Accenture Technology Labs

Elvis Hovor
@kofibaron

Shimon Modi
@shimonmodi

Shaan Mulchandani
@alabama_shaan

Unstructured Threat Intelligence Processing using NLP

Enhancing Cyber Security Operations by Automating Threat Intelligence Extraction from Unstructured Sources

Human driven processes are inadequate to effectively utilize increasing amount of cyber threat intelligence

Challenges

- Threat intelligence provided in advisories, reports and other text formats require human analysts to parse and extract relevance.
- Difficult to scale human driven processes for increased amount of threat data that needs to analyzed.
- Multiple sources detail the same threat intelligence leading to wasted analysis effort and inefficiencies in general.
- Rely on human memory to spot connections among disparate advisories.

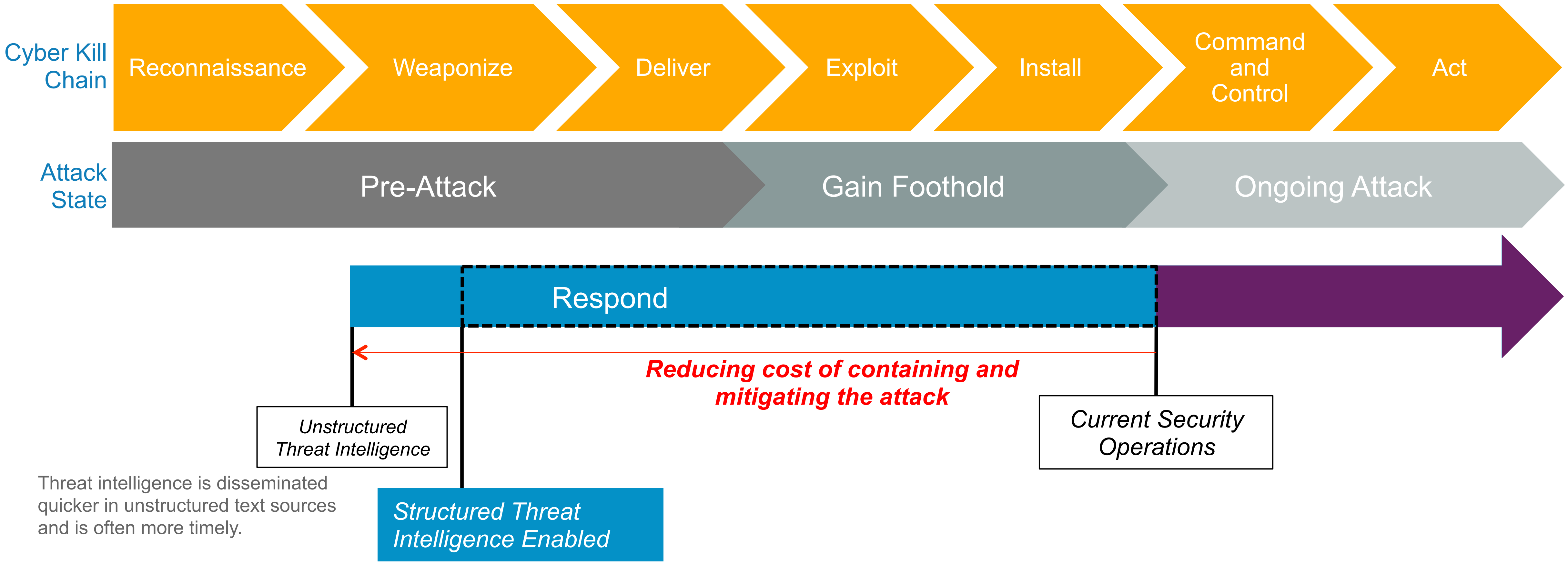
Desired Outcomes of Using Unstructured Threat Data

Solution that allows Threat Intelligence teams to utilize their human expertize by automating extraction of threat data from unstructured text sources

