ISACA®
Sistemi informativi: averne fiducia e trarne valore
Rome Chapter

# *La piattaforma cybersecurity Pragmema*
# *POC: architettura e servizi*

***Elisabetta Zuanelli***
*Pragmema srl*

Elisabetta Zuanelli
Presidente CReSEC/Università degli Studi di Roma «Tor Vergata»
Coordinatore del Piano di formazione nazionale in *cybersecurity*, *cyberthreat*, *privacy* (www.cybersecurityprivacy.it)

Roma 22/01/2018

# Gli argomenti

- Cybersecurity, *data analytics*, AI
- Le ontologie e le tassonomie della cybersecurity
- La soluzione POC: *tool*/piattaforma automatica di intelligenza artificiale *cybersecurity*/protezione dati e servizi

# Quale *cybersecurity*: una delimitazione del campo

La sicurezza informatica è la condizione in cui il **ciberspazio è protetto**, rispetto a **eventi volontari** o accidentali consistenti **nell'acquisizione ed esfiltrazione di dati, nella loro modifica o distruzione illegale o il blocco di sistemi informatici**, grazie ad appropriati sistemi di sicurezza.

Queste misure includono **verifiche di sicurezza, gestione degli aggiornamenti o correzioni, procedure di autenticazione, gestione degli accessi, analisi dei rischi, individuazione e risposta a incidenti/attacchi, mitigazione degli impatti, recupero delle componenti soggette ad attacco, addestramento ed educazione del personale**, verifica e incremento della sicurezza fisica dei locali dove sono situati i sistemi di informazione e comunicazione.

# La funzione dei dati: masse di dati del contesto/corpora di dati d'attacco, minacce, vulnerabilità (eventi e incidenti)

**Finalità**: pronti interventi, mitigazione, resilienza, prevenzione, predittività

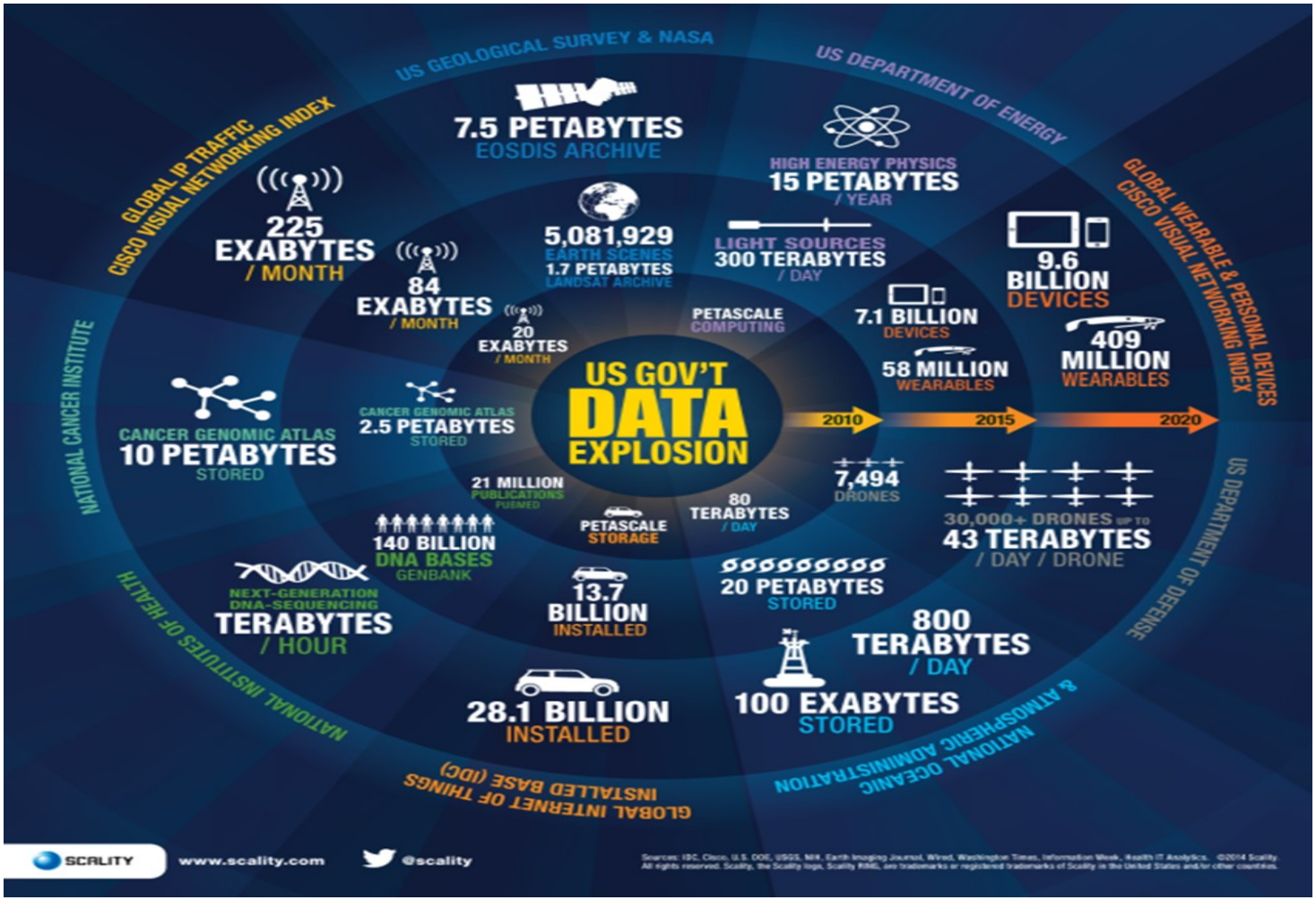
- Reportistica
- Metriche dei danni
- Statistiche
- Standard
- Certificazioni di sicurezza dei prodotti e dei servizi

**Percorsi d'azione:**

- L'acquisizione dei dati (detentori dei dati)
- Le tassonomie dei dati
- Le classificazioni dei dati

# ***Cybersecurity*, dati, *big data analytics***

# L'esplosione dei dati



## ***Big data analytics*: masse astrutturate (e strutturate) di dati da analizzare e strutturare (ri-strutturare)**

- Obiettivi: analisi e classificazione per scopi rimediali e preventivi nei diversi settori: militari, industriali, istituzionali, economico-finanziari, sanitari, ecc.
- Tecnologie disponibili per l'analisi: *libraries*, *big data software*, nuove logiche di *storage* ed elaborazione
- Limiti: carenza di modellizzazioni e applicazioni
- Utilità: destinazioni tecniche (CERT, CSIRT, SIEM, ecc.) e laiche
- Formalizzazione e gestione automatica dei dati

