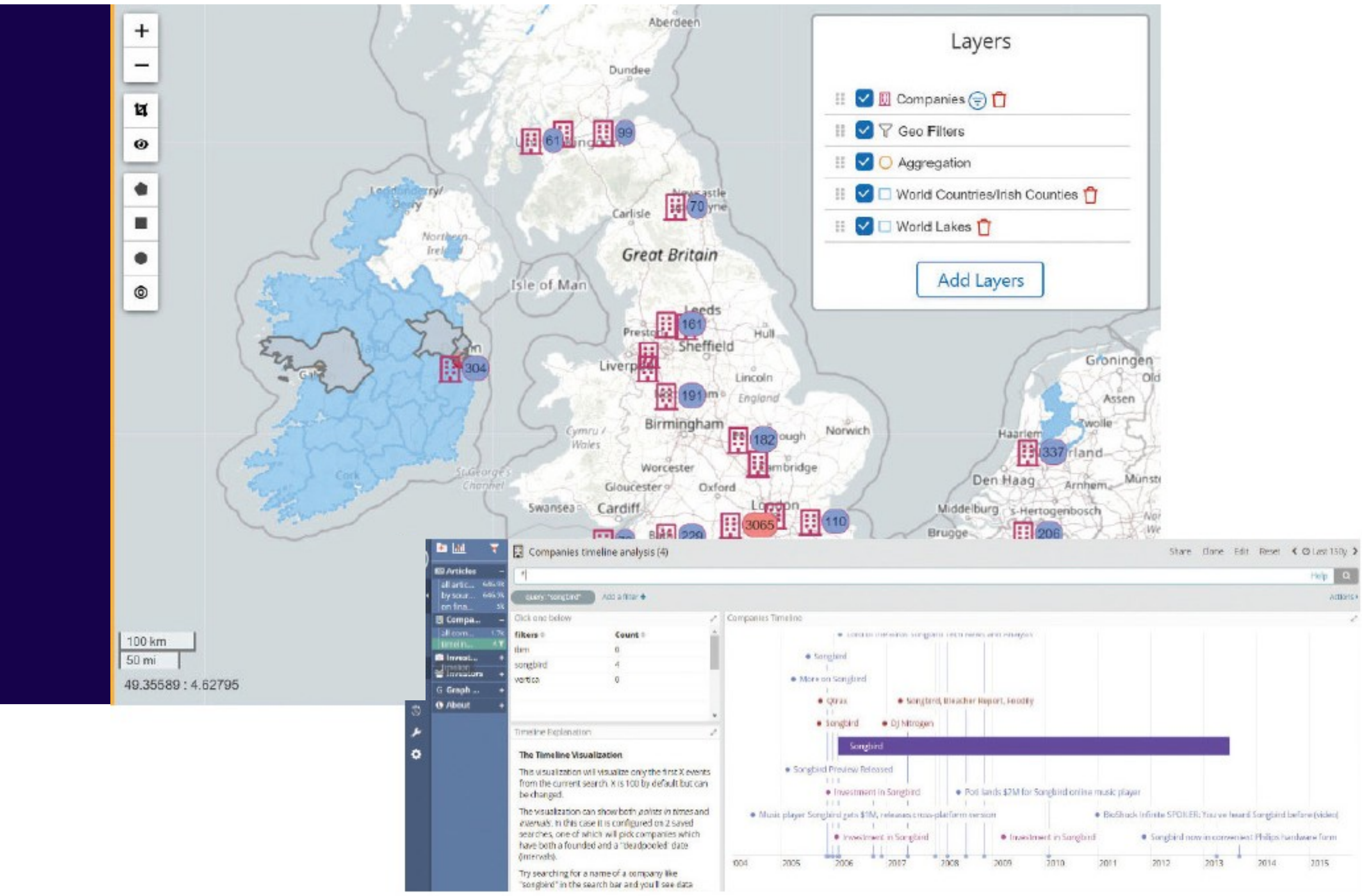
Siren has partnered with Esri to help organizations make smarter, faster decisions by solving business problems that require location intelligence. Siren Investigate is equipped with advanced spatial and temporal exploration capabilities.

esri Partner Network Bronze

- **From Big Data to individual records:** Interactive Big Data aggregations and drill downs to individual records.
- **Layers with enterprise security:** Browse thousands of layers based on the user's security level.
- **Track movements over time** and perform a historical movement analysis.
- **Run advanced algorithms** written in Python or other languages. For example, find meetings between people.
- **Visual time series builder**
- **Event overlap analysis**
- **Vector tile layer**: Create basemaps with custom styles and faster response time. Add data sets like census tracts and land parcels.







filters	Count
dem	0
songbird	4
vertica	0

Timeline Explanation

The Timeline Visualization

This visualization will visualize only the first X events from the current search. X is 100 by default but can be changed.

The visualization can show both *points in times* and *intervals*. In this case it is configured on 2 saved searches, one of which will pick companies which have both a founded and a "deadpooled" date (intervals).

Try searching for a name of a company like "songbird" in the search bar and you'll see data

Supporting the investigation life cycle

Keeping investigations distinct with dataspaces

In Siren Investigate, dataspaces allow you to create and clone investigative environments, typically to fit the needs of each distinct investigation.

With dataspaces, analysts have the freedom to:

» Modify the data model, uploading and integrating new datasets per investigation.

» Create custom dashboards and visualizations to suit specific needs.

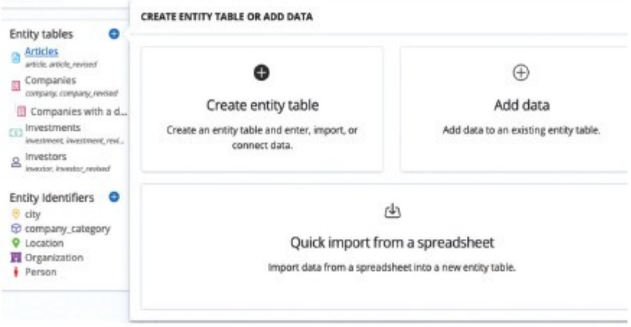
» Invite colleagues and work collaboratively.

Importing and transforming data

Data can enter the Siren Platform in the following ways:

» Importing spreadsheets
» Connecting to Elasticsearch indices
» Connecting to remote JDBC DataSources

» Invoking Web services



Create entity table
Create an entity table and enter, import, or connect data.

Add data
Add data to an existing entity table.

Quick import from a spreadsheet
Import data from a spreadsheet into a new entity table.