Desired Outcomes

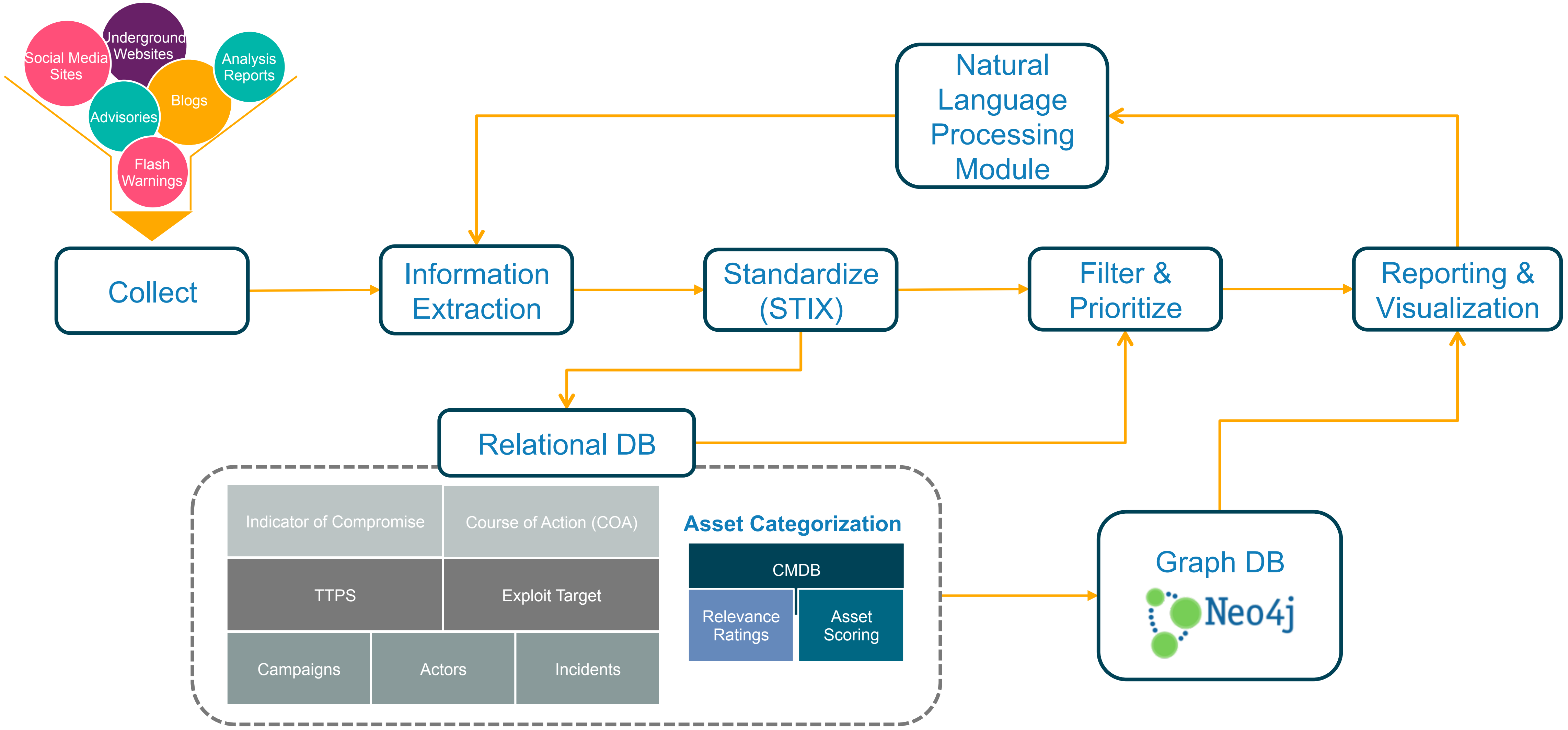
- Threat intelligence provided in advisories, reports and other text formats require human analysts to parse and extract relevance.
- Difficult to scale human driven processes for increased amount of threat data that needs to analyzed.
- Multiple sources detail the same threat intelligence leading to wasted analysis effort and inefficiencies in general.
- Rely on human memory to spot connections among disparate advisories.
- Conduct automated threat searches specific actors or TTPs
- Track threat data on indicators, actors and TTPs from multiple sources with minimal manual effort

Closing the Attacker’s Window of Opportunity



Threat intelligence is disseminated quicker in unstructured text sources and is often more timely.

Solution Framework



Using the Unstructured Threat Intelligence Processing Tool

☰ UTIP

Shaan

- Dashboard
- Prioritized List
- Data Upload
- Historical Analysis
- analytics-dasboard-viz
- analytics-doc
- analytics-node