# Soluzioni

- Modellistica ontologica di utilità generale
- Definizione, correlazione e valore univoco dei dati
- Sviluppo tecnologico interoperabile di piattaforme di analisi, *assessment* e valutazione del rischio
- Implementazione additiva dell'architettura

# L’intelligenza artificiale: *knowledge,* informazioni, dati

# Le nuove professionalità

- Business Intelligence Developer, Web Developer, Software Developer, Business Developer, Analyst Developer, Applications Developer, Database Developer and Front-End Developer

- The top three skills required of big data developers are

NoSQL, Java and SQL

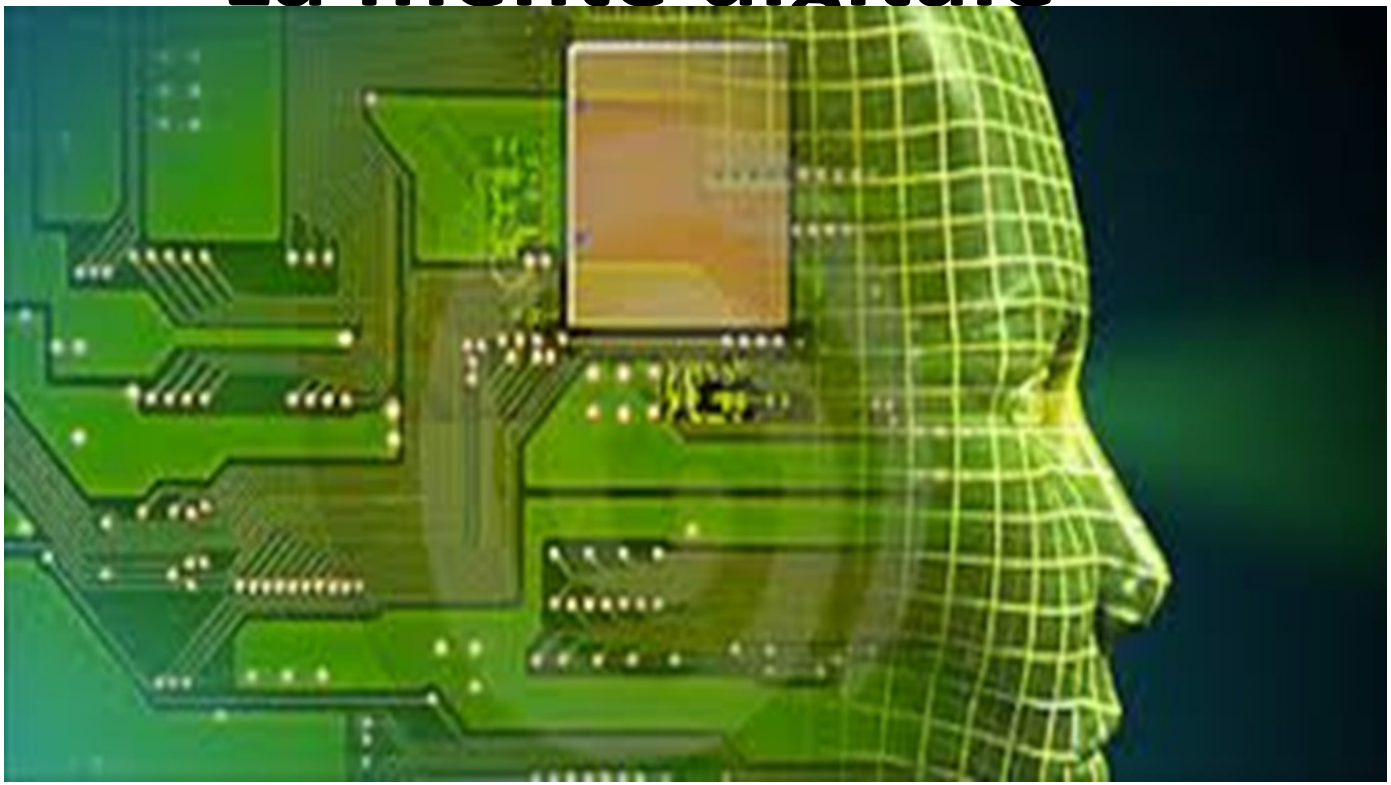
# Stima del mercato di AI

Crescita stimata del mercato AI al 2022: 16,06 miliardi di dollari. Top investitori Amazon e Google

## AI: emulazione del funzionamento della mente umana

- Immissione di informazioni sensoriali multiple dal mondo esterno
- Utilizzo di diverse tipologie di input sensoriale/modale
- Canalizzazioni sensoriali: specificità dei dati
- Interpretazione e modellizzazione dei dati
- Linguaggi ibridi
- Linguaggi traduttivi

# La mente digitale



# L’intelligenza artificiale e i dati

- Modellizzazione dei dati e delle relazioni logico-semantiche
- Definizione e sviluppo del modello
- Definizione dei metadati
- Linguaggi di metadati
- Formati di rappresentazione dei dati

# Ontologie e tassonomie: la modellistica

**Un riferimento: N. Guarino (ed.), *Formal Ontology in Information Systems*, IOS Press, Amsterdam, 1998**

Some twenty years ago Guarino postulated the increasing relevance of ontology in the fields of **Artificial Intelligence, Computational Linguistics and Database Theory** and mentioned specific research fields such as **knowledge engineering, knowledge representation, qualitative modelling, language engineering, database design, information modelling and integration, object oriented analysis, information retrieval and extraction, knowledge management and organization, agent-based systems design**.

