GraphIt

Ansh Chandnani, Joshua Ramos, Summer Beasley, Andrew Shoffler, Jeannette Kube, & Zhaoxiong Li
College of Computing and Informatics, Drexel University

What's GraphIt?

A Cybersecurity search engine that automates Open Source Intelligence and creating a query-driven knowledge graph.

Background

- Open Source Intelligence (OSINT) is the process of finding *publicly* available information on a target to gain insights.
- Unfortunately, this process is time consuming, cumbersome, and feels like following a trail of breadcrumbs.

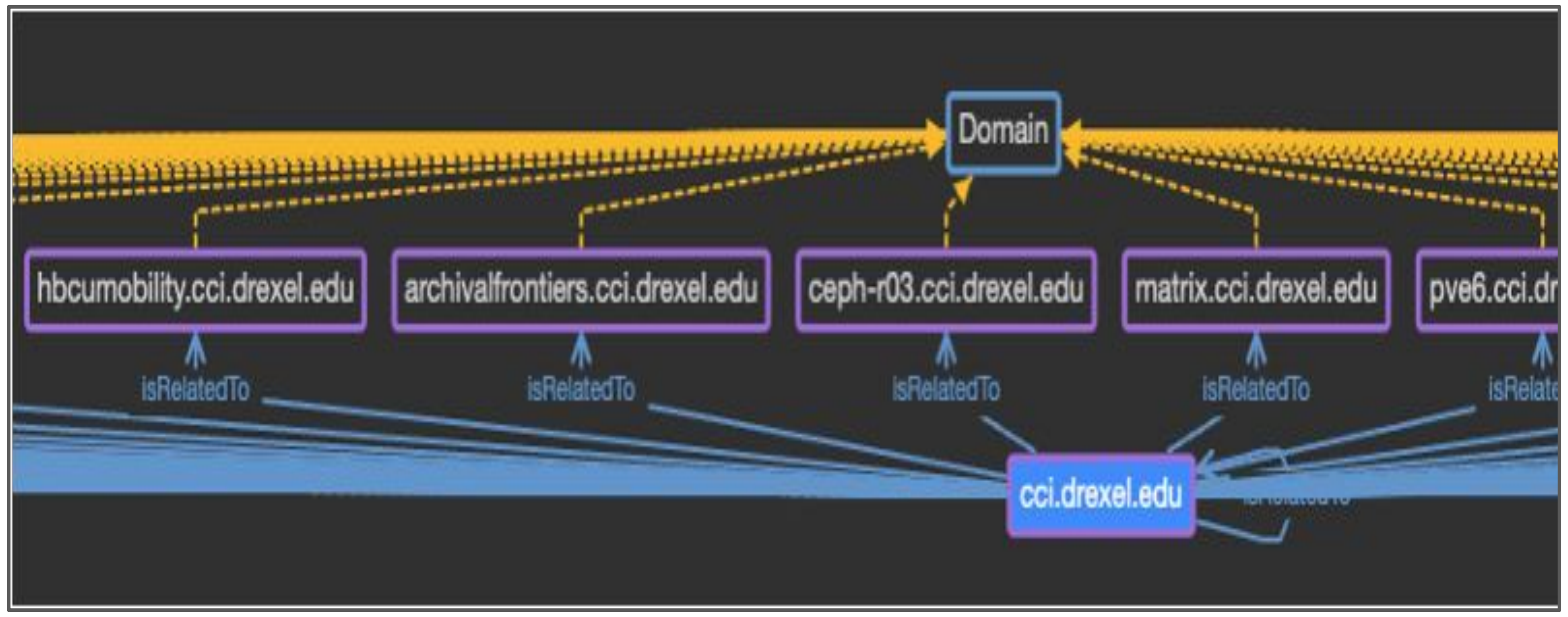


Fig 1: Snippet of 114 domains related to cci.drexel.edu.

Terms Used:

- CVE (Common Vulnerabilities and Exposures) - Unique identifier for each vulnerability.
- CWE (Common Weakness Enumeration) - category system used to group CVE's on their weakness type.
- OSINT (Open Source Intelligence) - Process of finding *publicly* available information on a target to gain insights about a target.

