VI Conferenza
Intelligenza artificiale, sanità, protezione dei dati: stato dell'arte e prospettive di sviluppo
Presentazione del Master in
«Competenze digitali per la Protezione dei Dati, la Cybersecurity e la Privacy»
Università degli Studi di Roma «Tor Vergata»
Roma **22 aprile 2024**

Intelligenza artificiale, sanità, protezione dei dati: temi e problemi

Elisabetta Zuanelli
Coordinatore del Partenariato per il Piano di formazione in *Cybersecurity*, *Cyberthreat* e *Privacy*
Esperto ONU e NATO di Intelligenza artificiale

1

Gli assi tematici

L'INTELLIGENZA ARTIFICIALE (AI)

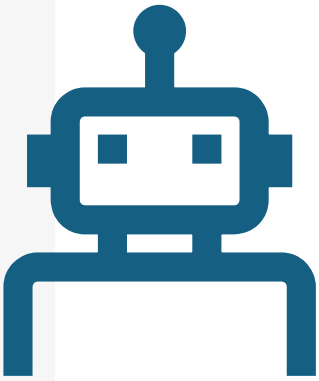
AI E SANITÀ

AI *CYBERSECURITY* E *CYBERSECURITY* SANITÀ

AI NORME E *COMPLIANCE*

DIGITAL HEALTH TECH: UN MERCATO CHE VARRÀ $ 1,5 TRILIONI ENTRO IL 2030

2



AI: temi e problemi

Le definizioni

Le funzioni attività dell'AI

La modellistica tecnologica di sviluppo: Chat GPT 4, ontologie e tassonomie

Gli strumenti:

Deep neural network machine learning

3

Le definizioni e le applicazioni AI

Temi

Dai dati, modelli di dati e basi di dati alla rappresentazione di conoscenza e alle **knowledge bases**

soluzioni tecnologiche **architetturali in AI**

strumenti e applicazioni/**prodotti e servizi AI**

Machine learning, deep learning: acquisizione ed elaborazione di *knowledge*

Problemi

L'emulazione delle funzioni intellettive e operative umane:

- reti neurali, funzioni cerebrali, funzioni cognitive, funzioni linguistiche
- *deep neural network machine learning*

4

WHO
Regulatory considerations

Machine learning-enabled medical devices: a subset of AI-enabled medical devices: key terms and definitions. Proposed document posted for public consultation, 16 September 2021. International Medical Device Regulators Forum; 2021 (https://www.imdrf.org/sites/default/files/202110/Machine%20Learningenabled%20Medical%20Devices%20%20A%20subset%20of%20Artificial%20Intelligenceenabled%20Medical%20Devices%20-%20Key%20Terms%20and%20Definitions.pdf).

1. Artificial Intelligence

AI is a branch of computer science, statistics and engineering that **uses algorithms or models** to perform tasks and exhibit behaviours such as **learning, making decisions and making predictions.**

The subset of AI known as ML allows computer algorithms to learn through data, **without being explicitly programmed to perform a task**

5

Meglio forse???

11. AI system

The IMDRF defines an AI system as a **software** that is developed with one or more of the techniques and approaches listed below* and can, for a given set of human-defined objectives, generate outputs such as **content**, predictions, recommendations or decisions that **influence** the environments they interact with.

AI techniques and approaches:

(a) **machine learning approaches**, including supervised, unsupervised and reinforcement learning, using a wide variety of methods, including deep learning;

(b) **logic- and knowledge-based approaches**, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems;

(c) **statistical approaches**, Bayesian estimation, search and optimization methods.