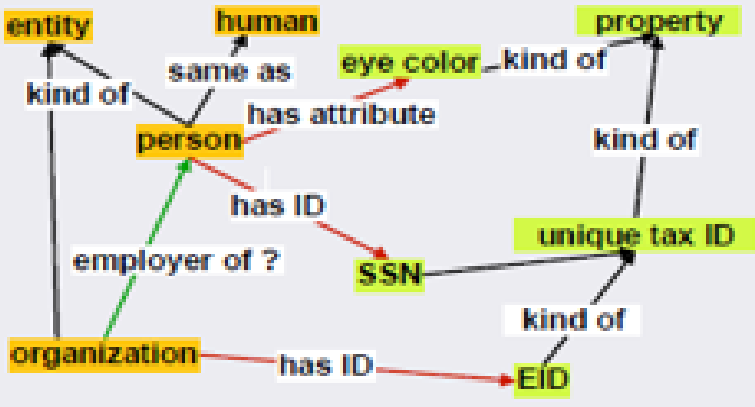
- At the methodological level he stressed the main peculiarity of an **ontology as its being a highly interdisciplinary approach where philosophy and linguistics play a fundamental role.**

# L'ontologia della vulnerabilità nella *cybersecurity* (NISTIR 8138 Draft 2016, Tocsa, Capec, ecc.)

- Elementi di conoscenza costitutivi della rappresentazione della **vulnerabilità** intesa come «debolezza nella logica computazionale dei prodotti e servizi che può essere sfruttata da una fonte di minaccia»
- Teatri remoti esterni: Internet (DNS *services*, *web browsers*, ecc.)
- Remoti limitati: cellulari, *wireless*, *bluetooth*, ecc.
- *Bug* degli applicativi interni (sistemi operativi, programmi installati)
- Localizzazione degli attacchi esterni: piattaforme ecommerce, social, cloud, WoT, mobile, ecc.
- Tipologia malware e finalità degli attacchi: per esfiltrazione dati, corruzione, controllo di sistema, ecc.
- Logiche di percorso: es. Kill Chain

# Esempio di ontologia, CV e interoperabilità semantica/MITRE

## Ontologies, Controlled Vocabularies and Semantic Interoperability

| | Controlled Vocabulary | Ontology |
|---|---|---|
| **Definition** | A controlled vocabulary (CV) is a set of lexical expressions that are vetted according to some criteria, such as their accepted usage in a community.<br>• CVs are structured by one or more ordering relations, such as "narrower-than," "broader-than," or "related-to."<br>• Structure is machine processable and semantics are human interpretable. | An ontology specifies the meaning of a controlled vocabulary in the form of a conceptual model.<br>• Ontologies can be independent of any given controlled vocabulary.<br>• Structure is machine processable and semantics are machine interpretable. |
| **Example** | **Terms** / **Relation**<br>entity — broader-than person, broader-than organiz.<br>> person — narrower-than entity<br>>> eye color — related-to person<br>>> SSN — related-to person<br>>> employer — related-to person<br>> organization — narrower-than entity<br>>> EID — related-to organization |  |

# Vocabolari controllati per gli standard

Gli standard richiedono vocabolari controllati CV a livello di contenuti e rappresentazione– Standard principali (NIST/MITRE)

- – CEE: Common Event Expression
- – CPE: Common Platform Enumeration
- – CRE: Common Remediation Enumeration
- – CVE: Common Vulnerability Enumeration
- – CWE: Common Weakness Enumeration
- – MAEC: Malware Attribute Enumeration and Characterization
- – OVAL: Open Vulnerability and Assessment Language
- – XCCDF: Extensible Configuration Checklist Description Format

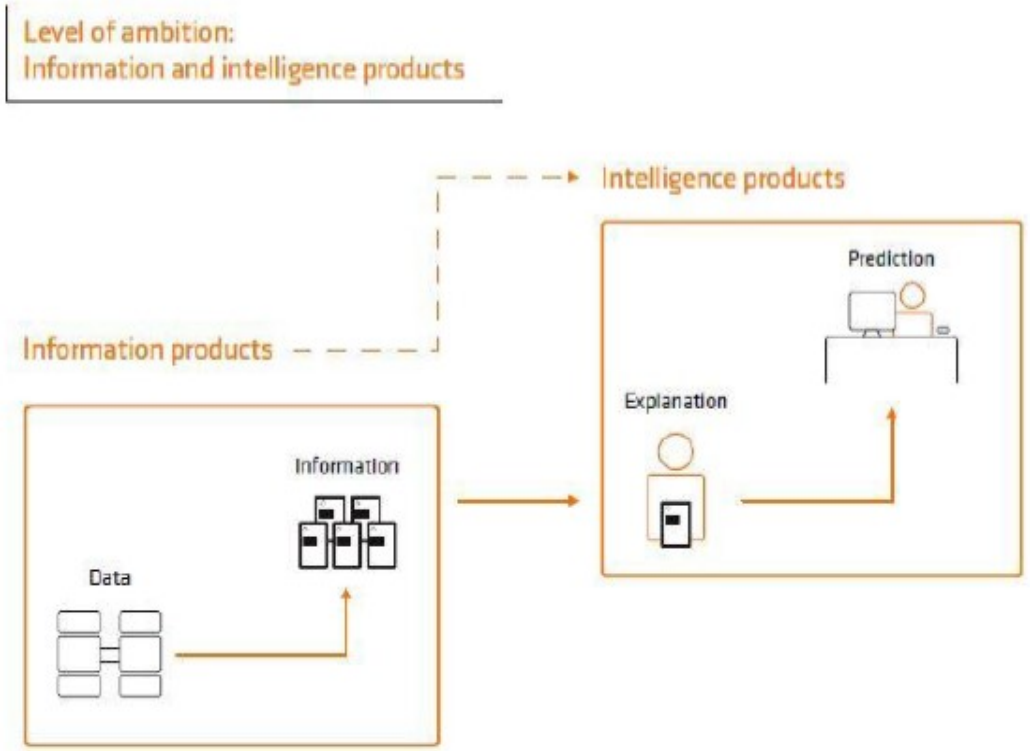
■ Both MITRE and NIST maintain public repositories and Web sites for

the various standards: http://nvd.nist.gov/
http://oval.mitre.org/repository/
http://measurablesecurity.mitre.org/