Importing data into entity tables in the Data model app

Data entry and data revision

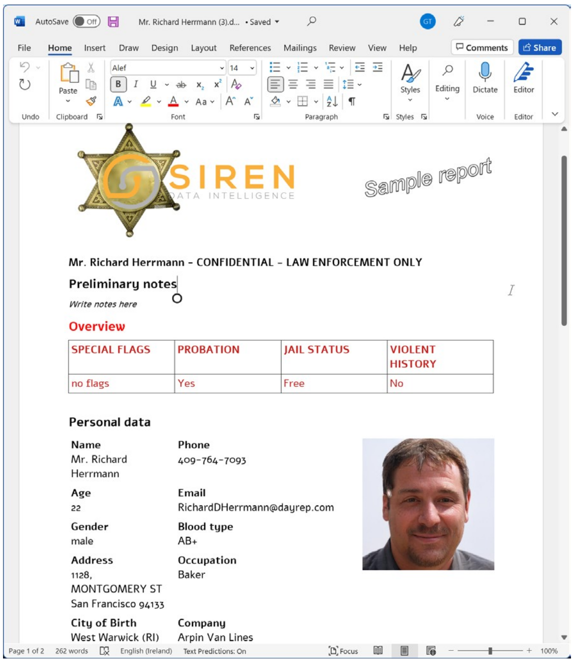
Authorized users can create new records and edit existing ones. This effectively allows data entry and the correction of mistakes. At any time, previous versions of the data are preserved, and all operations are monitored by the auditing system.

Exporting and reporting

Siren Platform includes the following features:

» Link sharing with filter statuses.

» PDF and PNG exports, which can be scheduled, triggered by conditions, or exported on demand. Siren Scripting API can extract PNGs from visualization to produce ad hoc reports from dashboards.



Mr. Richard Herrmann – CONFIDENTIAL – LAW ENFORCEMENT ONLY

Preliminary notes

Write notes here

Overview

SPECIAL FLAGS	PROBATION	JAIL STATUS	VIOLENT HISTORY
no flags	Yes	Free	No

Personal data

Name: Mr. Richard Herrmann
Phone: 409-764-7093
Age: 22
Email: RichardDHerrmann@dayrep.com
Gender: male
Blood type: AB+
Address: 1128, MONTGOMERY ST San Francisco 94133
Occupation: Baker
City of Birth: West Warwick (RI)
Company: Arpin Van Lines

Alerts by investigation

Siren Investigate can generate alerts by investigation based on scheduled queries, threshold alerts, and complex conditions. Alerts can generate emails, send PDFs, or call APIs to execute actions.

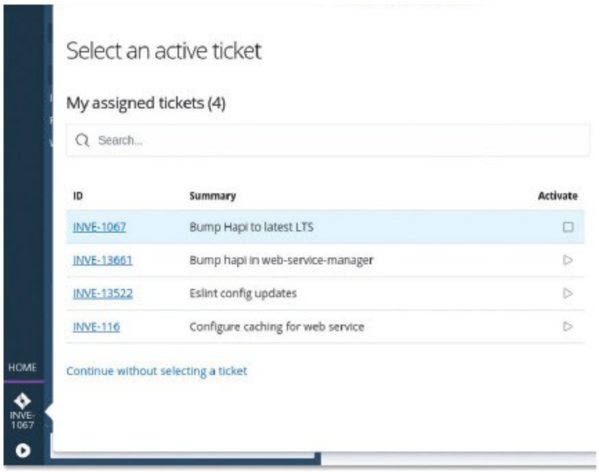
Alerts can be templated, with templates becoming available based on the context. In the adjacent screenshot, three alert templates are made available to the analyst.

Integrated Jira workflow

The Jira plugin allows you to connect Siren Investigate to a Jira cloud or server instance.

You can keep track of assigned tickets and export dashboard information as attachments to the tickets, directly from Siren Investigate.

The key features include searching for tickets, selecting an active ticket, and exporting dashboard information.



Select an active ticket

My assigned tickets (4)

Search...

ID	Summary	Activate
INVE-1067	Bump Hapi to latest LTS	□
INVE-13661	Bump hapi in web-service-manager	▷
INVE-13522	Eslint config updates	▷
INVE-116	Configure caching for web service	▷

Continue without selecting a ticket

Selecting a Jira ticket to track work in investigations

Augment your organizational data with external data providers

Siren has the ability to seamlessly integrate your internal organization data with external data providers both automatically and driven by analyst actions. Thanks to a flexible Web service driver model, Siren makes data coming from external providers naturally blend into the organization data model so that it becomes immediately available to complement the local knowledge.

There are countless uses for this capability:

» In cybersecurity, Siren data partners, such as Q6, can check emails and IP addresses for possible indication of compromise and warning investigators of potential illicit activities associated with them.

» Siren data partners such as Sayari, give access to over 500 million records of companies and registered executives worldwide.

» Open source and Dark Web intelligence specialist partner can be used to monitor mentions of people on social media as well as the Dark Web directly from Siren Investigate.

Web service invocations can happen seamlessly on the graph browser (driven by the analyst) but can also be automatically scheduled or used in ad hoc scripts automating investigative activities - for example, called automatically as part of report creations.

3 Security in Siren Platform

Access control, encryption, and integration with LDAP and Active Directory.

Siren Investigate has enterprise-level access control at the **index, record**, and **field** levels.

The following security is supported:

» **Elastic.co subscription**: This is the recommended option and is integrated in the Siren Platform Platinum Edition.

» **Alternative providers**: Siren Platform is also compatible with Elasticsearch clusters that are secured with Open Distro Security or Search Guard™.

With these security providers, Siren Investigate can connect with LDAP and Microsoft® Active Directory®. It can establish individual- and role-level access control for specific UI components, such as dashboards, visualizations, and saved search, and it can implement restrictions on data access.

End-to-end encryption exists from the user interface down to inter-cluster communications, and Elasticsearch can be encrypted at rest.