Upload

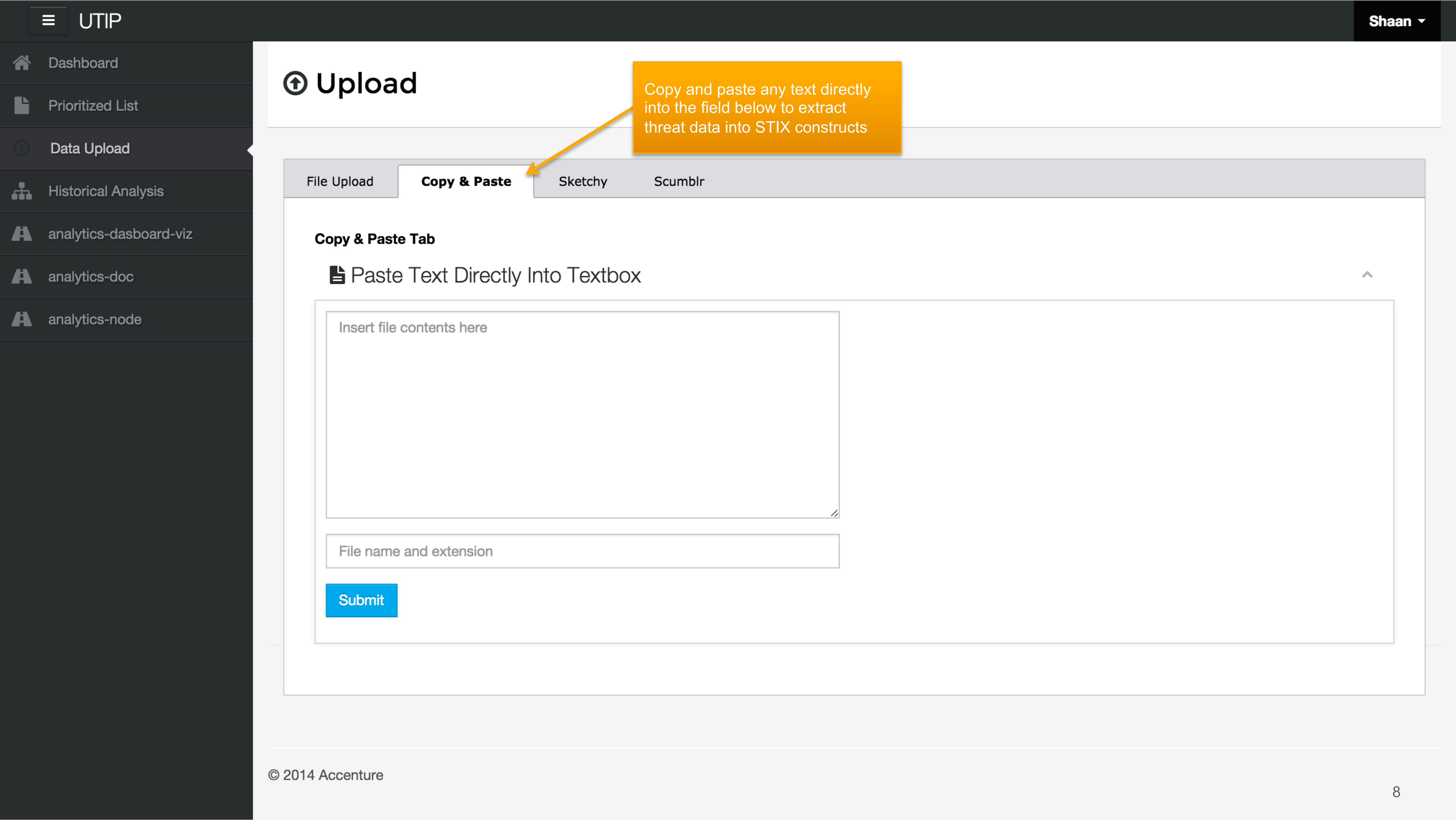
Upload a file for the parsing of threat data into STIX data constructs

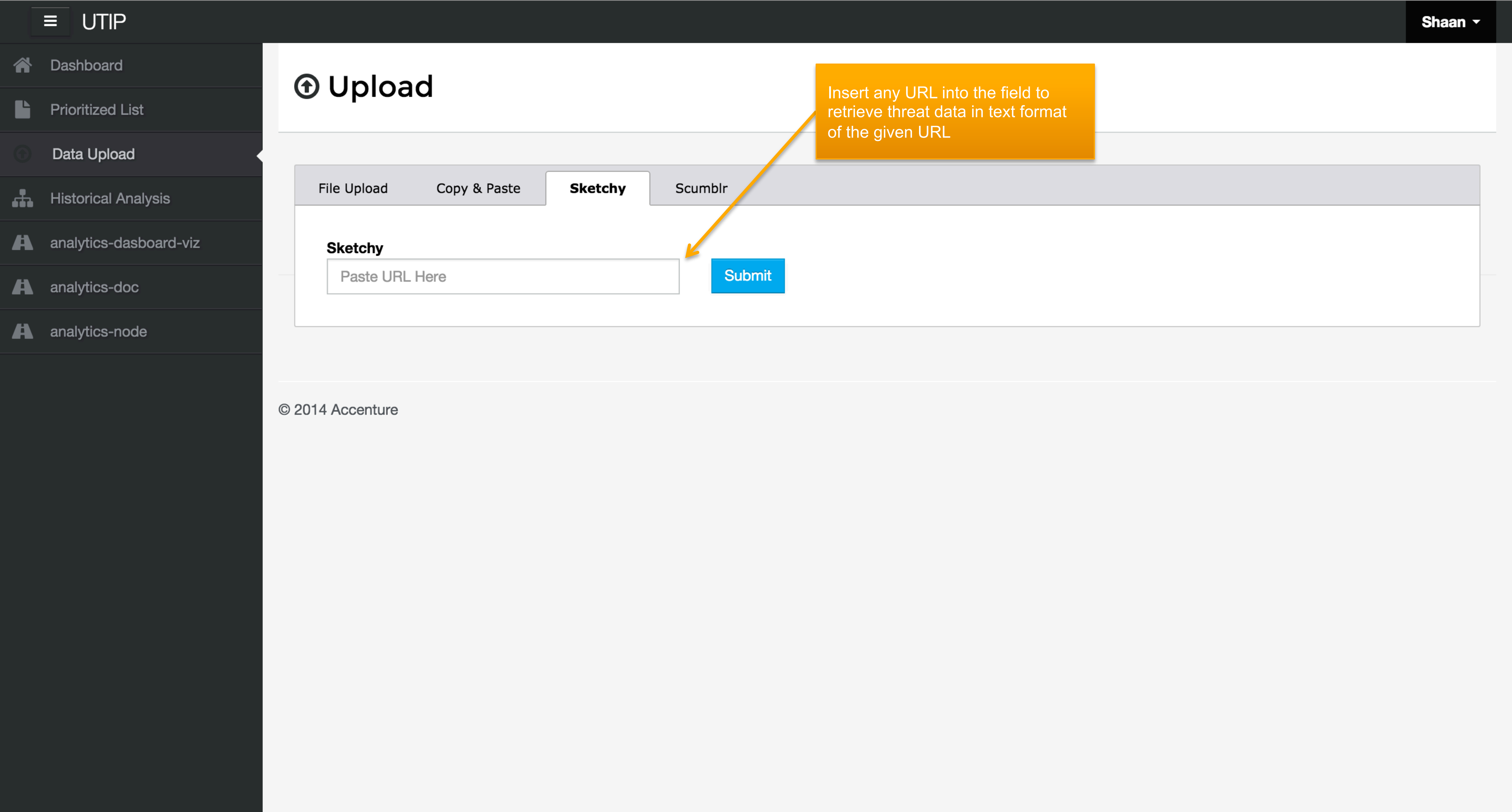
File Upload | Copy & Paste | Sketchy | Scumblr

