InTIME: Cyber-Threat Intelligence Extraction & Sharing

- An entirely Open Source framework for Cyber-Threat Intelligence (CTI) Extraction & Sharing
  - manages the complete CTI lifecycle
  - enhances security preparedness & awareness

- Machine learning-based web data gathering
  - efficient data discovery (e.g., 0-day vulnerabilities, exploits) from the clear, social, and dark web (e.g., forums, marketplaces, security-related websites, pastebins)

- Language model-based content ranking to assess its usefulness
  - statistical language modelling to represent content in low-dimensional space
  - ranks crawled content according to its potential to be useful

InTIME: Cyber-Threat Intelligence Extraction & Sharing

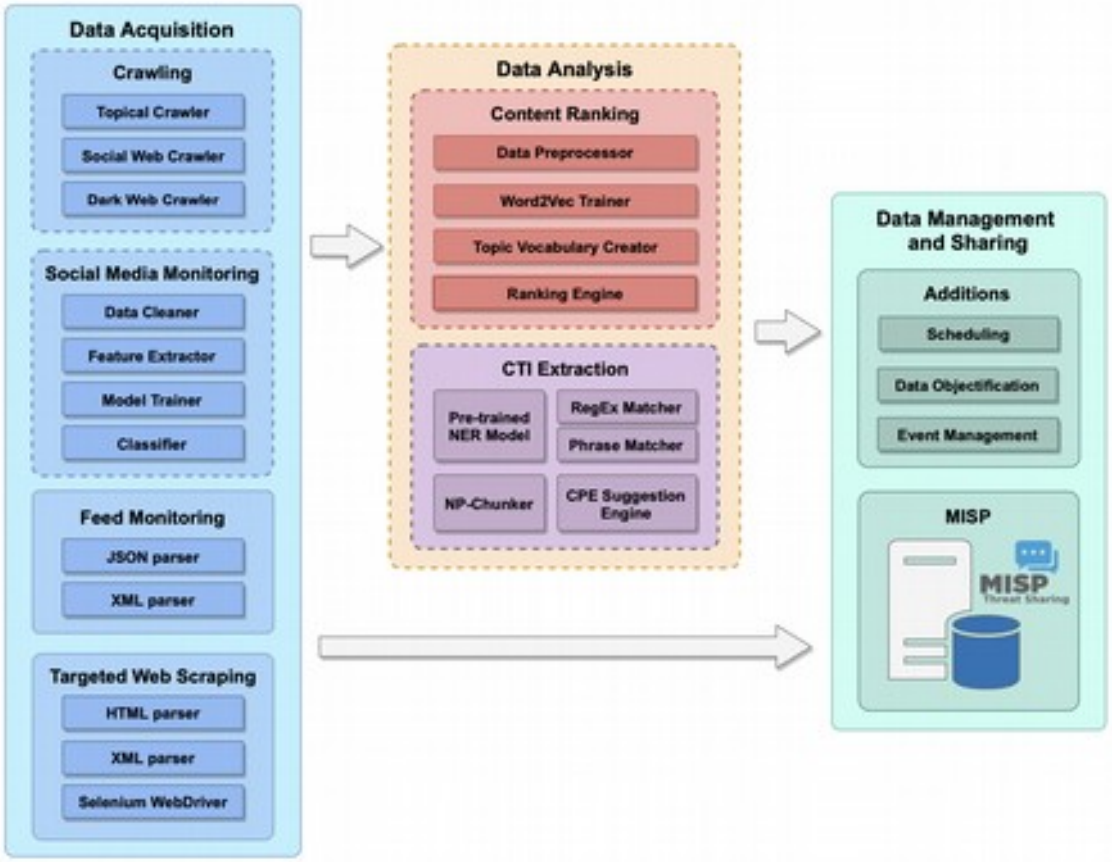
- Natural language understanding and named entity recognition/ disambiguation
  - extracts CTI from crawled free text
  - domain specific entities
  - dependency parsing for more accurate extraction
  - specialised tools for semi-automated linking to known platform/ vulnerability naming schemes (e.g., CPEs. CVEs)

- Scalable CTI management and sharing
  - extends MISP platform
  - supports CTI storage, consolidation, exploration, inspection and visualisation
  - supports CTI sharing in both human and machine-based formats

InTIME architecture

Three main modules:

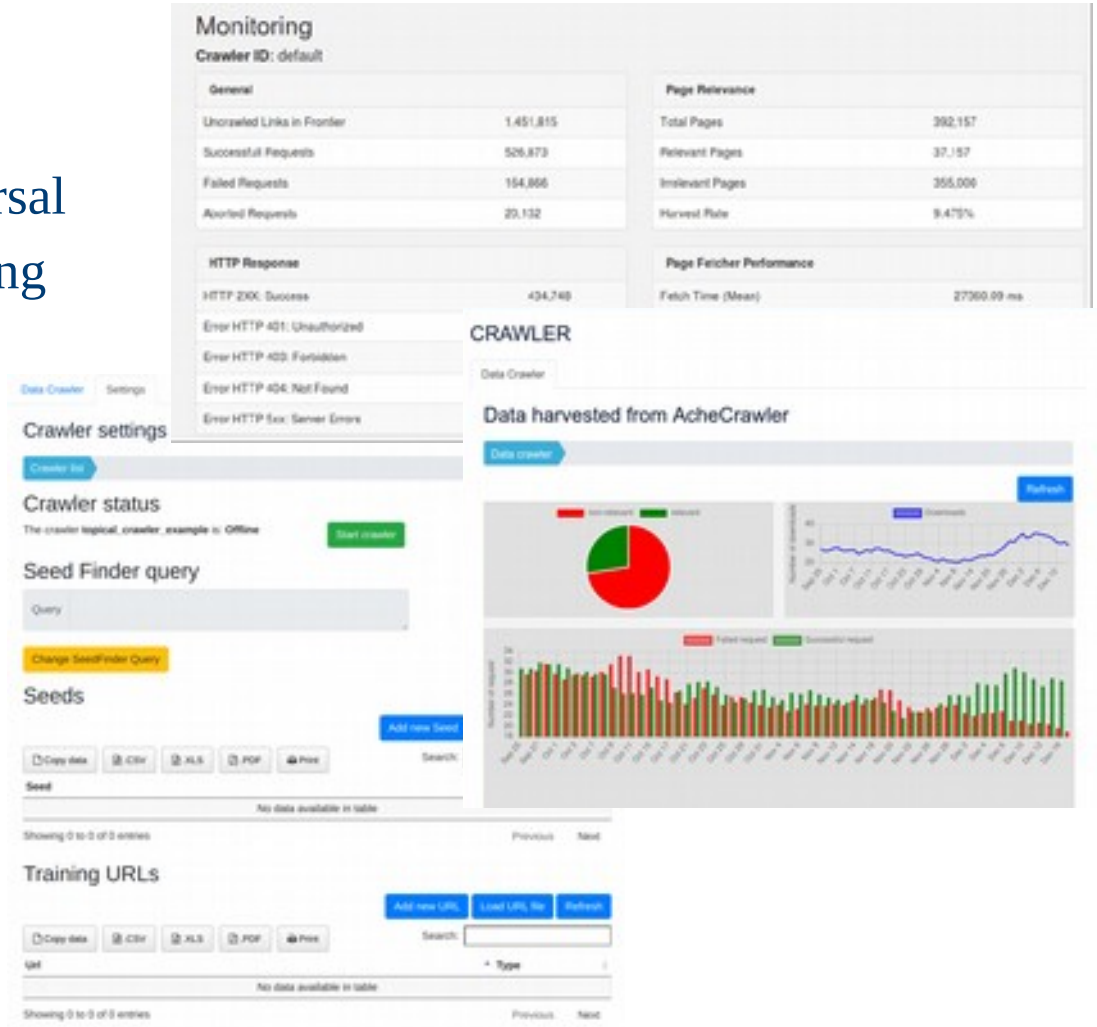
- data acquisition – content harvesting (crawlers/scrappers)
- data analysis
  - content ranking
  - CTI extraction
- data management & sharing