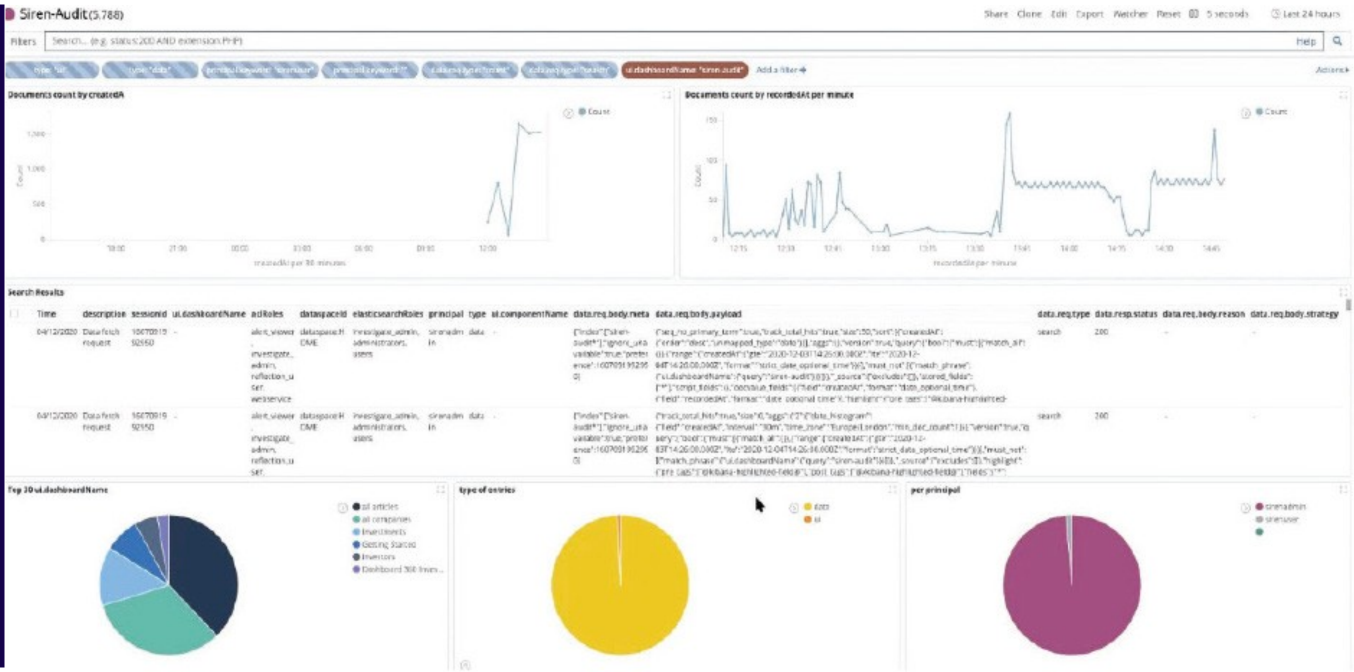
Siren also natively supports OpenID authentication and authorization, automatically generating a valid token from a client certificate without passing through external Identity and Access Management software.

User auditing

The Siren Investigate session audit feature allows you to log and perform internal audits on session user data. The system can log the following:

» **UI:** Actions that are performed by the user.
» **Saved objects:** Access to system configuration objects.
» **Records:** Store every record that was displayed to the user.
» **ACl:** User login and logout activity.

You can then configure a dashboard to search and track user activity for auditing purposes.



Time	description	sessionId	ui.dashboardName	adRoles	dataspaceId	elasticsearchRoles	principal	type	ui.componentName	data.req.body.meta	data.req.body.payload	data.req.type	data.resp.status	data.req.body.reason	data.req.body.strategy
04/12/2020	Data fetch request	1607091992950	-	alert_viewer, investigate_admin, reflection_user, webservice	dataspace:HOME	investigate_admin, administrators, users	sirenadmin	data	-	{"index":["siren-audit*"],"ignore_unavailable":true,"preference":1607091912960}	[illegible]	search	200	-	-
04/12/2020	Data fetch request	1607091992950	-	alert_viewer, investigate_admin, reflection_user,	dataspace:HOME	investigate_admin, administrators, users	sirenadmin	data	-	{"index":["siren-audit*"],"ignore_unavailable":true,"preference":1607091912960}	[illegible]	search	200	-	-

A dashboard configured to monitor and audit user activity

4 Investigative AI in Siren Platform

Siren Platform supports six typical AI scenarios:

1. Entity Resolution (Siren ER) – ML that recognizes that multiple records may be about the same entity (for example, a person or a company). Also finds suspiciously connected records.
2. Natural Language Processing (NLP) –Extracts locations, events, dates, amounts, sentiment, and salience (prominence) from text, and is multilingual.
3. Topic clustering – Real time clustering of corpuses of documents to extract key phrases and concepts.
4. Advanced graph pattern detection –Detects (and alerts about) patterns or new occurrences in the data.
5. Time series anomaly detection and individual behavior anomaly detection: (Federal US Market) Powered by Elastic Machine Learning, uses roll up indexes and anomaly detection for time series and for individual, based on summarizations of individual behavioral logs.
6. Apply arbitrary data science algorithms – Custom external algorithms can be invoked directly from the UI as a part of workflows.

Siren Entity Resolution

Siren Entity Resolution (Siren ER) is a Machine Learning component capable of:

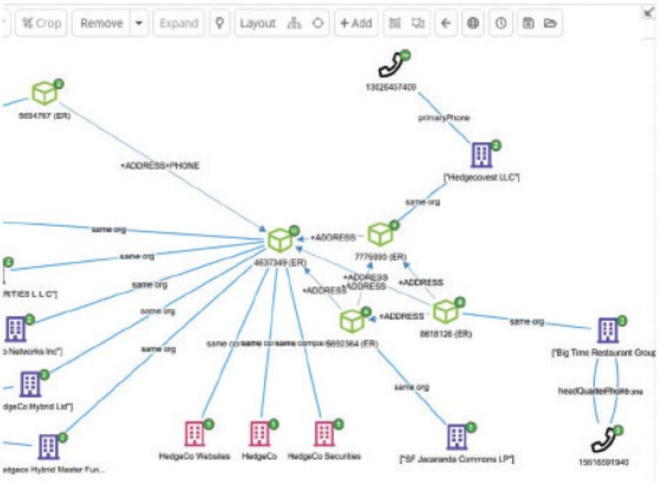
» Recognizing that two or more records are very likely to be referring to the same real-world entity - for example, the same person.

» Recognizing that two or more separate entities are interestingly connected - for example, they share an unusual combination of attributes.

Siren ER enables real-time (no re-calculation required) fuzzy matching on records at billion-record scale.

In addition, Siren ER can obtain suggested links across entities, which are likely different but linked by some strong similarity attributes - for example, entities with very similar addresses.

In the following screenshot, several records are joined by entity resolution cubes which, in turn, connect with other separate, but highly-related entities.



The data model graph in a financial investigation

Highlights of the Siren ER technology:

» Works across **schemata, languages**, and **cultural conventions** (Bob = Robert = Роберц) and plain human error.

» **Real time**: Corrects previous assertions based on new facts without recomputing all data.

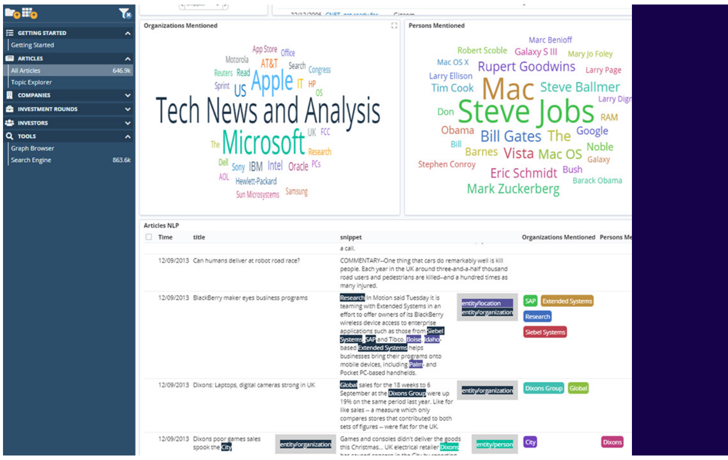
» **Machine learning based**: True ML thrives on as much data as it can be fed.

