# About me

Zainab is a Data Scientist at TruSTAR.

She builds Machine Learning models to augment core services in the security platform and loves bringing the latest and greatest technologies to her work at TruSTAR.

Prior to this, Zainab received her Masters in Data Science from University of San Francisco.

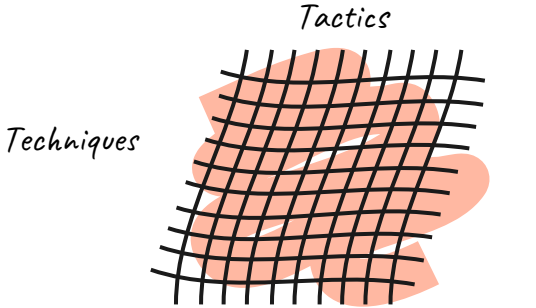
# Making Sense of Unstructured Data

An NLP approach

# Purpose

To establish links between:

MITRE Att&ck Framework





NIST Vulnerability Data

**CVE-2014-0622**

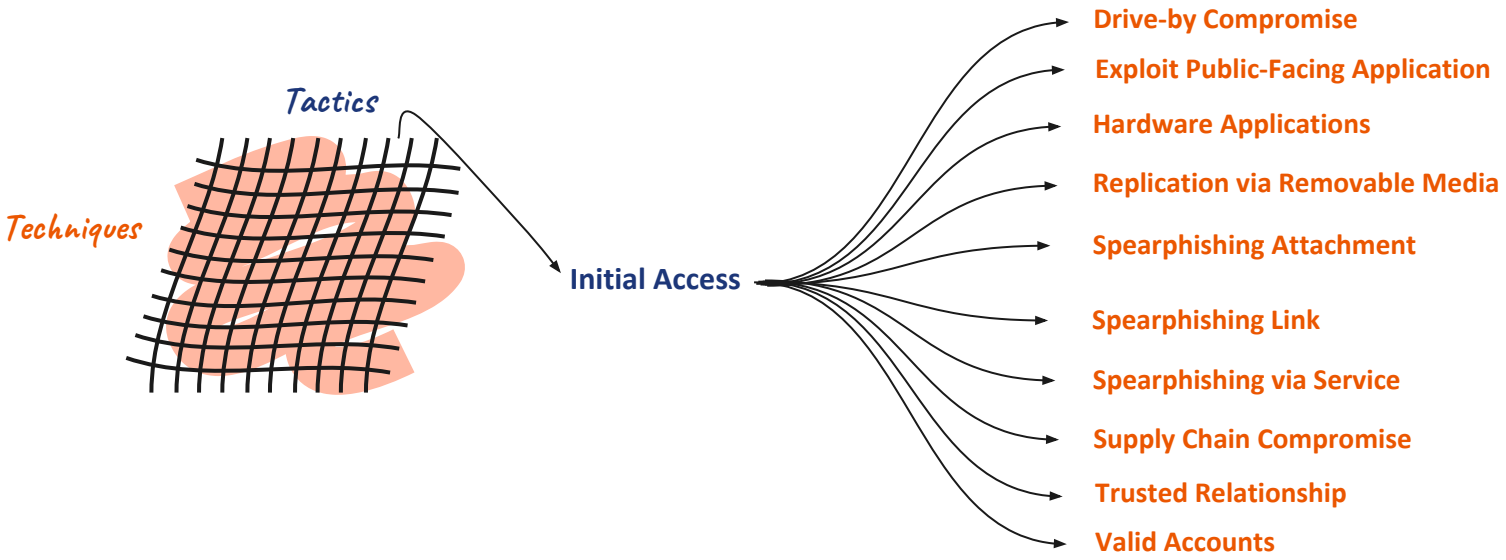
*Description: The web service in EMC Documentum Foundation Services does not properly which allows remote authenticated users to bypass intended content access restrictions via unspecified vectors. Foundation Services does not properly which allows remote authenticated users to bypass intended content access restrictions via unspecified vectors.*

*The web service in EMC Documentum Foundation Services does not properly which allows remote authenticated users to bypass intended content access restrictions via unspecified vectors. Foundation Services does not properly which allows remote authenticated users to bypass intended content access restrictions via unspecified vectors.*