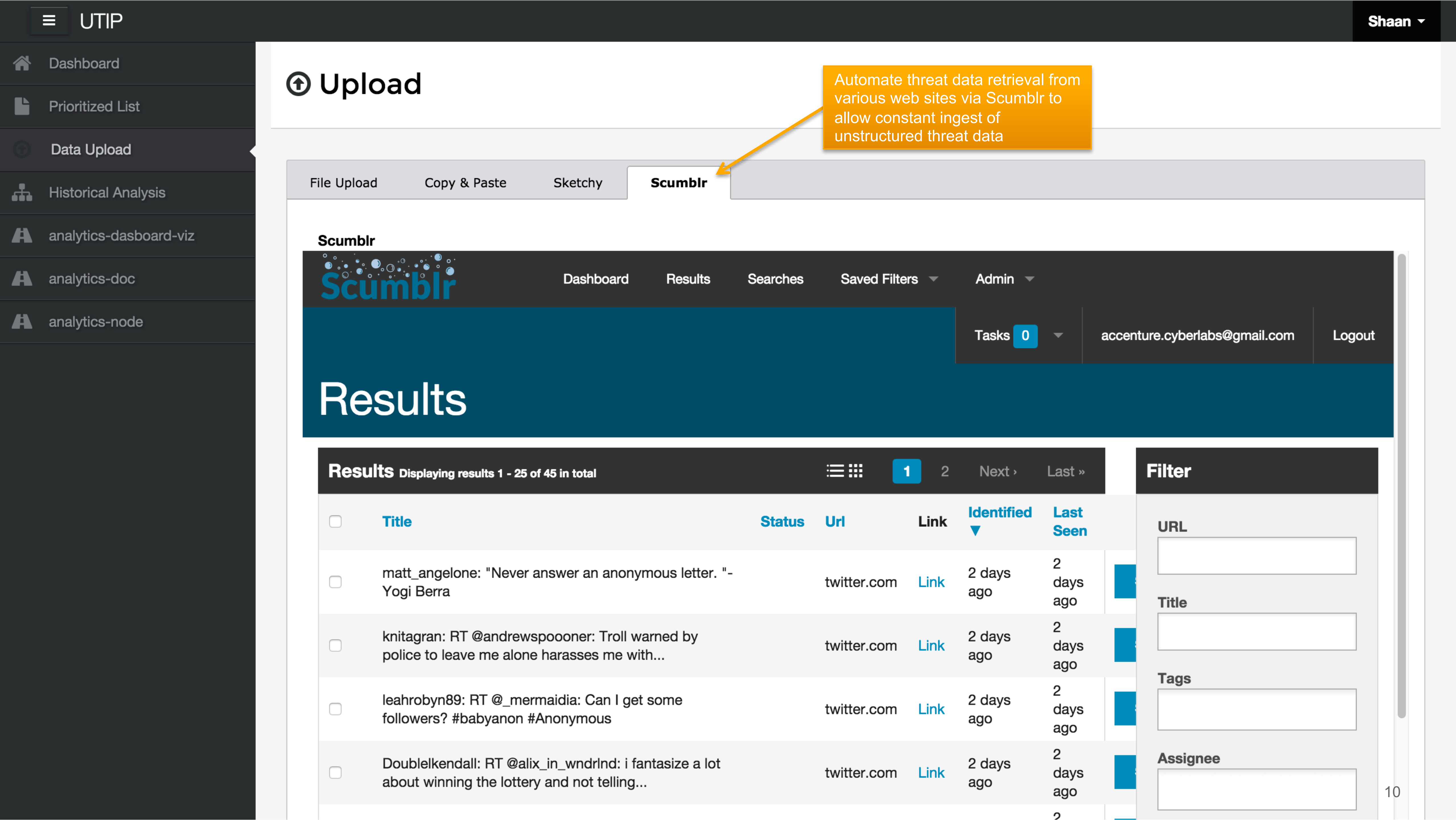
File Upload Tab

To upload a file, click Choose Files. Select files from the pop-up menu, or drag files from your computer on to the box. To upload multiple files at once, hold down the Shift or Control key as you select files.

+ Add files... | Start upload


UTIP
Shaan
Dashboard
Prioritized List
Data Upload
Historical Analysis
analytics-dasboard-viz
analytics-doc
analytics-node
Upload
Copy and paste any text directly into the field below to extract threat data into STIX constructs
File Upload
Copy & Paste
Sketchy
Scumblr
Copy & Paste Tab
Paste Text Directly Into Textbox
Insert file contents here
File name and extension
Submit
© 2014 Accenture


UTIP
Shaan
Dashboard
Prioritized List
Data Upload
Historical Analysis
analytics-dasboard-viz
analytics-doc
analytics-node
Upload
Insert any URL into the field to retrieve threat data in text format of the given URL
File Upload
Copy & Paste
Sketchy
Scumblr
Sketchy
Paste URL Here
Submit
© 2014 Accenture


UTIP
Shaan
Dashboard
Prioritized List
Data Upload
Historical Analysis
analytics-dasboard-viz
analytics-doc
analytics-node
Upload
Automate threat data retrieval from various web sites via Scumblr to allow constant ingest of unstructured threat data
File Upload
Copy & Paste
Sketchy
Scumblr
Scumblr
Scumblr
Dashboard
Results
Searches
Saved Filters
Admin
Tasks 0
accenture.cyberlabs@gmail.com
Logout
Results
Results Displaying results 1 - 25 of 45 in total
1
2
Next ›
Last »
Title
Status
Url
Link
Identified
Last Seen
matt_angelone: "Never answer an anonymous letter. "- Yogi Berra
twitter.com
Link
2 days ago
2 days ago
knitagran: RT @andrewspoooner: Troll warned by police to leave me alone harasses me with...
twitter.com
Link
2 days ago
2 days ago
leahrobyn89: RT @_mermaidia: Can I get some followers? #babyanon #Anonymous
twitter.com
Link
2 days ago
2 days ago