Natural Language Processing

Extract entities from text and visualize the connections across the corpus with the full-featured Natural Language Processing (NLP) engine, integrated in Siren Platform.

Extract mentions of **companies, people, locations, hashtags, phone numbers, addresses, IBANs, IPs, MDSs**, as well as tag terms and concepts that you can provide in custom taxonomies and lists.

In the screenshot to the right, Siren's NLP capabilities are at work on our demo news corpus, extracting mentions of the main entities among which are people and companies.



Articles NLP

□ Time	title	snippet	Organizations Mentioned	Persons Me
		a call.		
12/09/2013	Can humans deliver at robot road race?	COMMENTARY--One thing that cars do remarkably well is kill people. Each year in the UK around three-and-a-half thousand road users and pedestrians are killed--and a hundred times as many injured.		
12/09/2013	BlackBerry maker eyes business programs	Research In Motion said Tuesday it is teaming with Extended Systems in an effort to offer owners of its BlackBerry wireless device access to enterprise applications such as those from Siebel Systems, SAP and Tibco. Boise, Idaho-based Extended Systems helps businesses bring their programs onto mobile devices, including Palm- and Pocket PC-based handhelds. entity/location entity/organization	SAP Extended Systems Research Siebel Systems	
12/09/2013	Dixons: Laptops, digital cameras strong in UK	Global sales for the 18 weeks to 6 September at the Dixons Group were up 19% on the same period last year. Like for like sales -- a measure which only compares stores that contributed to both sets of figures -- were flat for the UK. entity/organization	Dixons Group Global	
12/09/2013	Dixons poor games sales spook the City entity/organization	Games and consoles didn't deliver the goods this Christmas... UK electrical retailer Dixons has caused concern in the City by reporting entity/person	City	Dixons

Natural language processing in Siren Investigate

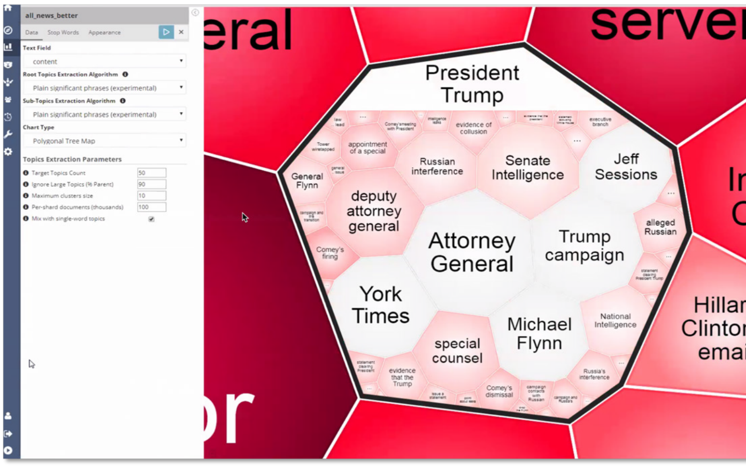
Editable NLP annotation

As NLP can be a noisy process, investigators can also manually add, remove, and fix annotations. This feature is part of a more general Siren feature which enables record editing in a safe and versioned way.

Real-time topic clustering

Siren topic clustering capabilities offer visual, real-time, interactive, large-scale, topic clustering and key phrase detection on a corpus of documents. The exploration requires no pre-processing, and works dynamically on any filtered set. In the screenshot to the right, we can see this at work on a political news corpus.

The components extract the key terms from the corpus, maximizing both coverage and significance.


all_news_better

Data | Stop Words | Appearance

Text Field: content

Root Topics Extraction Algorithm: Plain significant phrases (experimental)

Sub-Topics Extraction Algorithm: Plain significant phrases (experimental)

Chart Type: Polygonal Tree Map

Topics Extraction Parameters

- Target Topics Count: 50
- Ignore Large Topics (% Parent): 90
- Maximum clusters size: 10
- Per-shard documents (thousands): 100
- Mix with single-word topics: ☑

» **Coverage** – The resulting terms must represent as much as possible of the corpus (and can be read in the tooltip).
» **Significance** – How the terms in a sub-selection are strongly correlated with those in the parent (reflected in the color scheme).

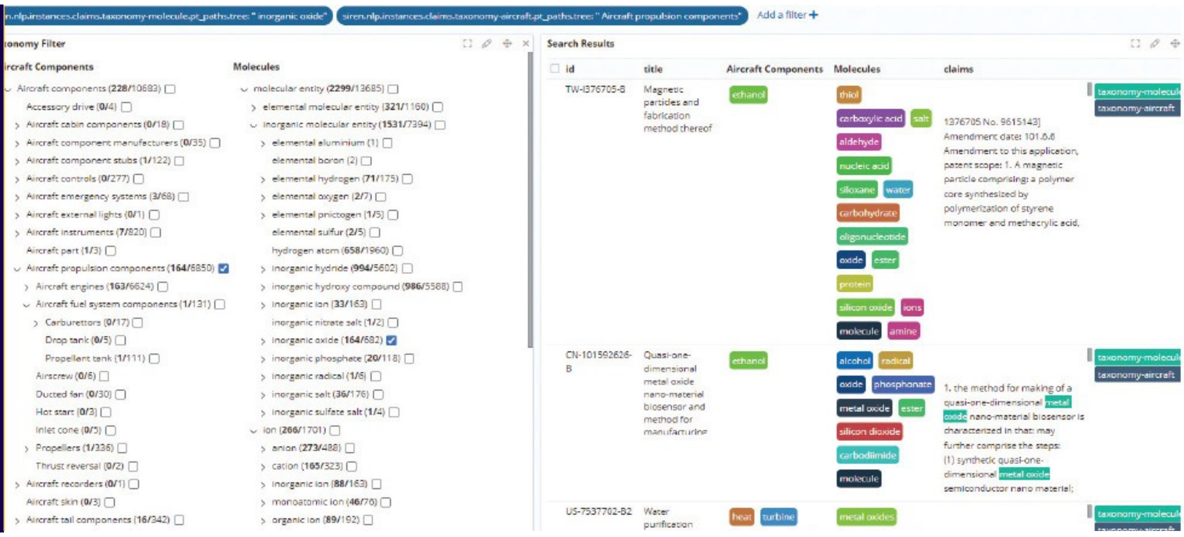
Taxonomy data browsing

For customers who are dealing with advanced textual content, such as technical or scientific news and documentation, Siren Platform can perform deep taxonomy tagging and browsing.

»Taxonomies with over tens of thousands of terms. » Providing synonyms at every level.