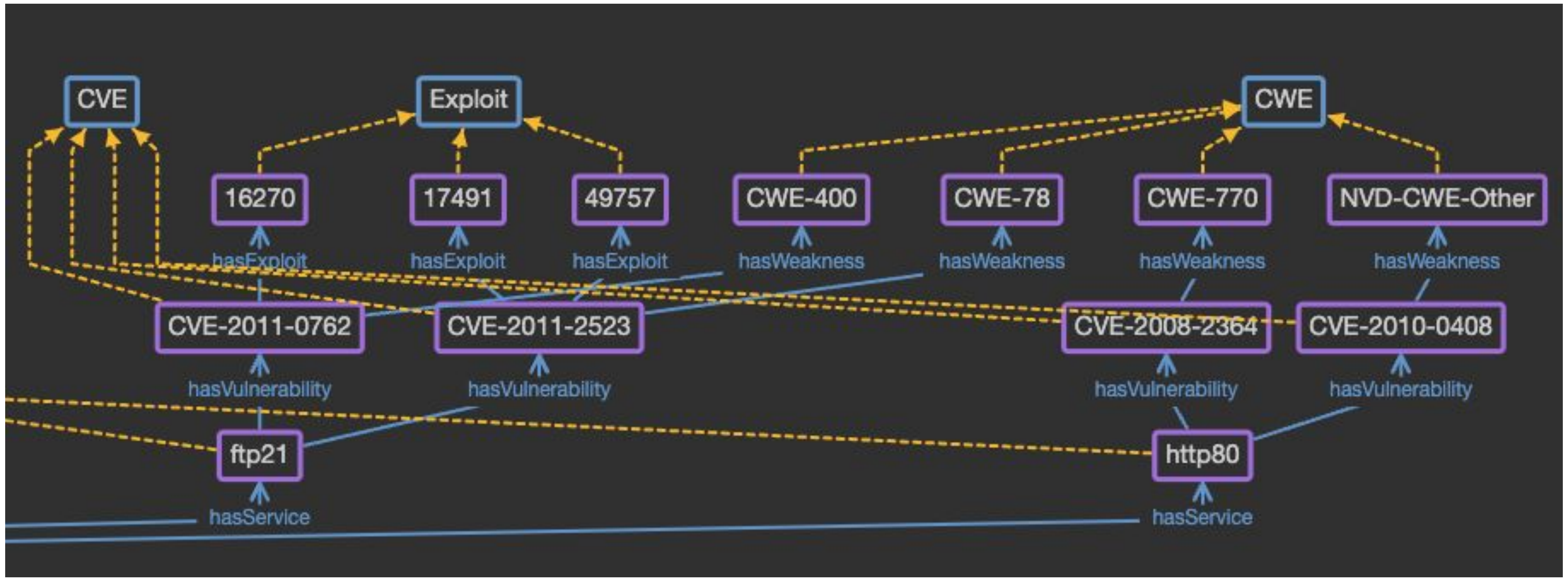
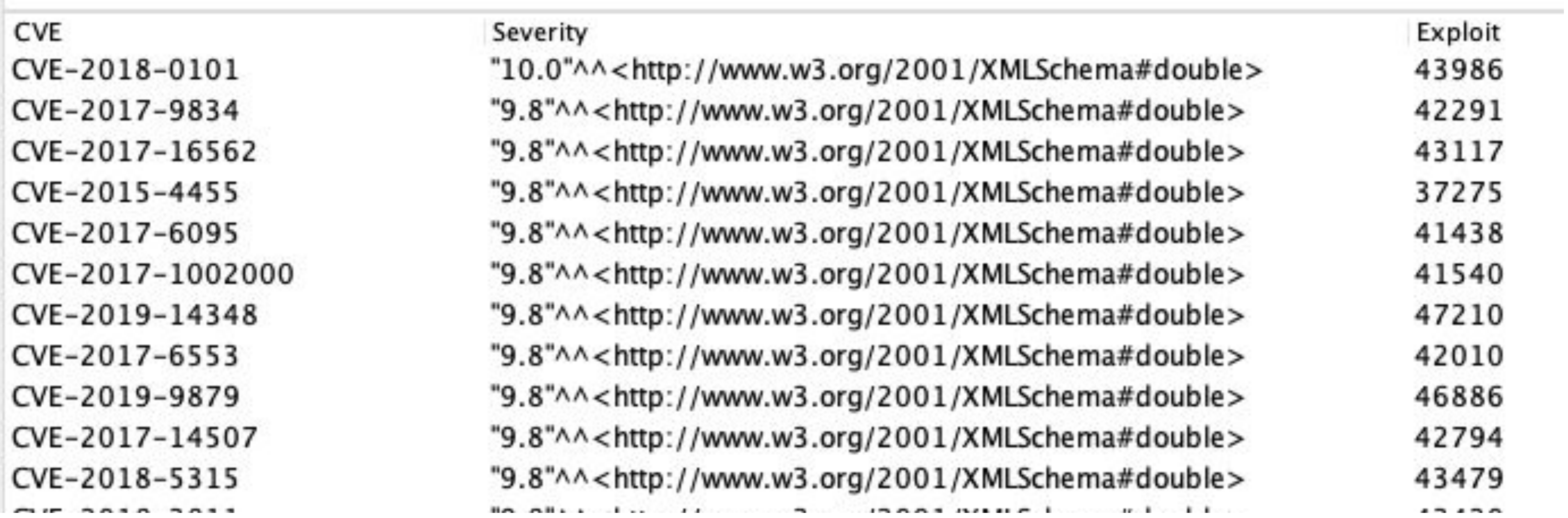