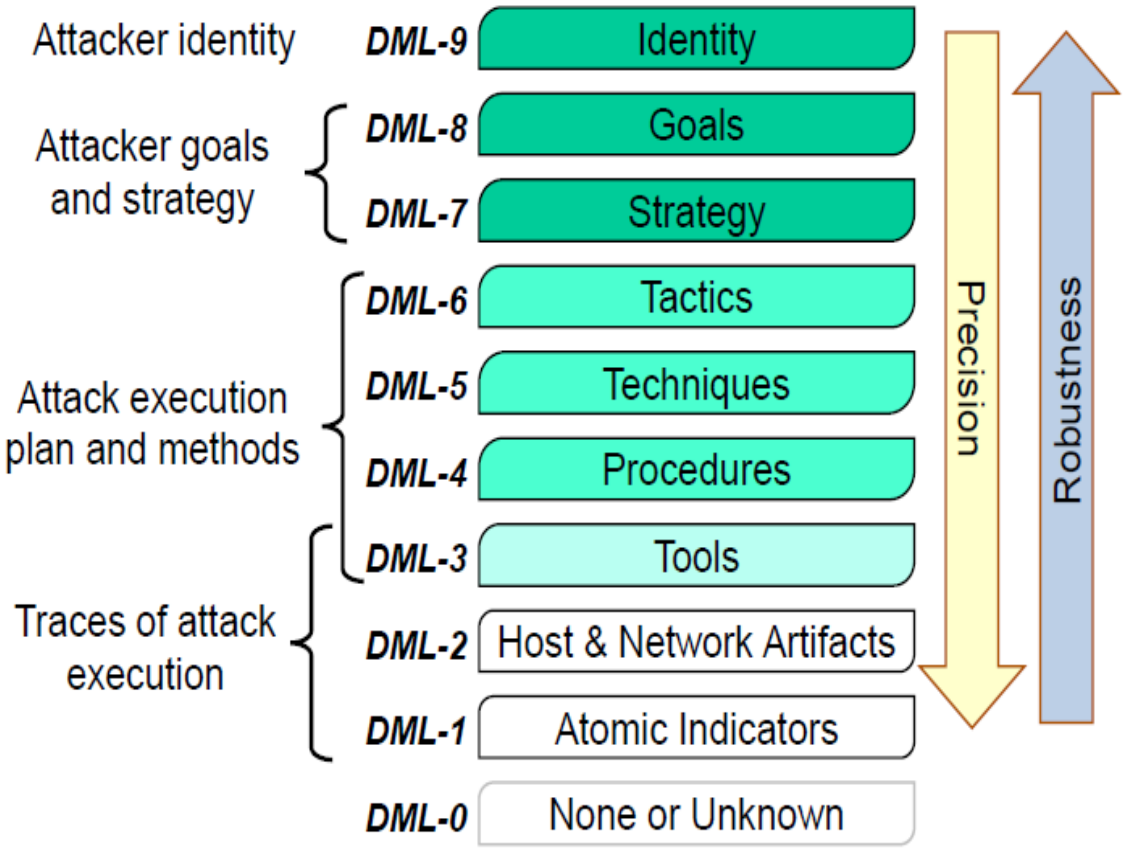
# ACT, TOCSA and Oslo Analytics (2017)

•Semi-Automated Cyber Threat Intelligence (ACT)
-Open Source Threat Intelligence Platform
-https://www.mnemonic.no/research-and-development/semi-automated-cyber-threat-intelligence/
•Threat Ontologies for Cyber Security Analytics (TOCSA)
-Ontologies
-PhD Project
-https://www.mnemonic.no/no/research-and-development/threat-ontologies-for-cybersecurity-analytics/
-http://www.mn.uio.no/ifi/english/research/projects/tocsa/
•Operable Subjective Logic Analysis Technology for Intelligence in Cybersecurity (Oslo Analytics)
-Analytics
-Subjective Logic (quantifying uncertainty)
-Trust Networks
-Academic
-http://www.mn.uio.no/ifi/english/research/projects/oslo-analytics/

# Threat Information vs Threat Intelligence

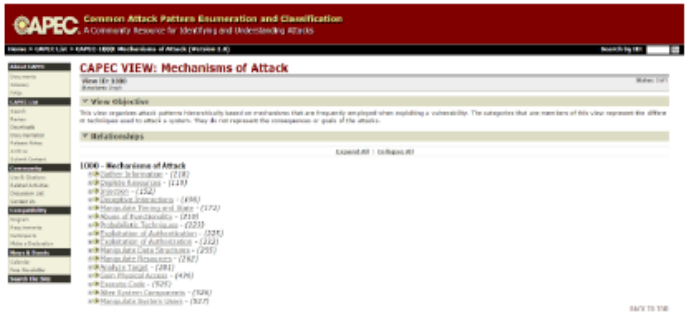


# The Detection Maturity Level (DML) Model



# Semantic Feature Extraction

- Formal definitions of
  - -Goals
  - -Strategy
  - -Tactics
  - -Techniques
  - -Procedures
- Relevant initiatives
  - -MITRE CAPEC
    - https://capec.mitre.org
  - -MITRE ATT&CK
    - https://attack.mitre.org
  - -MITRE CAR
    - https://car.mitre.org





mnemonic

## ENISA 2016
## Analisi comparativa tassonomie
## *cybersecurity*

**“There is currently no consensus on concepts and definitions related to taxonomies”.**

Attualmente le tassonomie esistenti

“lack terms to properly handle the impact of an incident, incidents with no malice intended, explicit fields for ransomware, whether the incident is confirmed, and the differentiation between intrusion attempts and intrusions”.

# Enisa Threat Taxonomy

Collecting

Sorting
Consolidating


Asset exposure


Vulnerability
exploitation

Figure 1: ENISA Threat Taxonomy and its use-cases

# La piattaforma ontologica per la cybersecurity POC Pragmema

- La piattaforma POC su infrastruttura Liferay, è la prima nel contesto europeo e internazionale

- POC integra la *Vulnerability ontology* del NIST, classificazioni ENISA, analisi e proposte ontologiche e classificatorie specifiche (TOCSA) riformulandone la valenza e l'impianto strutturale e ampliandone la portata conoscitiva e applicativa

- Necessaria per l'analisi e la classificazione dei dati e la descrizione univoca degli eventi e incidenti