6

Le funzioni cognitive e la mente: il grande motore delle attività umane

- percezione
- analisi
- trasduzione
- integrazione/memorizzazione
- elaborazione
- *retrieval*
- applicazioni

7

AI technological modeling

- Neural, cognitive and linguistic approaches
- CHAT GPT 4: (transformer attentive models) NLP, *data base/ knowledge base*, *linguistic processing* e servizi: *extraction*, *composition*, *information queries*, analisi multimodali, ecc.
- Deep neural network machine learning: Gitta Kutyniok sui *mathematical foundational support* e *criticism for data architectures*
- Ontologie e tassonomie: def e metodi

  Guarino, MITRE, Protégés/Stanford University
- POC Cybersecurity defense systems
- Linguaggi misti/*blended*
- Riconoscimento immagini: *labelled images* e *polisemy of images* (Google *image analytics*)

8

Gitta Kutyniok “a comprehensive theoretical mathematical foundation in AI is completely lacking at present”

In AI, …, in the case of deep neural networks, “the search results is a **timeconsuming work for a suitable network architecture**,

a highly delicate trial-and-error-based (training) process,

and **missing error bounds** for the performance of the trained neural networks”.

Layers, *data sets (numerical, bivariate, multivariate, categorical, correlational, ecc.)*, architetture: criteri per la coerenza tipologica dei dati e l’architettura delle basi di conoscenza

9

Klecha & Co. Insights Report, “Digital Health Tech: Paving the way for a healthier future”, May 2022

10

AI e sanità: applicazioni

Abbondanza di «ontologie», «tassonomie», *thesauri*

siti

Applicazioni

radiografie

oncologia

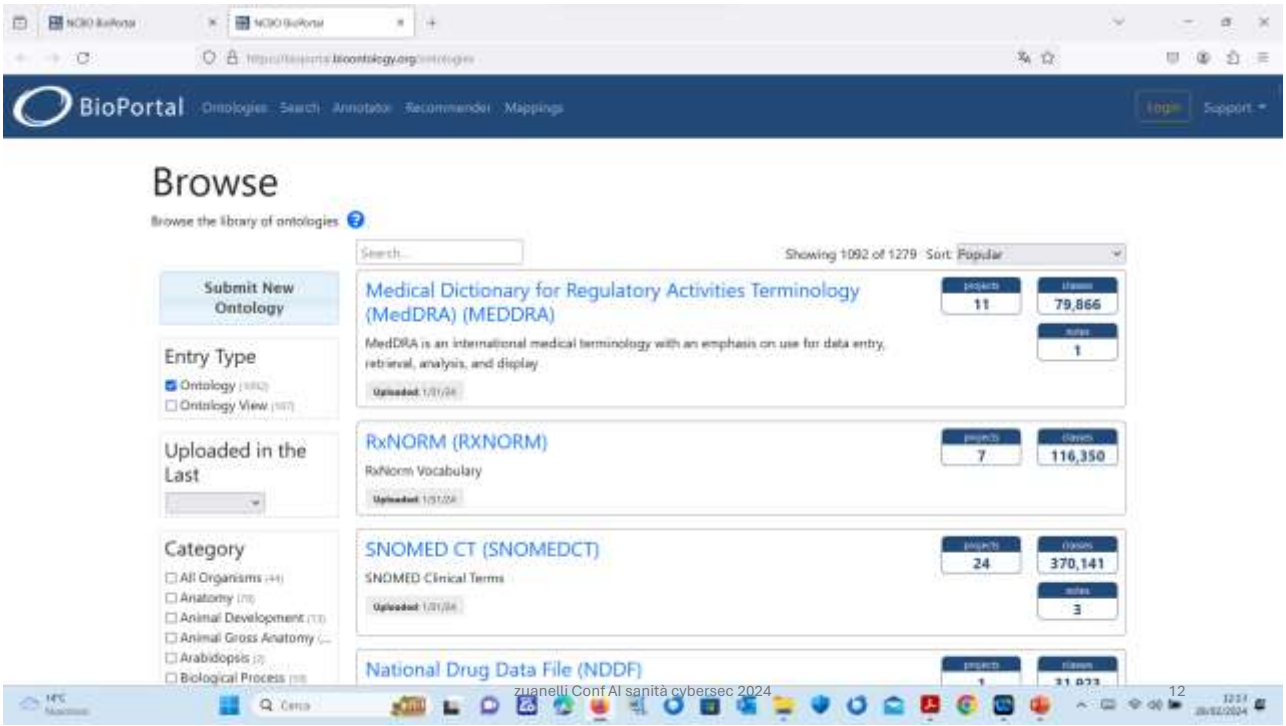
cardiologia

fisiatria

depressione (MIT speech patterns)

etc.