Doublelkendall: RT @alix_in_wndrlnd: i fantasize a lot about winning the lottery and not telling...
twitter.com
Link
2 days ago
2 days ago
Filter
URL
Title
Tags
Assignee

☰ UTIP | Shaan

- Dashboard
- Prioritized List
- Data Upload
- Historical Analysis
- analytics-dasboard-viz
- analytics-doc
- analytics-node

Dashboard

The dashboard provides a view of extracted threat data from documents mapped to STIX data constructs. Each individual piece of threat data is counted to determine the overall threat's trend within the specified date range.

2015/07/01 - 2015/07/2 GO

TTP	
The following updates ar...	5
Successful exploitation ...	3
• In a web-based attack ...	3
• The vulnerability canno...	2
A remote attacker could ...	2
The vulnerabilities are ca...	2
An attacker could exploi...	2
An attacker could exploi...	1
An attacker who succes...	1
This fixes a vulnerability,...	1

COA	
[Update to version 8.0.1,...	5
An update is available fo...	4
The following updates ar...	4
[Updated packages are ...	3
[When considering softw...	2
[Customer must upgrad...	2
[Update to version 7.0.1...	2
any necessary updates	1
an update for rpm	1
an update for java7 in h...	1

IOC	
cve-2014-4452	3
cve-2014-6363	3
cve-2014-6511	3
cve-2014-3511	3
http://helpx.adobe.com/...	2
https://rhn.redhat.com/e...	2
cve-2014-6504	1
cve-2014-6531	1
http://www.secunia.com...	1
cve-2014-3509	1

TARGET	
[Windows Server 2003 S...	4
The vulnerabilities are ca...	2
which	2
the way bind	2

THREATACTOR	
malicious people	3
malicious, local users	1
the legislation	1
remote attackers	1

CAMPAIGNS	
[HP]	4
[IBM]	3
[Microsoft]	3
[Apple]	3

UTIP

Shaan

- Dashboard
- Prioritized List
- Data Upload
- Historical Analysis
- analytics-dasboard-viz
- analytics-doc
- analytics-node

Prioritized List

This page shows a prioritized list of threat documents that were uploaded based on their overall relevance score. Relevance is determined by the score of the