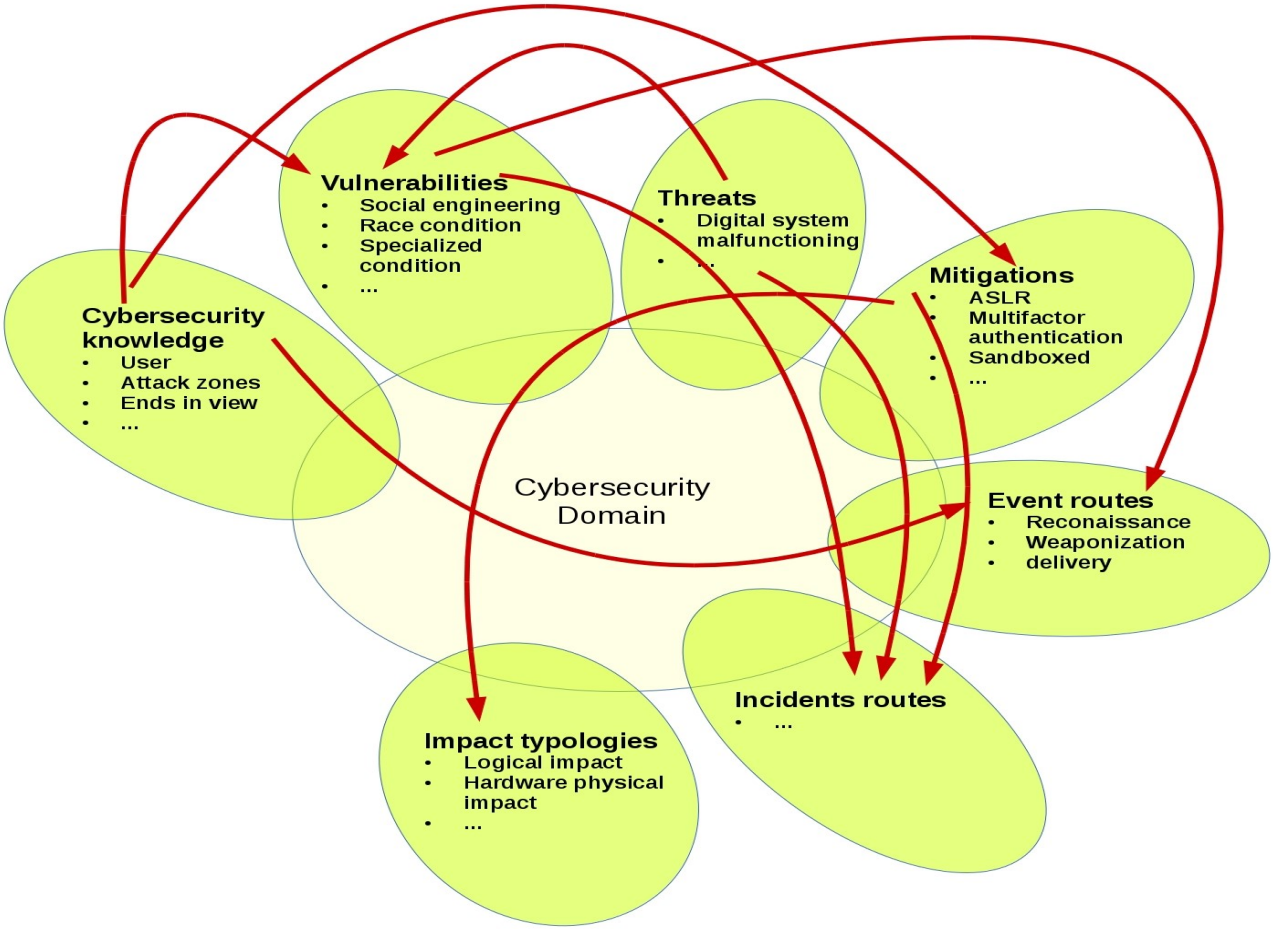
## The POC solution

- La prima piattaforma tecnologica in ambito internazionale per la gestione della *cybersecurity* e la protezione dei dati
- Di immediato interesse applicativo nel contesto normativo di ***compliance*** per la sicurezza nella Direttiva NIS e nel GDPR
- Customizzabile per clienti diversi nel pubblico e nel privato
- Con funzioni di interoperabilità