11



12

World Health Organization

WHO (World Health Organization)
Regulatory considerations on artificial intelligence for health

Classificazioni

https://www.who.int/standards/classifications/who-fic-maintenance

13

WHO Aspetti regolatori

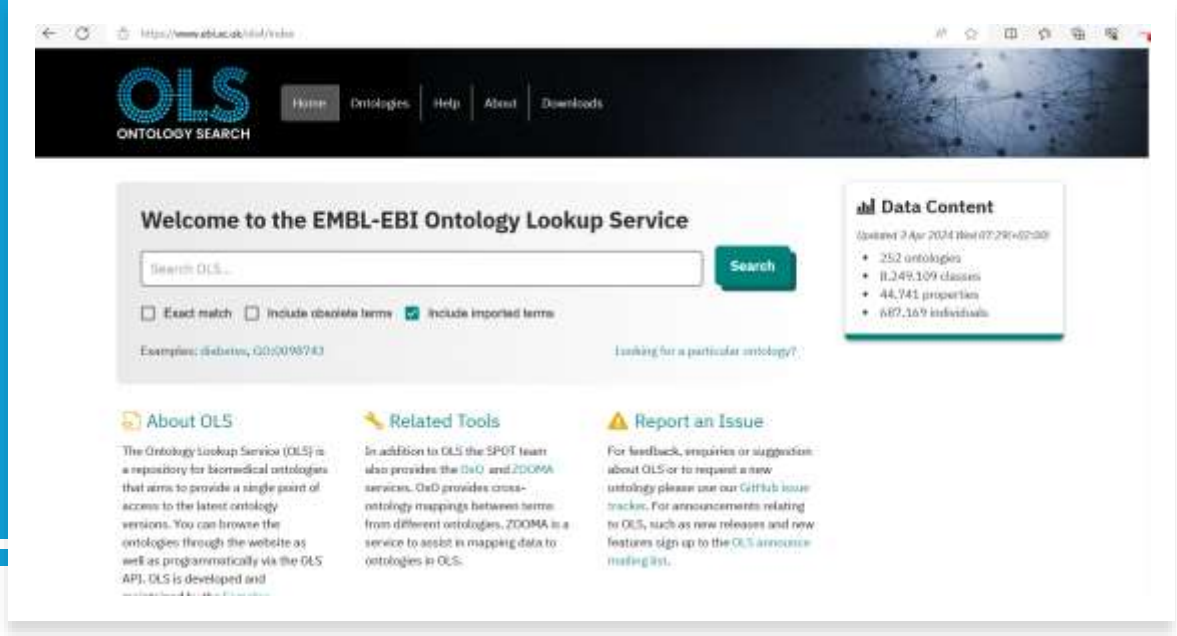
Health technology evaluation

Clinical evaluation components

FIGURE 9. Domains of health technology regulation, assessment and management for drugs and devicesh

FIGURE 10. IMDRF description of clinical evaluation components (4)

14



15

An overview of software tools and applications involved in machine learning with biomedical ontologies