On the tagging side, the built-in Siren NLP component allows leveraging deep taxonomies, specifically:

In the following example, patent documents are filtered to display any inorganic oxide and any aircraft propulsion component. We find that the first document mentions metal oxide, an inorganic oxide, and propellant tank, which is an aircraft propulsion component.



n.nlp.instances.claims.taxonomy-molecule.pt_paths.tree: " inorganic oxide"
siren.nlp.instances.claims.taxonomy-aircraft.pt_paths.tree: " Aircraft propulsion components"
Add a filter +

onomy Filter

ircraft Components
Aircraft components (228/10683) □
Accessory drive (0/4) □
Aircraft cabin components (0/18) □
Aircraft component manufacturers (0/35) □
Aircraft component stubs (1/122) □
Aircraft controls (0/277) □
Aircraft emergency systems (3/68) □
Aircraft external lights (0/1) □
Aircraft instruments (7/820) □
Aircraft part (1/3) □
Aircraft propulsion components (164/6850) ☑
Aircraft engines (163/6624) □
Aircraft fuel system components (1/131) □
Carburettors (0/17) □
Drop tank (0/5) □
Propellant tank (1/111) □
Airscrew (0/6) □
Ducted fan (0/30) □
Hot start (0/3) □
Inlet cone (0/5) □
Propellers (1/336) □
Thrust reversal (0/2) □
Aircraft recorders (0/1) □
Aircraft skin (0/3) □
Aircraft tail components (16/342) □

Molecules
molecular entity (2299/13685) □
elemental molecular entity (321/1160) □
inorganic molecular entity (1531/7394) □
elemental aluminium (1) □
elemental boron (2) □
elemental hydrogen (71/175) □
elemental oxygen (2/7) □
elemental pnictogen (1/5) □
elemental sulfur (2/5) □
hydrogen atom (658/1960) □
inorganic hydride (994/5602) □
inorganic hydroxy compound (986/5588) □
inorganic ion (33/163) □
inorganic nitrate salt (1/2) □
inorganic oxide (164/682) ☑
inorganic phosphate (20/118) □
inorganic radical (1/6) □
inorganic salt (36/176) □
inorganic sulfate salt (1/4) □
ion (266/1701) □
anion (273/488) □
cation (165/323) □
inorganic ion (88/163) □
monoatomic ion (46/76) □
organic ion (89/192) □

Search Results

□ id	title	Aircraft Components	Molecules	claims	
TW-I376705-B	Magnetic particles and fabrication method thereof	ethanol	thiol carboxylic acid salt aldehyde nucleic acid siloxane water carbohydrate oligonucleotide oxide ester protein silicon oxide ions molecule amine	1376705 No. 9615143] Amendment date: 101.6.6 Amendment to this application, patent scope: 1. A magnetic particle comprising: a polymer core synthesized by polymerization of styrene monomer and methacrylic acid,	taxonomy-molecul taxonomy-aircraft
CN-101592626-B	Quasi-one-dimensional metal oxide nano-material biosensor and method for manufacturing	ethanol	alcohol radical oxide phosphonate metal oxide ester silicon dioxide carbodiimide molecule	1. the method for making of a quasi-one-dimensional metal oxide nano-material biosensor is characterized in that: may further comprise the steps: (1) synthetic quasi-one-dimensional metal oxide semiconductor nano material;	taxonomy-molecul taxonomy-aircraft
US-7537702-B2	Water purification	heat turbine	metal oxides		taxonomy-molecul taxonomy-aircraft

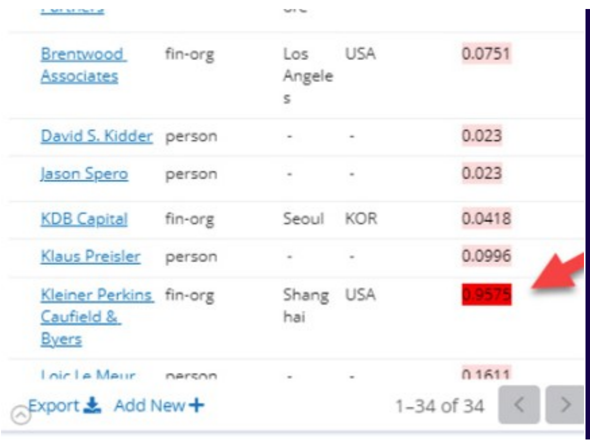
Elasticsearch transforms and outlier detection

When an analyst zones in on 100 individual companies or people in an investigation there is no clear way of knowing which one to investigate first. For scenarios like this, AI comes in handy and helps analysts prioritize their efforts and focus their attention.

Siren Platform offers a set of scripts and procedures to determine which entities stand out from others and why.

It works by leveraging the UI-driven procedures that are in Siren Platform Platinum Edition (bundled with Elasticsearch platinum subscription nodes) to run AI outlier detection and to enrich existing records with outlier scores.

The outlier score appears as an extra field in the target records, which can be used to sort records or to display a color coding scheme. Analysts can view details about how the anomaly score is composed.



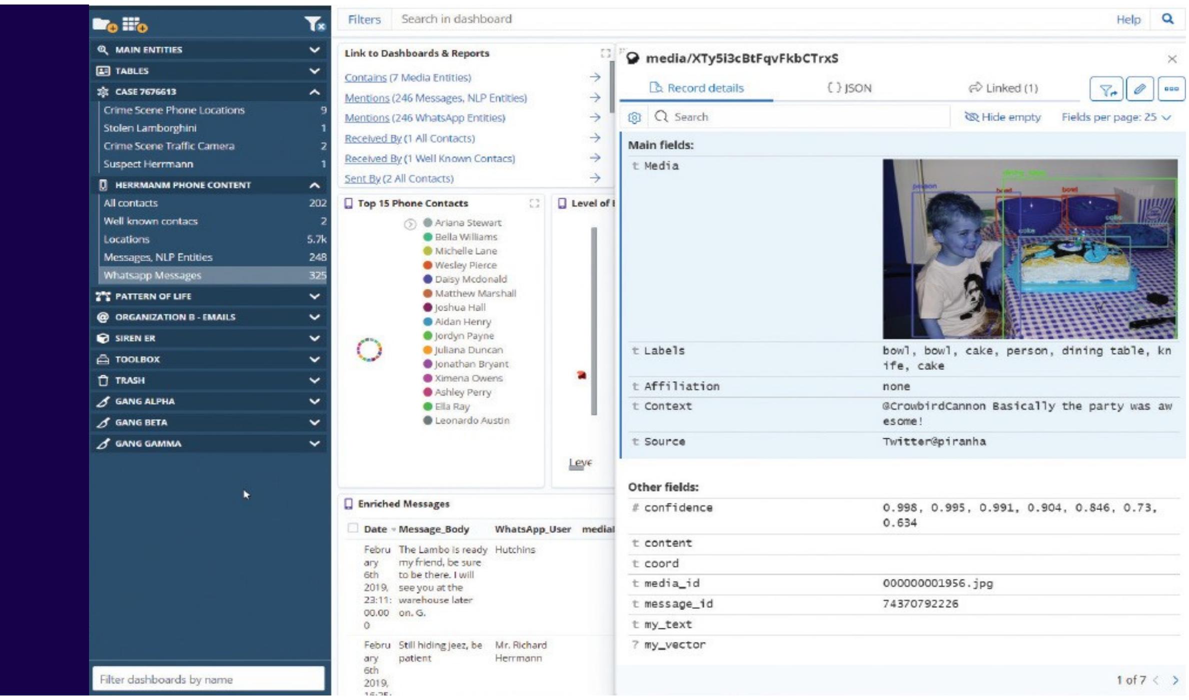
Brentwood Associates	fin-org	Los Angeles	USA	0.0751
David S. Kidder	person	-	-	0.023
Jason Spero	person	-	-	0.023
KDB Capital	fin-org	Seoul	KOR	0.0418
Klaus Preisler	person	-	-	0.0996
Kleiner Perkins Caufield & Byers	fin-org	Shanghai	USA	0.9575
Loic Le Meur	person	-	-	0.1611