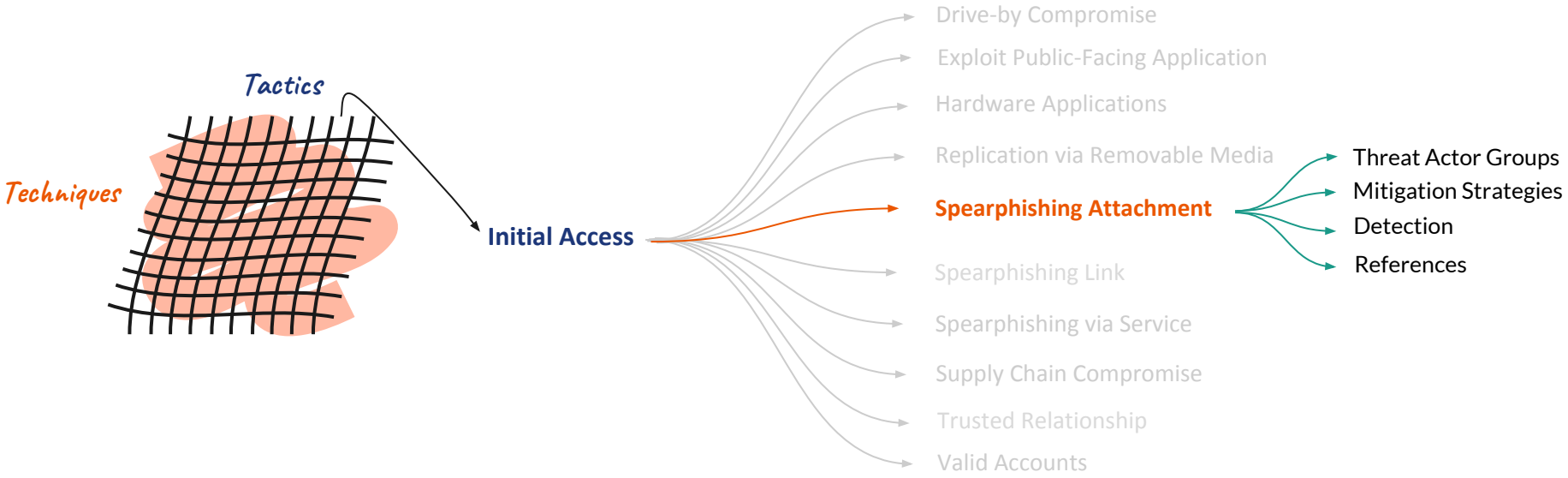
# MITRE ATT&CK Framework

- Tactic = Why?
- Technique = How?

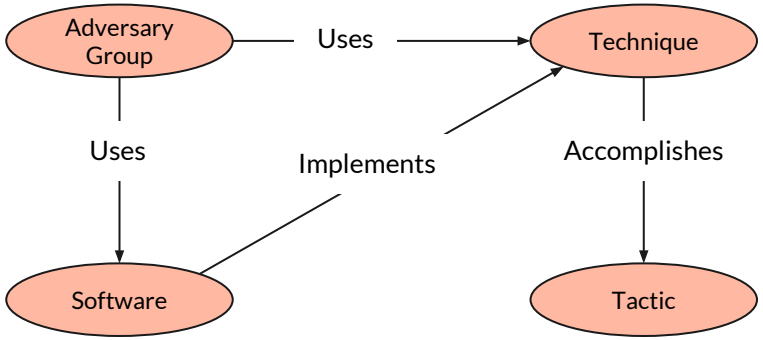
# MITRE ATT&CK Framework



# MITRE ATT&CK Framework

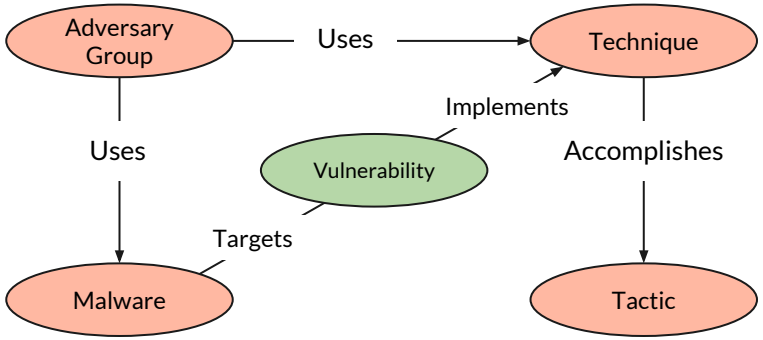


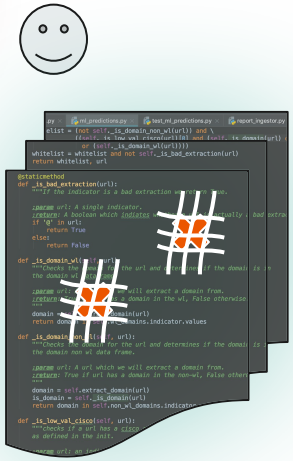
# Method to the madness?