Content harvesting

A family of crawlers/scrappers

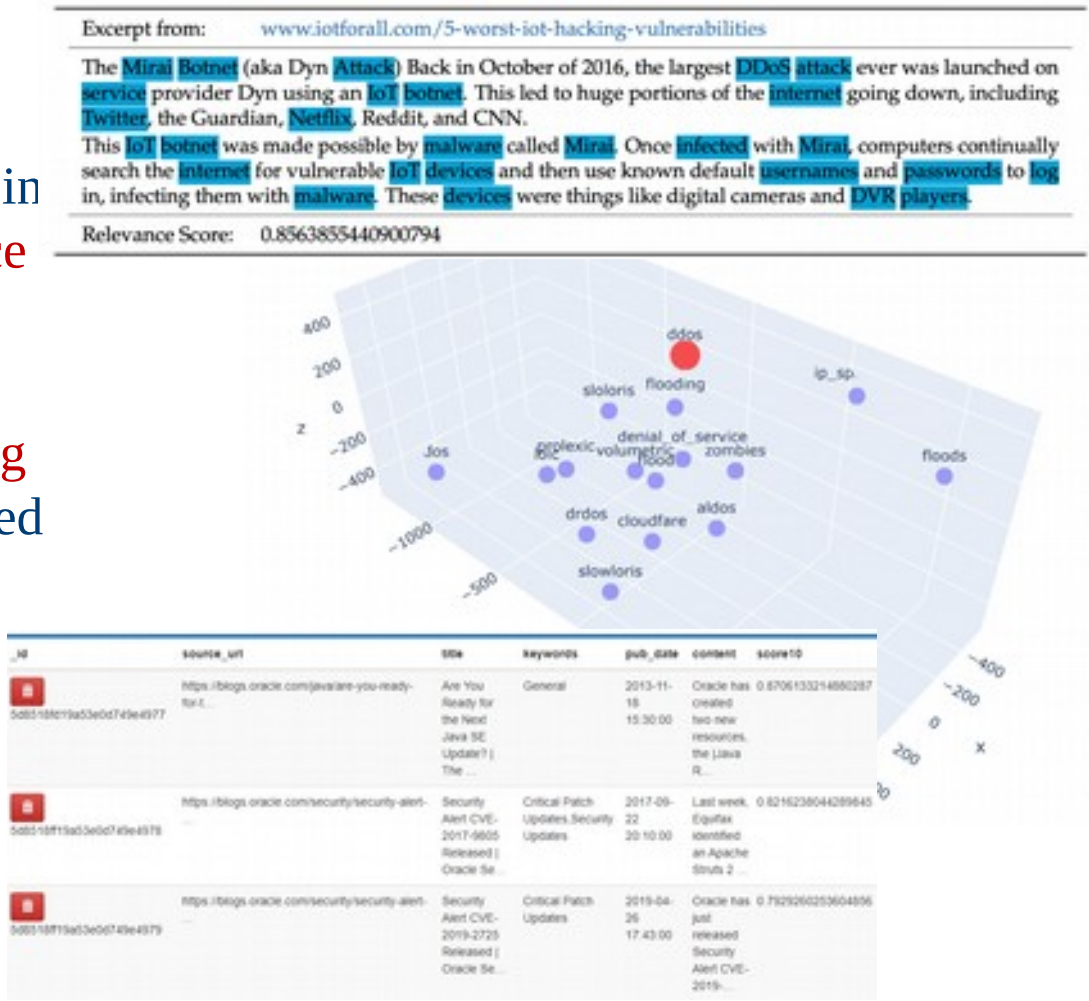
- tunable & automated clear/ deep/social/dark web traversal
- thematic & in-depth crawling
- machine-learning based (SVMs/random forests direct the crawls)
- regex-support (structured domains)
- Integration with public VDBs (JVN, NVD, kbcert, vuldb, exploit db)



Content ranking

A novel ranking/classification module:

- represent harvested content in latent low-dimensional space (adaptive to vocabulary changes)
- statistical language modeling to assess usefulness harvested content (captures salient concepts)
- machine-learning based (word embeddings/2-layer NNs & CNNs)
- noSQL data store



CTI extraction

Information extraction module:

- natural language understanding (machine-learning & rule based)
- named entity recognition & disambiguation (e.g., CPE suggestion)
- introduced CTI-specific entities (e.g., shell_cmd, version, file, cpe/cve/cwe) alongside standard ones (e.g., org, date, ordinal)
- REST API w/ extracted CTI (JSON objects w/ entities & metadata)