Export ⭳ Add New + 1–34 of 34 < >

5 Integrating with image and video processing (beta)

The incredible volume of media generated by the digitized society we live in can no longer be analyzed by people without technical aid.

It is easy to integrate Siren with industry leading open source or commercial image analysis solutions to provide investigators with automated audio/visual content analysis.



Link to Dashboards & Reports

- Contains (7 Media Entities) →
- Mentions (246 Messages, NLP Entities) →
- Mentions (246 WhatsApp Entities) →
- Received By (1 All Contacts) →
- Received By (1 Well Known Contacs) →
- Sent By (2 All Contacts) →

Enriched Messages

□	Date	Message_Body	WhatsApp_User	medial
	February 6th 2019, 23:11:00.000	The Lambo is ready my friend, be sure to be there. I will see you at the warehouse later on. G.	Hutchins	
	February 6th 2019, 16:25:	Still hiding jeez, be patient	Mr. Richard Herrmann	

media/XTy5i3cBtFqvFkbCTrxS

Main fields:

t Media	
t Labels	bowl, bowl, cake, person, dining table, knife, cake
t Affiliation	none
t Context	@CrowbirdCannon Basically the party was awesome!
t Source	Twitter@piranha

Other fields:

# confidence	0.998, 0.995, 0.991, 0.904, 0.846, 0.73, 0.634
t content	
t coord	
t media_id	000000001956.jpg
t message_id	74370792226
t my_text	
? my_vector	

Typical capabilities that can be integrated include:

» Object and person tagging.

» Face vector extraction and search.

» Ability to train the tool for custom object categories, such as guns, and drugs.

» Image similarity search based on smart deep learning extracted vectors.

» Text-in-image search, such as license plates, addresses, and names.

» General image category detection.

Additional capabilities:

» Audio transcript from videos.
» Audio event extraction.
» Integration with third-party curated datasets
that are law enforcement and intelligence specific.

6 Solution templates: Kick-start your use case

Siren Platform comes with three solution templates, which are pre-made example configurations that can quickly kick-start a Siren project.

Cybersecurity and operational analytics

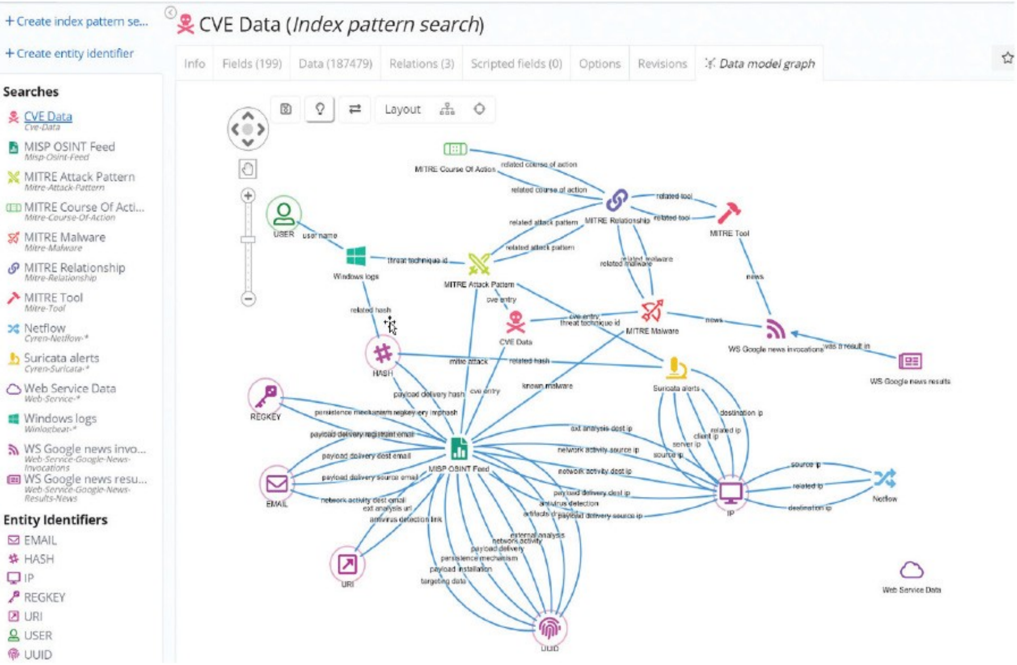
Connectors, data model, and rules for top of the pyramid threat hunting and monitoring, also powered by Elastic Common Schema and MITRE ATT&CK data model, see the screenshot below.

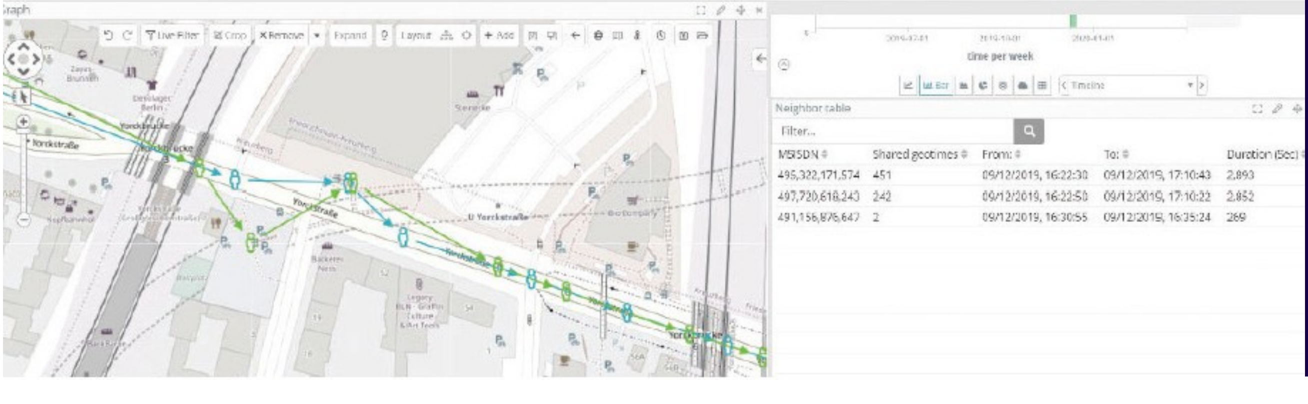
For more information, see Operational intelligence and cybersecurity on the Siren website.