Uploaded Files

July 18, 2015 - July 24, 2015

10 **records per page**

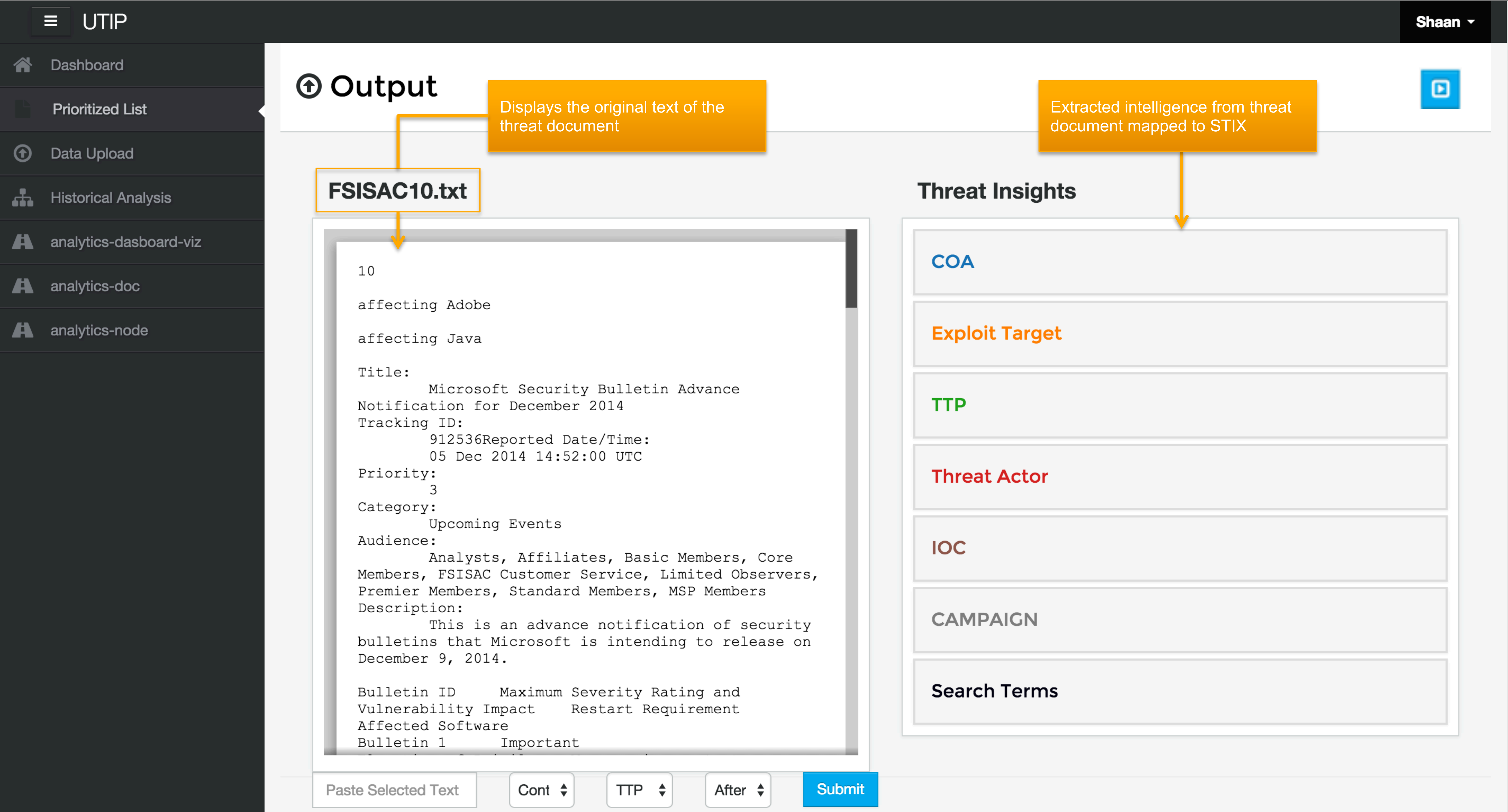
Search: Search

Score	Name	Type	Uploaded	Uploaded By	Status	Actions
550.0	FSISAC10.txt	TXT	July 22, 2015, 10:10 p.m.	Shaan	Processed	
550.0	Shaan Test 2.txt	TXT	July 23, 2015, 7:46 p.m.	Shaan	Processed	
235.0	FTest.txt	TXT	July 20, 2015, 11:10 p.m.	Shaan	Processed	