Type	Method/Tool	Description	URL
Processing and preprocessing ontologies	OWLAPI	Reference library to process OWL ontologies, supports most OWL reasoners [13]	https://github.com/owlcs/owlapi
	funowl	Python library to process OWL ontologies	https://github.com/hsolbrig/funowl
	owlready2	Python library to process OWL ontologies	https://pypi.org/project/Owlready2/
	Apache Jena ontAPI	RDF library with OWL support	https://jena.apache.org/
	Protégé	Python RDF library with OWL support	https://github.com/RDFLib/owlrl
		Ontology editor and knowledge engineering environment [14]	https://protege.stanford.edu/
Computing entailments, reasoning	ELK	Very fast reasoner for the OWL 2 EL profile with polynomial worst-case time complexity [15]	https://github.com/liveontologies/elk-reasoner
	HermiT	Automated reasoner supporting most of OWL ontologies with exponential worst-case complexity [16]	http://www.hermit-reasoner.com/
	Pellet	OWL reasoner supporting most of the OWL constructs and supporting several additional features [17]	https://github.com/stardog-union/pellet
Generating graphs from ontologies	OBOGraphs	Syntactic conversion of ontologies to graphs, targeted at OBO ontologies	https://github.com/geneontology/obographs
	Onto2Graph	Semantic conversion of OWL ontologies to graphs, following the axiom patterns of the OBO Relation Ontology [18]	https://github.com/bio-ontology-research-group/Onto2Graph
Computing Semantic Similarity	Semantic Measures Library	Comprehensive Java library to compute semantic similarity measures over ontologies [19]	http://www.semantic-measures-library.org/sml/
	semantic	Python library to compute semantic similarity on knowledge graphs [20]	https://github.com/gpu-mooc/semantic
	DiShIn	Python library for semantic similarity on ontologies [21]	https://github.com/lasigeBioTM/DiShIn
Embedding graphs	OWL2Vec	Method that combines generation of graphs from ontologies, random walks on the generated graphs, and generation of embeddings using Word2Vec. Syntactically processes most OWL axioms [22]	https://github.com/bio-ontology-research-group/owl2vec-star
	DL2Vec	Method that combines generation of graphs from ontologies, random walks on the generated graphs, and generation of embeddings using Word2Vec. Syntactically processes most OWL axioms [23]	https://github.com/bio-ontology-research-group/DL2Vec
	Walking RDF&OWL	Method that combines generation of graphs from ontologies, random walks on the generated graphs, and generation of embeddings using Word2Vec. Only considers the ontology taxonomy. [24]	https://github.com/bio-ontology-research-group/walking-rdf-and-owl
	RDF2Vec	Method to embed RDF graphs [25]	https://github.com/IBCNServices/pyRDF2Vec, https://github.com/dwslab/jRDF2Vec
	Node2Vec	Method to embed graphs using biased random walks [26]	http://snap.stanford.edu/node2vec/
	PyKEEN, BioKEEN	Toolkit for generating knowledge graph embeddings using several different approaches [27,28]	https://github.com/SmartDataAnalytics/PyKEEN
	OpenKE	Library and toolkit for generating knowledge graph embeddings	https://github.com/thunlp/OpenKE
	PyTorch Geometric	Library for graph neural networks which can be used to generate graph embeddings [29]	https://github.com/rusty1s/pytorch_geometric
Embedding ontologies	Onto2Vec	Embeddings based on treating logical axioms as a text corpus [30]	https://github.com/bio-ontology-research-group/onto2vec
	OPA2Vec	Embeddings that combine logical axioms with annotation properties and the literature [31]	https://github.com/bio-ontology-research-group/opa2vec
	EL Embeddings	Embeddings that approximate the interpretation function and preserve semantics for intersection, existential quantifiers, and bottom [32]	https://github.com/bio-ontology-research-group/el-embeddings
Ontology-based constrained learning	DeepGO	Implements an ontology-based hierarchical classifier for function prediction. The hierarchical classification module is generic and can be used with other ontologies and applications [33]	https://github.com/bio-ontology-research-group/deepgo
	DEEPred	Automated Protein Function Prediction with Multi-task Feed-forward Deep Neural Networks [34]	https://github.com/cellmapper/DEEPred
	DeepMir2GO	Inferring Functions of Human MicroRNAs Using a Deep Multi-Label Classification Model [35]	https://github.com/Shanshan-Wang/DeepMir2GO