**Att&ck Objects**

# Method to the madness?



# Fun NLP Stuff

# Word2Vec

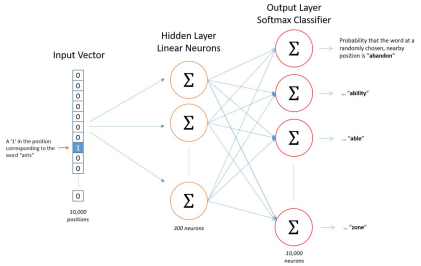
**What?** NLP technique that seeks to teach the computer to understand, interpret and manipulate human language.

**Why?** Translate words into vectors for mathematical manipulation.

**How?** By leveraging context and calculating probabilities.

| I | see | ants | on | the | tree. ⇨

(ants, I)
(ants, see)
(ants, on)
(ants, the)




ants = $[a_1\ a_2\ a_3\ a_4\ a_5 \ldots\ldots\ldots\ldots\ a_d]$

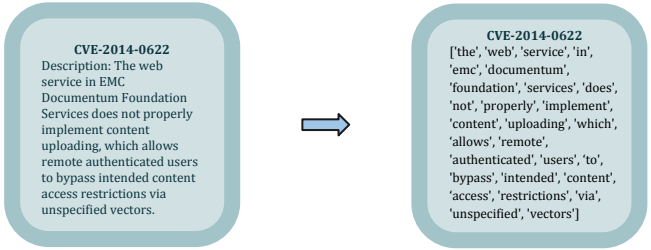
Le and Mikolov - Efficient Estimation of Word Representations in Vector Space (https://arxiv.org/abs/1301.3781)

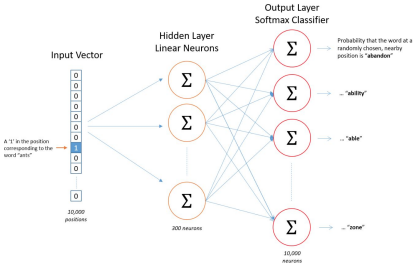
# Doc2Vec

**What?** NLP technique that seeks to teach the computer to understand, interpret and manipulate human language.

**Why?** Translate words and documents into vectors for mathematical manipulation.

**How?** By leveraging context and calculating probabilities.

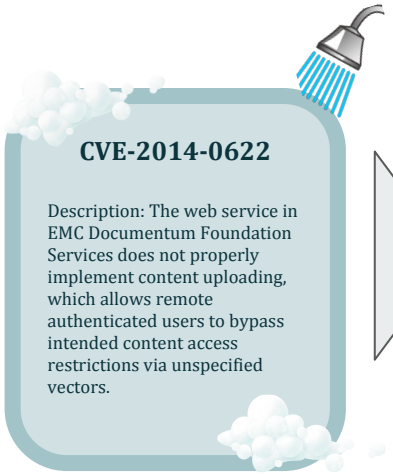


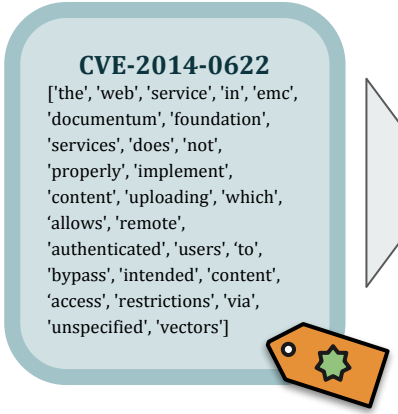

CVE-2014-0622 = $[a_1\ a_2\ a_3\ a_4\ a_5 \ldots\ldots\ldots\ a_d]$

Le and Mikolov - Distributed Representations of Sentences and Documents (https://arxiv.org/abs/1405.4053)

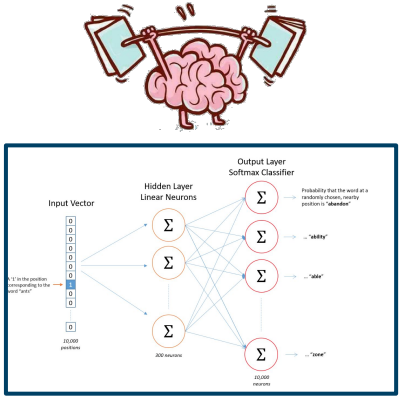
# Process




Data Cleaning




Tokenization




Model Training

CVE-2014-0622 $[a_1\ a_2\ a_3\ a_4\ a_5 \ldots\ldots\ldots\ a_d]$
CVE-2015-0765 $[a_1\ a_2\ a_3\ a_4\ a_5 \ldots\ldots\ldots\ a_d]$
...
attack-pattern12 $[a_1\ a_2\ a_3\ a_4\ a_5 \ldots\ldots\ldots\ a_d]$