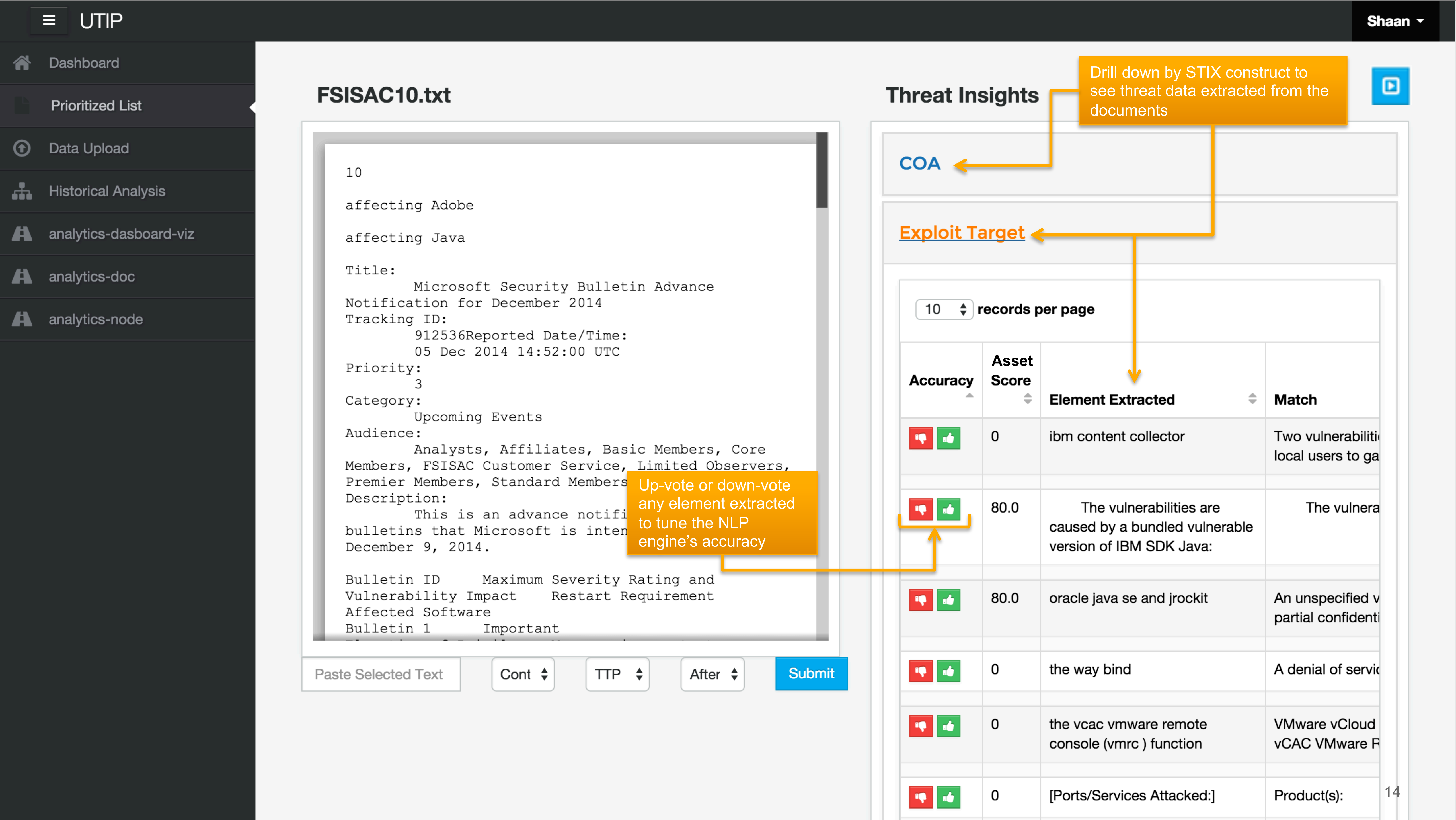
Showing 1 to 3 of 3 entries

← Previous 1 Next →

Additional actions can be taken on each ingested threat document. (e.g. show details or display visualizations of analytics on data from the documents)



UTIP
Shaan
Dashboard
Prioritized List
Data Upload
Historical Analysis
analytics-dasboard-viz
analytics-doc
analytics-node
Output
Displays the original text of the threat document
Extracted intelligence from threat document mapped to STIX
FSISAC10.txt
Threat Insights
10
affecting Adobe
affecting Java
Title:
Microsoft Security Bulletin Advance Notification for December 2014
Tracking ID:
912536Reported Date/Time:
05 Dec 2014 14:52:00 UTC
Priority:
3
Category:
Upcoming Events
Audience:
Analysts, Affiliates, Basic Members, Core Members, FSISAC Customer Service, Limited Observers, Premier Members, Standard Members, MSP Members
Description:
This is an advance notification of security bulletins that Microsoft is intending to release on December 9, 2014.
Bulletin ID Maximum Severity Rating and Vulnerability Impact Restart Requirement Affected Software
Bulletin 1 Important
COA
Exploit Target
TTP
Threat Actor
IOC
CAMPAIGN
Search Terms
Paste Selected Text
Cont
TTP
After
Submit