16

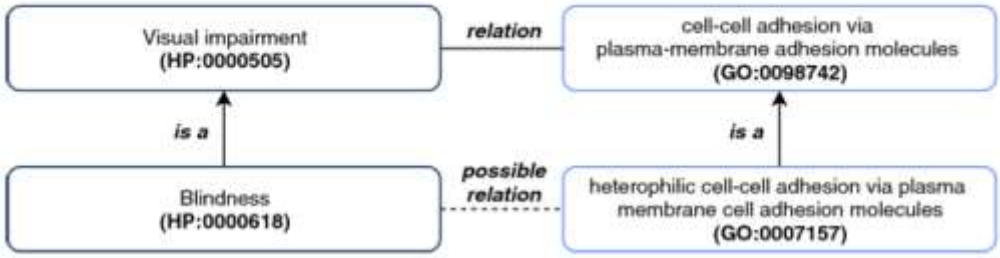
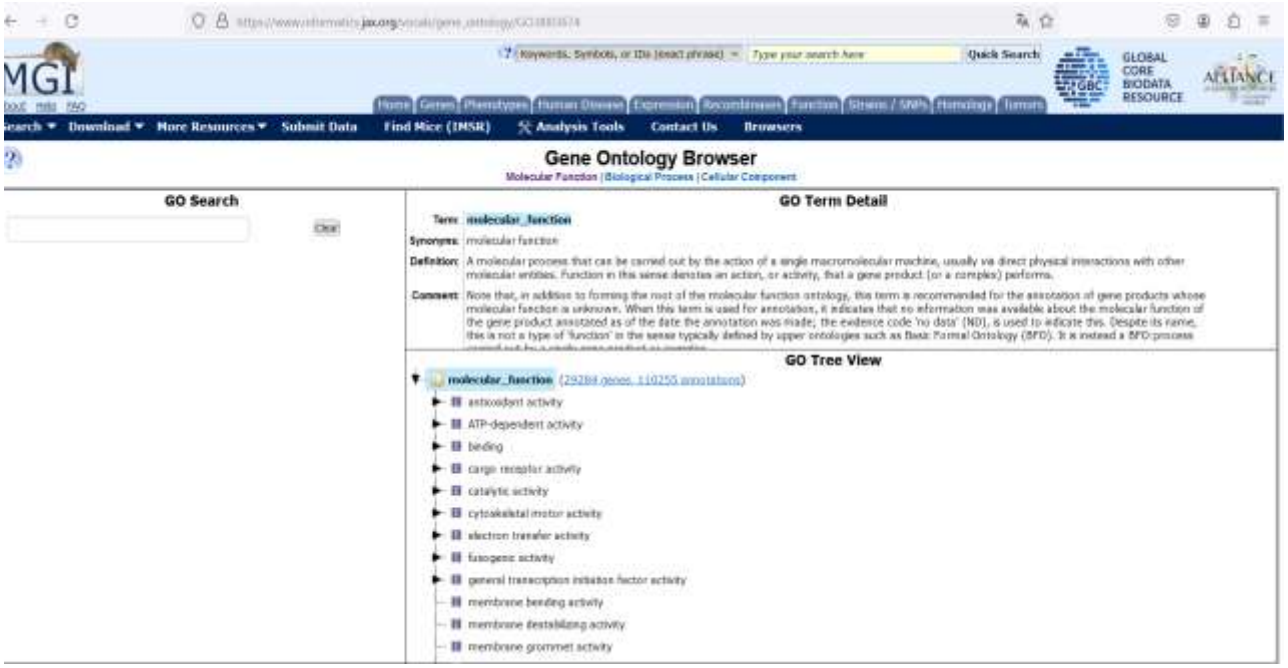


Figura 14 *- Esempio di ontology embedding in BiOnt basato sulle ontologie HPO e GO, per la relazione candidata tra il fenotipo umano "cecità" e il gene CRB1 (rappresentato dal termine GO:0007157 " heterophilic cell-cell adhesion via plasma membrane cell adhesion molecules"). (Fonte: https://link.springer.com/chapter/10.1007%2F978-3-030-45442-5_46#Sec2).*