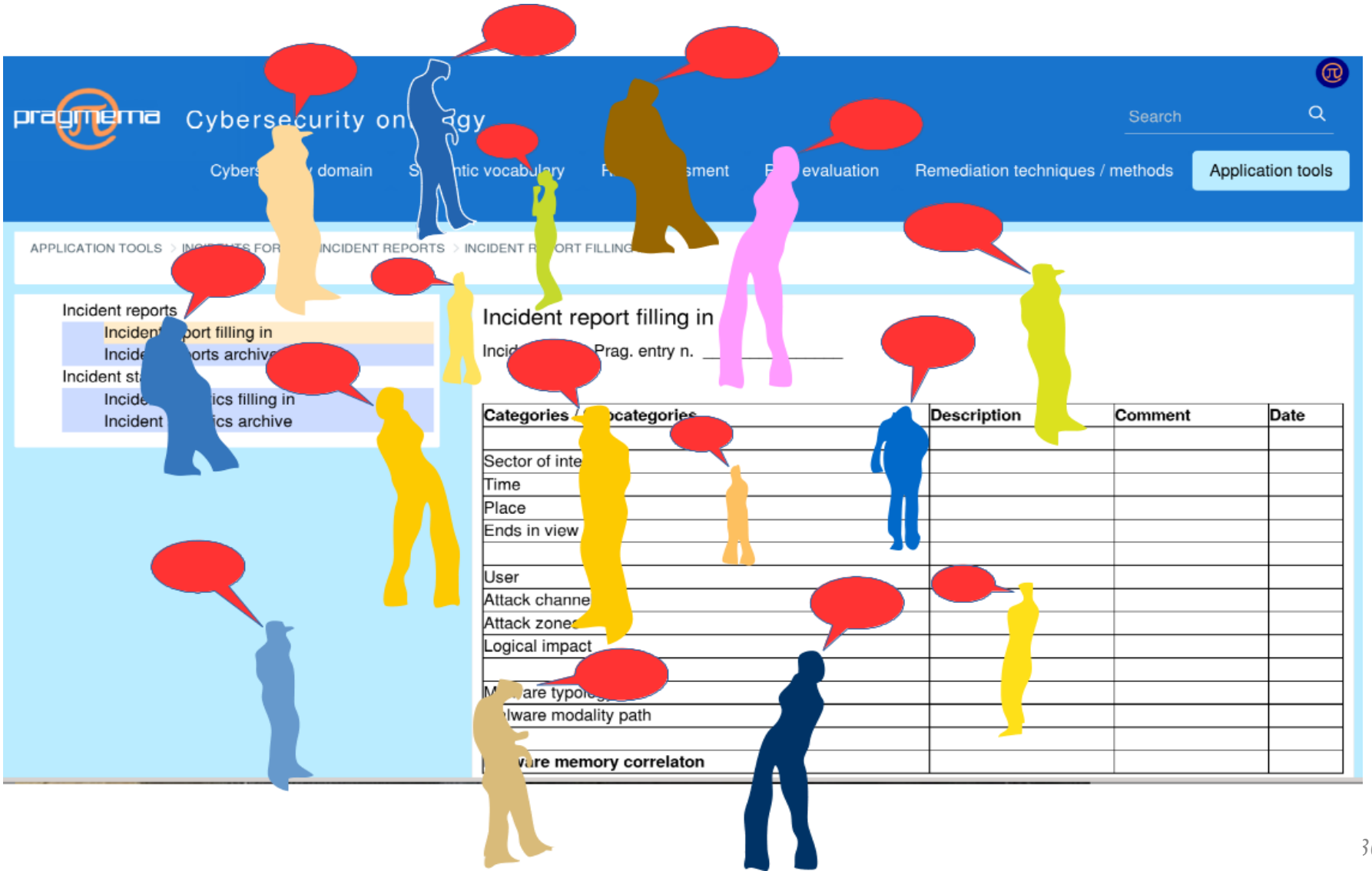
## Criteri definitori ontologici

- Entità
- Attributi
- Proprietà
- Relazioni logico-semantiche
- Vocabolari controllati
- Traduzione tecnologica

# La rete di relazioni su entità di primo livello



# Uno strumento per normalizzare i big data

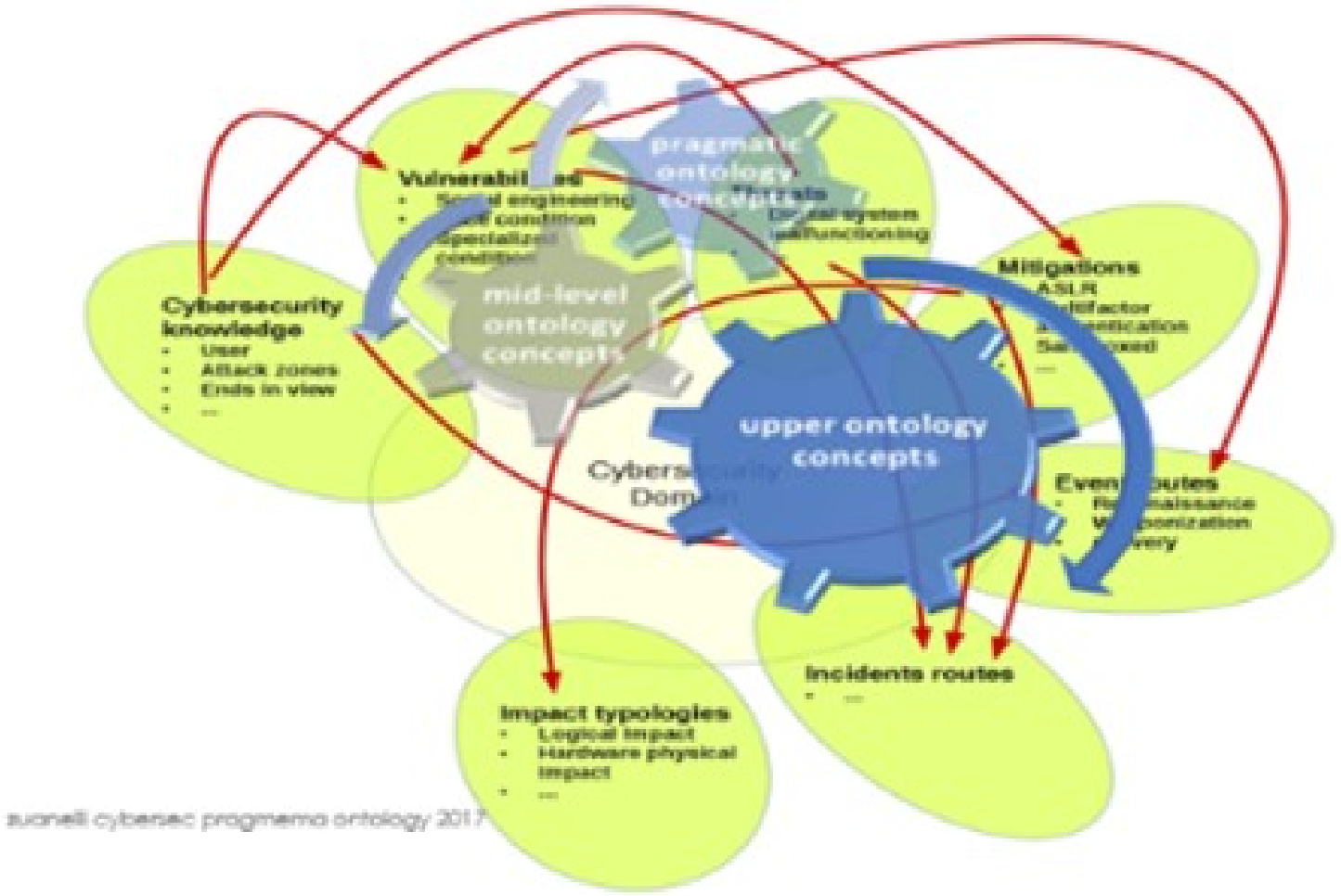


# La struttura

- POC dispone di due livelli corrispondenti a due oggetti ontologici/tecnologici: l'**ontologia di dominio cybersecurity** e l'**ontologia pragmatica di servizi** alla cybersecurity

- I due oggetti sono **commerciabili separatamente o in combinazione**. Possono essere **integrati ai dispositivi aziendali cybersecurity già installati o da installare**, costituendo un livello di automazione e *data analytics* unico

# *zuanelli 2018 AI Big data*



# Le funzioni dell'ontologia di dominio

Oltre 600 entità del dominio *cybersecurity* e relativi nodi relazionali di tipo logico-semantico

La rappresentazione ontologica della *cybersecurity knowledge* e delle vulnerabilità, *threat*, impatto, resilienza, ecc. consente:

- la possibilità di **prevenire e individuare** eventi e incidenti cibernetici
- la definizione di **metodi e tecnologie per il *risk assessment* e la *risk evaluation***
- l'applicazione di **sistemi rimediali**, tecnologici e comportamentali
- l'impiego di **standard** per l'automazione della sicurezza
- la conoscenza, la definizione e la rappresentazione degli **elementi costitutivi di eventi e di incidenti *cybersecurity***, necessarie per la reportistica dei dati
- l'**analisi automatica delle variabili tipologiche** che definiscono gli eventi e gli incidenti
- modalità di **rappresentazione dei dati**

# Le metriche

- L'ontologia ricomprende la rappresentazione strutturata di masse di dati
- Correlazione tra *asset tecnologici* dell'ente/azienda e *cybersecurity implementation*
- Tipologia di registrazione del valore di rischio e *assessment* degli incidenti

# Stakeholder privilegiati

- **CERT, CSIRT, SIEM, SOC e aziende di infrastrutture critiche e servizi essenziali** come da Direttiva NIS, GDPR e standard sicurezza internazionali
- I dati aziendali inseriti nel database customizzato consentono di analizzare le evoluzioni del sistema di sicurezza nella logica della ***big data analytics***. La correlazione dei dati consente di pervenire a una interpretazione preventiva e, in prospettiva, predittiva degli attacchi
- La piattaforma ontologica può alimentare la struttura delle varie *libraries*

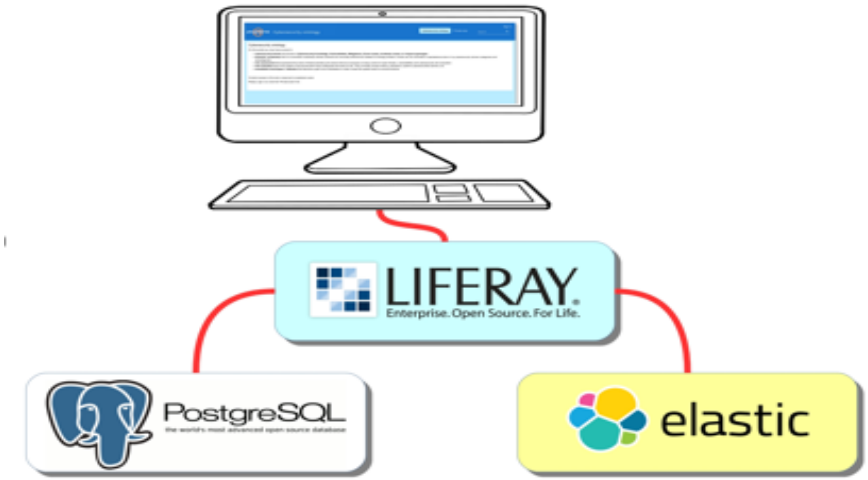
## La Direttiva NIS e il GDPR (PIA): correlazione e automazione *compliance*

- Operative da maggio 2018
- *Risk assessment* e *risk evaluation*
- Adozione standard tecnologici  sicurezza (NIS)per aziende di servizi essenziali e aziende di servizi digitali (*search engine*, *e commerce*, *cloud*, DNS e Registrar

## Architettura e infrastruttura tecnologica

POC è sviuppato su Liferay, portale orizzontale scritto in Java;

Post-gres,database relazionale e ricerca elastic; motore di ricerca scritto in Java



# La forza dell'architettura POC: il parametro strutturale

- Il **server applicazioni**: Liferay è costruito sui microservizi Java e progettato per essere un server a *clustered persistence*
- Il database relazionale Postgres è progettato per essere replicato
- Il server di indicizzazione e ricerca Elastic search è progettato per gestire ricerche distribuite

# Il parametro architetturale

I componenti sono scalabili e offrono un ampio margine di crescita

I tratti caratterizzanti

- **internazionalizzazione**: gestisce molteplici lingue per la navigazione e I contenuti
- la **categorizzazione** è indipendente dalle lingue: ogni categoria ha un identificatore e una descrizione multipla in varie lingue in modo che la scelta in base alle categorie dà lo stesso risultato in lingue diverse
- **integrabile:** basato su standard, può essere integrato con qualunque sistema *legacy*
- **gestione granulare del ruolo**: gestione utenza sofisticata per la separazione di gestione operative e visibilità

il **sistema di gestione interno può essere integrato** con sistemi esterni offrendo un segnale singolo per i sistemi e l'identita di sistemi

# I servizi tecnologici alla cybersecurity della piattaforma POC

# I servizi della piattaforma/ontologia pragmatica POC Pragmema

Tutti i servizi sono retti dall'ontologia di dominio *cybersecurity*:

- il **vocabolario semantico controllato** della *cybersecurity* (VSC) (e i vocabolari generici) è il servizio di definizione delle informazioni e dei dati utili per gestire univocamente le attività di cybersecurity

- ***risk assessment***: il servizio consente di verificare lo stato dell'arte preliminare e in progress del sistema e nel tempo l'andamento ovvero il *check* di sistema di sicurezza inerente dati e tecnologie delle aziende e delle istituzioni. Su richiesta del cliente il modello *check* può essere personalizzato e interrelato ai sistemi tecnologici di *cybersecurity* aziendali/istituzionali

- ***risk evaluation***: il servizio consente di valutare il rischio economico dell'azienda in caso di attacchi *cyber*. Può essere customizzato …

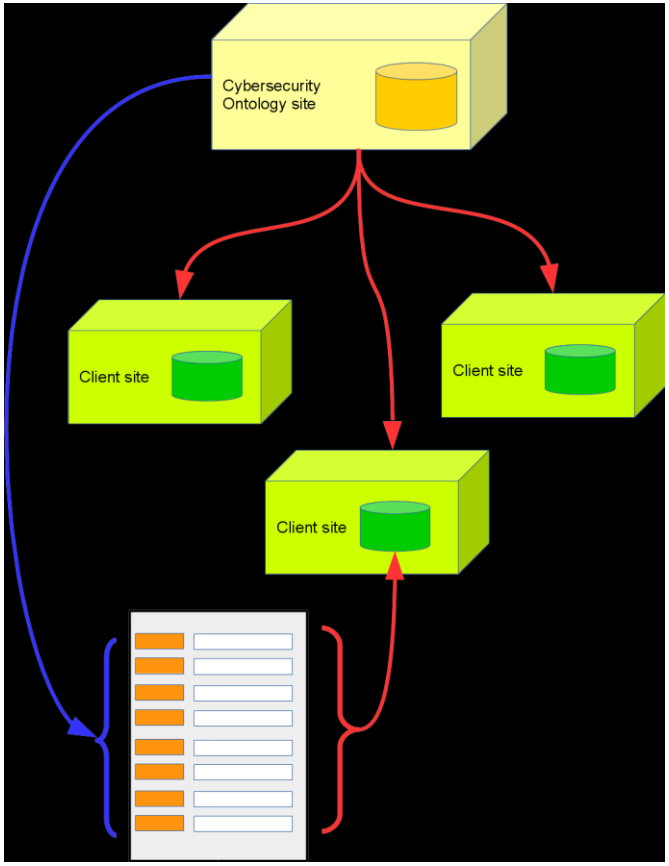
# I servizi alla cybersecurity

***remediation techniques***: il servizio consente di definire le attività di resilienza in caso di attacchi. Può essere customizzato