UTIP
Shaan
Dashboard
Prioritized List
Data Upload
Historical Analysis
analytics-dasboard-viz
analytics-doc
analytics-node
FSISAC10.txt
10
affecting Adobe
affecting Java
Title:
Microsoft Security Bulletin Advance Notification for December 2014
Tracking ID:
912536Reported Date/Time:
05 Dec 2014 14:52:00 UTC
Priority:
3
Category:
Upcoming Events
Audience:
Analysts, Affiliates, Basic Members, Core Members, FSISAC Customer Service, Limited Observers, Premier Members, Standard Members
Description:
This is an advance notifi
bulletins that Microsoft is inten
December 9, 2014.
Bulletin ID Maximum Severity Rating and Vulnerability Impact Restart Requirement Affected Software
Bulletin 1 Important
Paste Selected Text
Cont
TTP
After
Submit
Threat Insights
Drill down by STIX construct to see threat data extracted from the documents
COA
Exploit Target
10 records per page
Accuracy
Asset Score
Element Extracted
Match
0
ibm content collector
Two vulnerabiliti local users to ga
80.0
The vulnerabilities are caused by a bundled vulnerable version of IBM SDK Java:
The vulnera
80.0
oracle java se and jrockit
An unspecified v partial confidenti
0
the way bind
A denial of servic
0
the vcac vmware remote console (vmrc) function
VMware vCloud vCAC VMware R
0
[Ports/Services Attacked:]
Product(s):
Up-vote or down-vote any element extracted to tune the NLP engine's accuracy

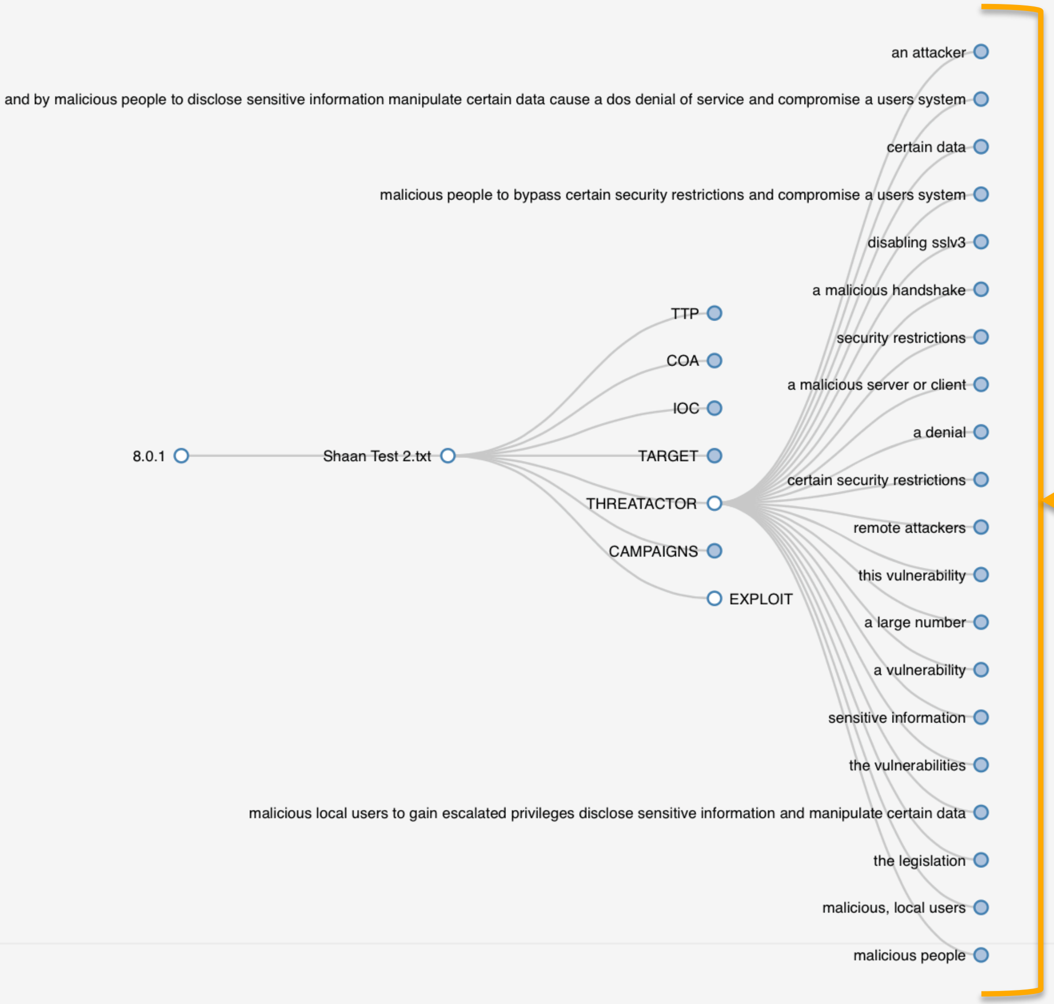
UTIP

Shaan

- Dashboard
- Prioritized List
- Data Upload
- Historical Analysis
- analytics-dasboard-viz
- analytics-doc
- analytics-node

Node Viz

2015/07/23 - 2015/07/24 GO



A custom tree node visualization is generated for each document and shows the elements extracted, their categories under STIX and all other documents that contain a related element