Numeric Vectors

# Ok cool, then?

# Interesting CVE Clusters

**Initial number of docs:**

~100,000

**Total discovered clusters:**

~100

Browser Vulnerabilities

Overflow Vulnerabilities

Privilege Escalation Vulnerabilities

# Interesting CVE Clusters: DLL Vulnerabilities

An uncontrolled search path element (DLL Hijacking) vulnerability has been identified in Fuji Electric Energy Savings Estimator versions V.1.0.2.0 and prior. Exploitation of this vulnerability could give an attacker access to the system with the same level of privilege as the application that utilizes the malicious DLL.

An Uncontrolled Search Path Element issue was discovered in Moxa SoftNVR-IA Live Viewer, Version 3.30.3122 and prior versions. An uncontrolled search path element (DLL Hijacking) vulnerability has been identified.

A uncontrolled search path element issue was discovered in Vyaire Medical CareFusion Upgrade Utility used with Windows XP systems, Versions 2.0.2.2 and prior versions. A successful exploit of this vulnerability requires the local user to install a crafted DLL on the target machine.

An Uncontrolled Search Path Element issue was discovered in Advantech WebAccess versions prior to V8.2_20170817. A maliciously crafted dll file placed earlier in the search path may allow an attacker to execute code within the context of the application.

# Bringing two worlds together

**Dylib Hijacking**

macOS and OS X use a common method to look for required dynamic libraries (dylib) to load into a program based on search paths. Adversaries can take advantage of ambiguous paths to plant dylibs to gain privilege escalation or persistence. A common method is to see what dylibs an application uses, then plant a malicious version with the same name higher up in the search path. This typically results in the dylib being in the same folder as the application itself. If the program is configured to run at a higher privilege level than the current user, then when the dylib is loaded into the application, the dylib will also run at that elevated level. This can be used by adversaries as a privilege escalation technique.

# Bringing two worlds together

**Dylib Hijacking**

macOS and OS X use a common method to look for required dynamic libraries (dylib) to load into a program based on search paths. Adversaries can take advantage of ambiguous paths to plant dylibs to gain privilege escalation or persistence. A common method is to see what dylibs an application uses, then plant a malicious version with the same name higher up in the search path. This typically results in the dylib being in the same folder as the application itself. If the program is configured to run at a higher privilege level than the current user, then when the dylib is loaded into the application, the dylib will also run at that elevated level. This can be used by adversaries as a privilege escalation technique.

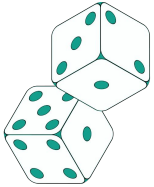
**CVE-2017-6329**

Symantec VIP Access for Desktop prior to 2.2.4 can be susceptible to a DLL Pre-Loading vulnerability. These types of issues occur when an application looks to call a DLL for execution and an attacker provides a malicious DLL to use instead. Depending on how the application is configured, the application will generally follow a specific search path to locate the DLL. The exploitation of the vulnerability manifests as a simple file write (or potentially an over-write) which results in a foreign executable running under the context of the application.

# Yeah but I need evidence

Github: https://github.com/zdanish1/trustar-daenerys

# Evaluation



100 naturally forming clusters

1/100 chance of getting it right at random

1% accurate associations

50% accurate associations

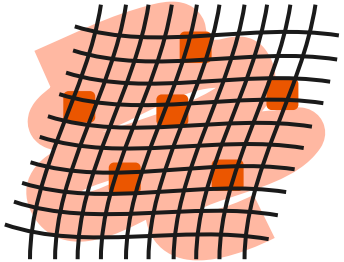
# Show me the money

# Organizational Prioritization

MITRE Att&ck Framework

*Tactics*

*Techniques*



**CVE-2017-5161**

An issue was discovered in Sielco Sistemi Winlog Lite SCADA Software. An uncontrolled search path element (**DLL Hijacking**) vulnerability has been identified. Exploitation of this vulnerability could give an attacker access to the system with the same level of privilege as the application that utilizes the malicious **DLL**.

- **Sielco Sistemi Winlog Lite SCADA**
- **Fuji Electric Energy Savings Estimator v1.0.0**
- **BLF-Tech LLC VisualView HMI v9.9.14.0**
- **SIMPlight SCADA Software version 4.3.0.27**
- **Moxa SoftNVR-IA Live Viewer v3.30.3122**

Weak posture

Strong posture

# Thank you