- ***standard references***: il servizio offre la gamma di standard internazionali richiesti per la *cybersecurity*. E' connesso ai **form/*application tools***
- gli **application tools** contengono form/schemi da compilare e memorizzare sul *data base* dell'azienda o dell'istituzione correlandosi ai servizi di sicurezza presenti nel sistema informativo informatico ed effettuare le seguenti attività: reportistica di eventi ed incidenti, statistiche, valutazione del rischio, ***assessment***, ecc. Consente **l'automazione del servizio** e dei **segnali di rischio**

# Features of all services

- *Every client has **his/her own service site***
  - *Only subscribed services are accessible*
  - ***Client data remain in the client site***
- *Each service displays:*
  - *Structured forms to **gather field data***
  - *Each field data has a **contextual explanation/implementation** based on the Cybersecurity Ontology*
  - *Forms data **are stored in the client site and represent a time series***



# Risk assessment service

- *Acquisition forms for:*
  - *Asset analysis*
  - *Reporting*
  - *Incidents identifiers*
  - *Event description*
  - *Metrics*
  - *Statistics*
- *Search and correlation of the acquired data*
- *Access to the related knowledge base in the Cybersecurity Pragmatic Domain*

# Risk evaluation service

- *Acquisition forms for:*
  - *Metrics*
  - *Statistics*
- *Search and correlation of the acquired data*
- *Access to the related knowledge base in the Cybersecurity Pragmatic Domain*

# Remediation techniques service

- *Acquisition forms for:*
  - *Preparation*
  - *Identification*
  - *Containment*
  - *Eradication*
  - *Recovery*
  - *Post-incident activity*
- *Search and correlation of the acquired data*
- *Access to the related knowledge base in the Cybersecurity Pragmatic Domain*

## Standards reference service

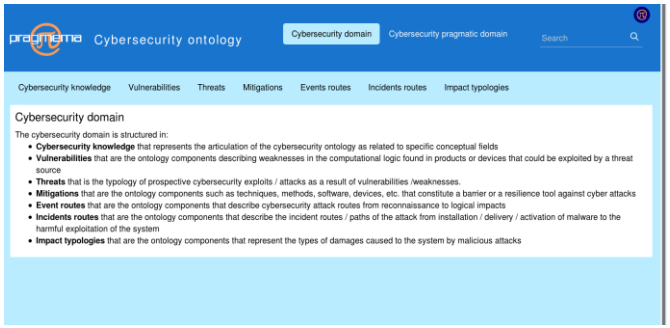
- *Access to the knowledge base in the Cybersecurity Pragmatic Domain*
  - *Access to relevant standards with annotated focal points*
  - *Standard documents correlation with Cybersecurity Ontology structure*
  - *Standard documents are organized following the Cybersecurity Ontology structure*

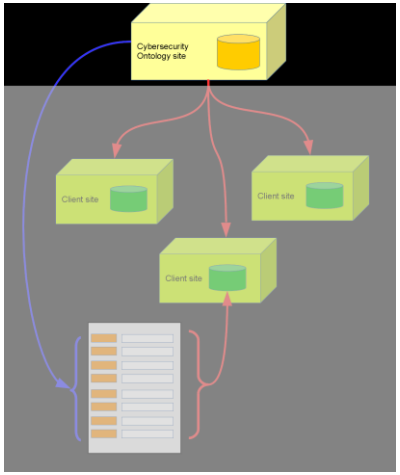
## *Controlled Semantic Vocabulary service*

- *Access to the knowledge base in the Cybersecurity Domain*
- *Multiple vocabularies with search functions*
  - *Cybersecurity Ontology Domain Vocabulary*
  - *Common attack Pattern Enumeration and Classification (CAPEC™)*
  - *Others*

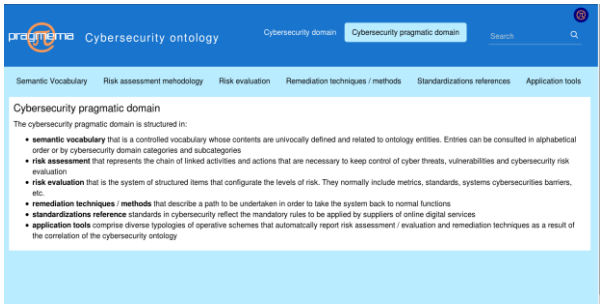
# Pragmema Cybersecurity Ontology Domain

- *Cybersecurity domain is structured in:*
  - *Cybersecurity knowledge*
  - *Vulnerabilities*
  - *Threats*
  - *Mitigations*
  - *Event routes*
  - *Incidents routes*
  - *Impact typologies*



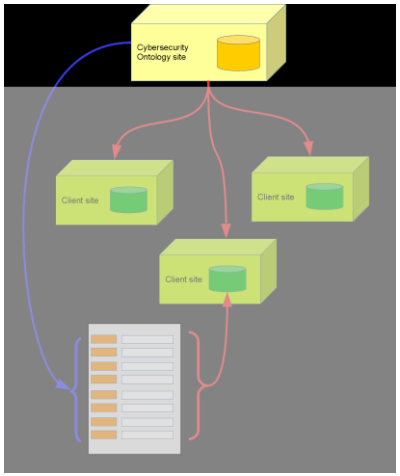



# Pragmema Cybersecurity Pragmatic Domain




*Cybersecurity pragmatic domain is structured in:*

- *semantic vocabulary*
- *risk assessment*
- *risk evaluation*
- *remediation techniques / methods*
- *standardizations reference*



# Grazie!

# Q&A