Fig 2: Semantic relationship between vulnerabilities, exploits, and weaknesses found on a vulnerable system.

```
Query Eg:
SELECT ?CVE ?Severity
?Exploit WHERE {
  ?CVE ns:hasExploit
?Exploit
  ?CVE ns:hasBaseScore
?Severity
} ORDER BY DESC(?Severity)
```



CVE	Severity	Exploit
CVE-2018-0101	"10.0"^^<http://www.w3.org/2001/XMLSchema#double>	43986
CVE-2017-9834	"9.8"^^<http://www.w3.org/2001/XMLSchema#double>	42291
CVE-2017-16562	"9.8"^^<http://www.w3.org/2001/XMLSchema#double>	43117
CVE-2015-4455	"9.8"^^<http://www.w3.org/2001/XMLSchema#double>	37275
CVE-2017-6095	"9.8"^^<http://www.w3.org/2001/XMLSchema#double>	41438
CVE-2017-1002000	"9.8"^^<http://www.w3.org/2001/XMLSchema#double>	41540
CVE-2019-14348	"9.8"^^<http://www.w3.org/2001/XMLSchema#double>	47210
CVE-2017-6553	"9.8"^^<http://www.w3.org/2001/XMLSchema#double>	42010
CVE-2019-9879	"9.8"^^<http://www.w3.org/2001/XMLSchema#double>	46886
CVE-2017-14507	"9.8"^^<http://www.w3.org/2001/XMLSchema#double>	42794
CVE-2018-5315	"9.8"^^<http://www.w3.org/2001/XMLSchema#double>	43479

Fig 3: Snippet from about 2,500 results related to Wordpress.

More Information and Research

https://tinyurl.com/AnshResearch

Features

- Parse NMap Port Scans to automate searching for related domains, vulnerabilities, and exploits.
- Search for vulnerabilities and exploits based on keywords.
- Find related domains and internal URLs.
- Ask questions to the knowledge graph, sort, and filter using SPARQL queries.

How Does It Work?

- Aggregates valuable information from various trusted public data sources like Certificate Transparency Logs, the National Vulnerability Database (NVD), and ExploitDB.
- Semantically organizes information and their relations in an Ontology.
- Users may begin their query with a keyword, CVE, Domain, or NMap scan.

Future Work

- Integrate more data sources into the ontology, i.e. Shodan, MITRE, and Large Language Models (LLMs).
- Build a unified front end allowing users to query and pan the graph simultaneously.