Open-Source Intelligence (OSINT) and Signals Intelligence (SIGINT)

At Siren, OSINT and SIGINT templates address some of the most critical problems in national security.

The templates allow analysts to perform inference, based on Big Data quantities of signals, such as positions and calls, and to detect alert conditions. Contact Siren for more information.





Neighbor table

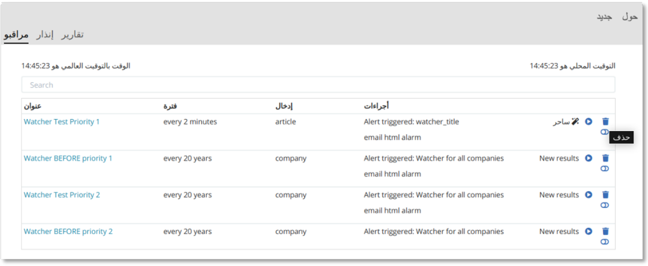
Filter...

MSISDN	Shared geotimes	From:	To:	Duration (Sec)
495,322,171,574	451	09/12/2019, 16:22:30	09/12/2019, 17:10:43	2,893
497,720,618,243	242	09/12/2019, 16:22:50	09/12/2019, 17:10:22	2,852
491,156,876,647	2	09/12/2019, 16:30:55	09/12/2019, 16:35:24	269

7 Siren Platform for integrators

Internationalization and white-labelling

The Siren Investigate UI currently supports seven locales, which include Spanish, French, German, Japanese, Korean, Chinese and Arabic. More can be added on demand.



حول جديد

مراقبو إنذار تقارير

الوقت بالتوقيت العالمي هو 14:45:23   التوقيت المحلي هو 14:45:23

Search

عنوان	فترة	إدخال	أجراءات	
Watcher Test Priority 1	every 2 minutes	article	Alert triggered: watcher_title email html alarm	ساحر
Watcher BEFORE priority 1	every 20 years	company	Alert triggered: Watcher for all companies email html alarm	New results
Watcher Test Priority 2	every 20 years	company	Alert triggered: Watcher for all companies email html alarm	New results
Watcher BEFORE priority 2	every 20 years	company	Alert triggered: Watcher for all companies	New results

Partners creating solutions with Siren can easily white- label Siren Investigate by changing names, logos, and icons. Together with a custom CSS, Siren can smoothly fit into larger integrations.

Creating custom behaviors with the Siren scripting API

Siren Platform is rooted in open source. You can easily extend it in a way similar to how you would extend the ELK (Elasticsearch/Logstash/Kibana) stack, from which Siren Platform is derived.

Furthermore, many parts of Siren Investigate offer the ability to write scripts for custom templating, transformations, workflows, and visualizations.

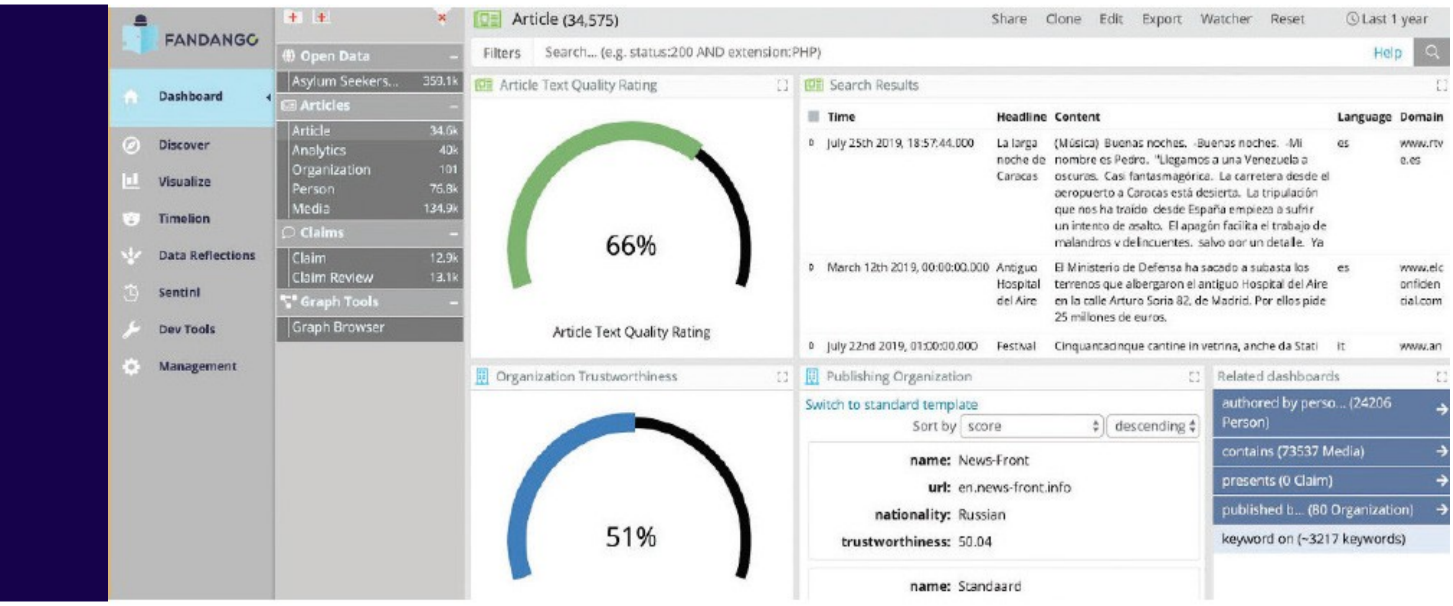
Scripts are written by administrators or experts in JavaScript and executed in a safety sandbox in dashboards.

The following are just a few examples of behaviors that can be scripted but the possibilities are endless.

Scriptable elements in Siren Investigate:

» Additional graph functions as administrative JavaScript extensions.
» Additional graph formatters.
» Dashboard special behaviors and custom-scripted
UIs.
» Record visualization templates.
» Custom interfaces to look up external services.

Siren also provides the full source code and training to qualifying partners to write their own plugins.