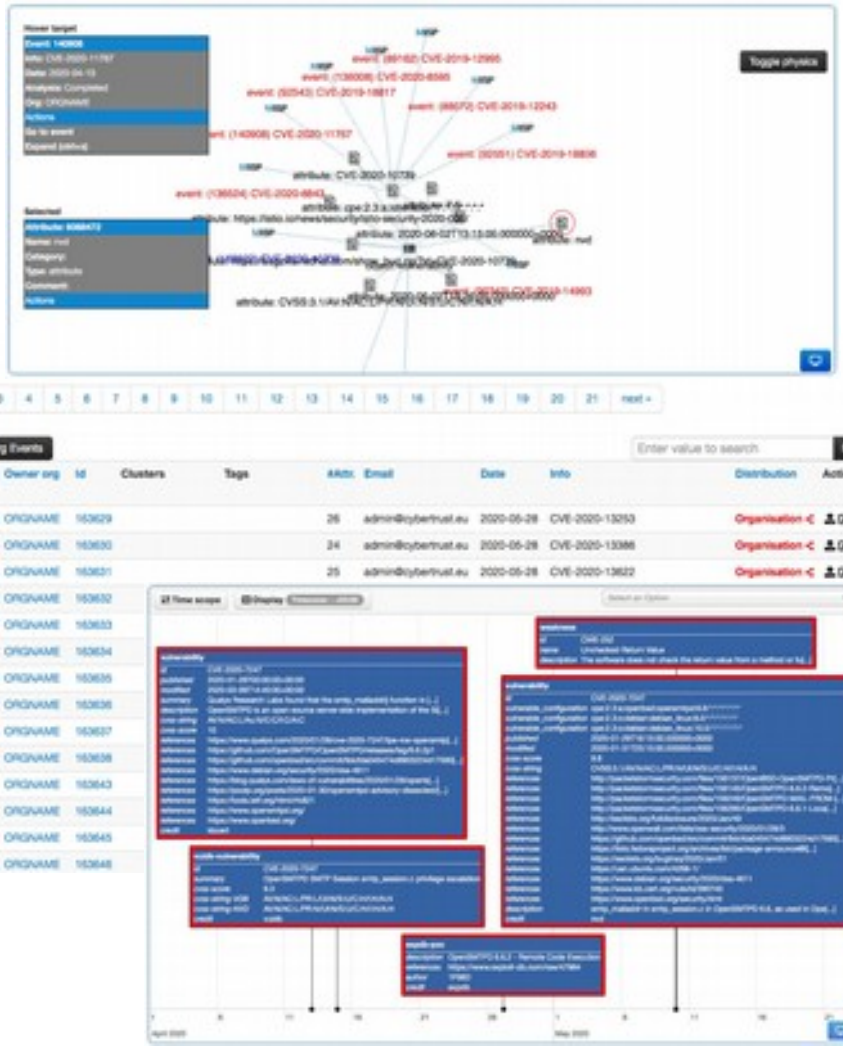

```
{
    "docId": 915240,
    "rawText": "Symantec Web Security Group (WSG) products
    "matches": [
        {
            "text": "Symantec Web Security",
            "entity": "PRODUCT",
            "start": 0,
            "end": 3
        },
        {
            "text": "WSG",
            "entity": "ORG",
            "start": 5,
            "end": 6
        },
        {
            "text": "OpenSSL",
            "entity": "PRODUCT",
            "start": 12,
            "end": 13
```


[redacted] Group (WSG ORG) products using affected versions of [redacted] may be susceptible to multiple vulnerabilities. A local or remote attacker can obtain private key or other secret key information. A remote attacker can also cause denial of service. [redacted] is exploitable in ASG (versions 6.7 VERSION , 7.1 VERSION , 7.2 VERSION), CA (versions 2.3 VERSION , 2.4 VERSION , 3.0 VERSION), [redacted] (cpe:/a:bluecoat:proxysg CPE : 6.5.8.7 VERSION), [redacted] (versions 10.4 VERSION , 10.5 VERSION), and SSLV ORG (versions 4.5 VERSION , 5.0 VERSION) only when customers configure the products' SSL/TLS ORG interfaces with 1024-bit RSA ORG keys.

CTI management & sharing

CTI management & sharing module:

- efficient CTI storage & indexing
- easy-to-use CTI consolidation & correlation
- intuitive CTI inspection & visualisation
- supports CTI sharing in both human and machine-based formats
- extends MISP platform



InTime & Open Source

- InTime has been built using solely open source tools
- Very positive experience from the Open Source Ecosystem
  - Good documentation
    - Clear installation instructions, descriptive APIs
    - More details could be provided for DB design
  - Good scalability
    - The MISP system scaled up well to hundreds of thousands of entries
    - The storage and management of correlations was suboptimal, leading to performance bottlenecks and manageability issues (e.g. transferring of DB)
      - A more efficient database design could serve well here
- InTime is in the process of being published as open source
  - Code cleanup and documentation improvements


Merci
Grazie
Chokrane
Kiitos
Spacibo
Obrigado
Asante
Multumesc
Dank Je
Thank You
Matondo
Juspaxar
Thank You
Arigato
Nuwun
Asante
Kiitos
Grazie
Chokrane
Dank Je Maake Kiitos
Juspaxar
Mochchakkeram
Merci
Maake
Dank Je
Spacibo
Grazie