Time	Headline	Content	Language	Domain
July 25th 2019, 18:57:44.000	La larga noche de Caracas	(Música) Buenas noches. -Buenas noches. -Mi nombre es Pedro. "Llegamos a una Venezuela a oscuras. Casi fantasmagórica. La carretera desde el aeropuerto a Caracas está desierta. La tripulación que nos ha traido desde España empieza a sufrir un intento de asalto. El apagón facilita el trabajo de malandros y delincuentes. salvo por un detalle. Ya	es	www.rtve.es
March 12th 2019, 00:00:00.000	Antiguo Hospital del Aire	El Ministerio de Defensa ha sacado a subasta los terrenos que albergaron el antiguo Hospital del Aire en la calle Arturo Soria 82, de Madrid. Por ellos pide 25 millones de euros.	es	www.elconfidencial.com
July 22nd 2019, 01:00:00.000	Festival	Cinquantacinque cantine in vetrina, anche da Stati	it	www.an

Sort by score descending

name: News-Front
url: en.news-front.info
nationality: Russian
trustworthiness: 50.04

name: Standaard

An example of a white-labelled Siren Investigate UI

8 Architecture: building your investigative capabilities on the Siren Platform

Siren Platform is a highly-configurable environment that organizations can use to create Investigative Intelligence solutions at scale.

In building such solutions, organizations can start from an empty installation or from a Siren domain template.

Siren Federate augments Elasticsearch as follows:

» Data model and graph capabilities such as exploration, shortest path, and central node detections.
» Real time / interactive speed big data joins (correlations) among data tables and streams.
» Virtualization (live connections) to other back-end systems and Web services.

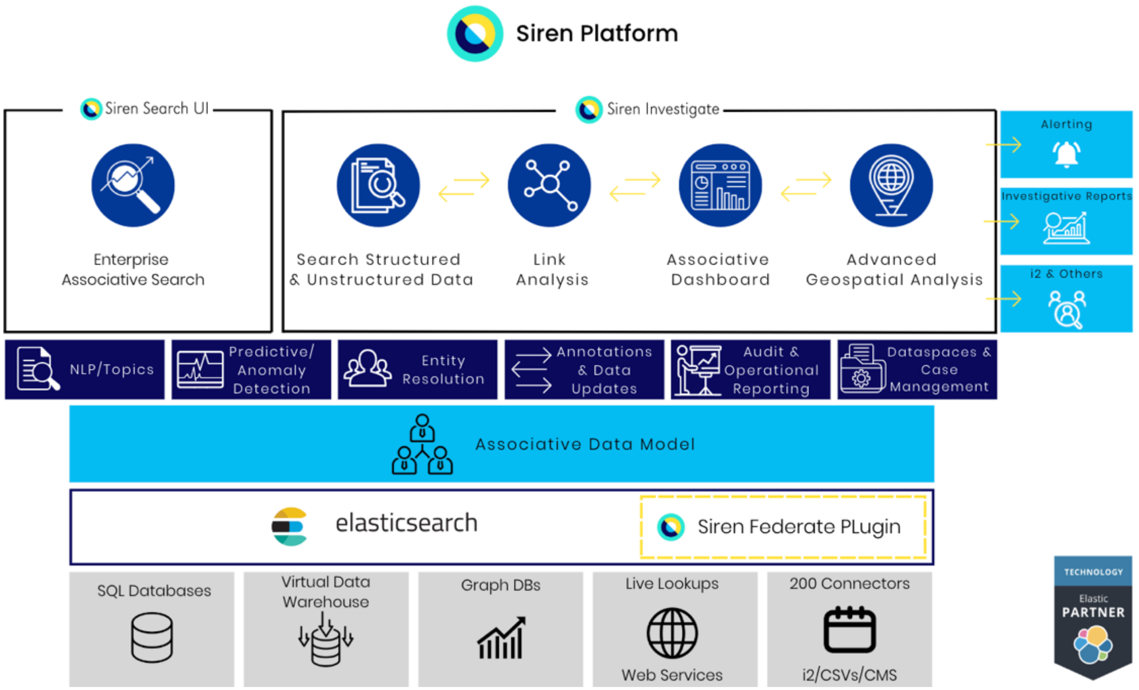
Architectural Overview

Siren is developed as an enhancement to the ElasticsearchTM big data search engine infrastructure, ElasticsearchTM, and developed in collaboration and partnership with Elastic.co.

The architecture centers around an Elasticsearch cluster, in which Siren Platform's proprietary technology - the Siren Federate plugin - is installed. Siren supports Elasticsearch 8.x as primary or remote clusters both on premise and Elastic Cloud.

At the core of the front end system, Siren Investigate is the search engine in which analysts can search and filter data. A mobile friendly search interface assists users who are on the move.

The data model enriches the data and enhances the knowledge graph and dashboard capabilities. Added to this is a layer of investigative AI: natural language processing, entity resolution, and anomaly detection. The platform is completed by alerting, auditing, and support for workflow management systems.



9 About Siren

Named a Gartner Cool Vendor in 2020 for Analytics and Data Science, Siren provides **investigative intelligence** software to some of the world's largest and most complex organizations.

Our product, Siren Platform, allows businesses to derive valuable insights from big data.

Founded in 2010 by thought leaders in Information Retrieval and distributed Knowledge Representation, Siren now employs a range of experts in data discovery, advanced search, data science, and engineering. Siren's headquarters are in Galway, Ireland, with offices in Europe and in the USA.

In 2018, Siren was recognized with awards for Technology Innovation of the Year and Start-up of the Year at the Ireland Industry and Tech Excellence Awards. In 2019, Siren was honored with the Enterprise of the Year award in the SME category at the ITAG awards. In 2022, Siren was recognized by Deloitte as one of the 50 fastest growing technology companies across the island of Ireland.

For more information, visit www.siren.io.

Headquarters

15 Market St,
Galway,
H91 TCX3,
Ireland
Tel: +353 (0)91 704 885

Legal Notices

Siren Platform™ is a trademark of Sindice Ltd. trading as Siren, with offices in 15 Market St, Galway, H91 TCX3, Ireland.

Elasticsearch™ is a trademark of Elasticsearch B.V., registered in the U.S. and in other countries.
IBM® i2® Analyst's Notebook® is a registered trademark of IBM.

Search Guard™ is a trademark of floragunn GmbH.

Microsoft® Active Directory® is a registered trademark of Microsoft Corporation.
Docker® is a registered trademark of Docker, Inc.

Gartner Disclaimer: GARTNER is a registered trademark and service mark of Gartner, Inc., and/or its affiliates in the U.S. and internationally, and is used herein with permission. All rights reserved.
Gartner, Cool Vendors in Analytics & Data Science, Julian Sun, Jim Hare, Rita Sallam, James Richardson, Afraz Jaffri, 7 May 2020.

The GARTNER COOL VENDOR badge is a trademark and service mark of Gartner, Inc. and/or its affiliates and is used herein with permission. All rights reserved. Gartner does not endorse any vendor, product or service depicted in its research publications and does not advise technology users to select only those vendors with the highest ratings or other designation. Gartner research publications consist of the opinions of Gartner's Research & Advisory organization and should not be construed as statements of fact. Gartner disclaims all warranties, expressed or implied, with respect to this research, including any warranties of merchantability or fitness for a particular purpose.
Deloitte® and the Deloitte logo are registered trademarks of entities within the Deloitte Network

WEB

WWW.SIREN.IO

EMAIL

INFO@SIREN.IO