17



18

Un esempio di possibile riformulazione

TABLE 1: Characteristics of the Two Types of Femoroacetabular Impingement

Criteria	Pincer Impingement	Cam Impingement
Cause	Focal or general overcoverage	Aspherical head
Mechanism	Linear contact between overcovering rim and head–neck junction	Jamming of aspherical head portion into acetabulum
Sex distribution (M:F)	1:3	14:1
Average age (range) (y)	40 (40–57)	32 (21–51)
Typical location of cartilage damage	Circumferential with contrecoup	11- to 3-o'clock position
Average depth of cartilage damage (mm)	4	11
Associated disorders	Bladder extrophy	Slipped capital femoral epiphysis
	Proximal femoral focal deficiency	Legg-Calvé-Perthes disease
	Posttraumatic dysplasia	Posttraumatic retrotorsion of femoral head
	Chronic residual dysplasia of acetabulum	Coxa vara
	Legg-Calvé–Perthes disease	Pistol-grip deformity
	Slipped capital femoral epiphysis	Head-tilt deformity
	After acetabular reorientation procedures	Post-slip deformity
	Idiopathic retroversion	Femoral retroversion
		Growth abnormality of femoral epiphysis
Radiographic signs on anteroposterior radiographs	Coxa profunda	Pistol-grip deformity
	Protrusio acetabuli	CCD angle < 125°
	Focal acetabular retroversion (figure-8 configuration)	Horizontal growth plate sign
	Lateral center edge angle > 39°	
	Reduced extrusion index	
	Acetabular index ≤ 0°	
	Posterior wall sign	
Radiographic signs on cross-table radiographs	Linear indentation sign	Alpha angle > 50°
		Femoral head–neck offset < 8 mm
		Offset ratio < 0.18
		Femoral retrotorsion
Secondary changes	Herniation pits	
	Ossification of labrum	
	Appositional bone sign	
	Os acetabuli	
	Posterior inferior joint space loss (on faux profile in pincer hips)	
	Late, classic signs of osteoarthritis	

19

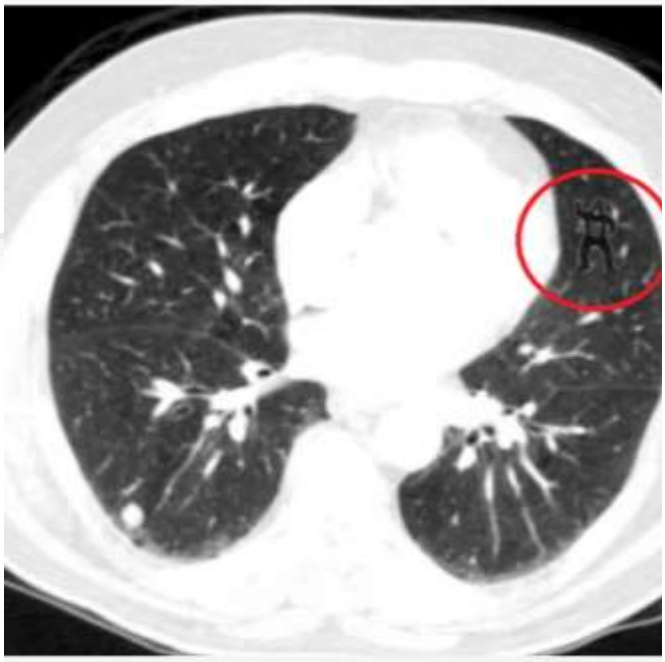
I linguaggi multimodali: large multimodal model (LMM)

La metodologia

- ❑La qualità delle *xrays*
- ❑Immagini e testi:
  - la combinazione relazionale logico-semantica: rappresentazione vettoriale (*embedding*) delle immagini compatibile con la rappresentazione vettoriale di una frase
  - il *machine learning*, i *data sets*
  - il profilo architetturale
- ❑Domini, classi, entità, relazioni logico-semantiche, vocabolari semantici controllati, metadati

20

Come viene percepita e analizzata l'immagine: il gorilla nascosto



21

L'Ontologia cybersecurity POC: *threat intelligence*, *information sharing*, *incident reporting* (2024) Pragmema

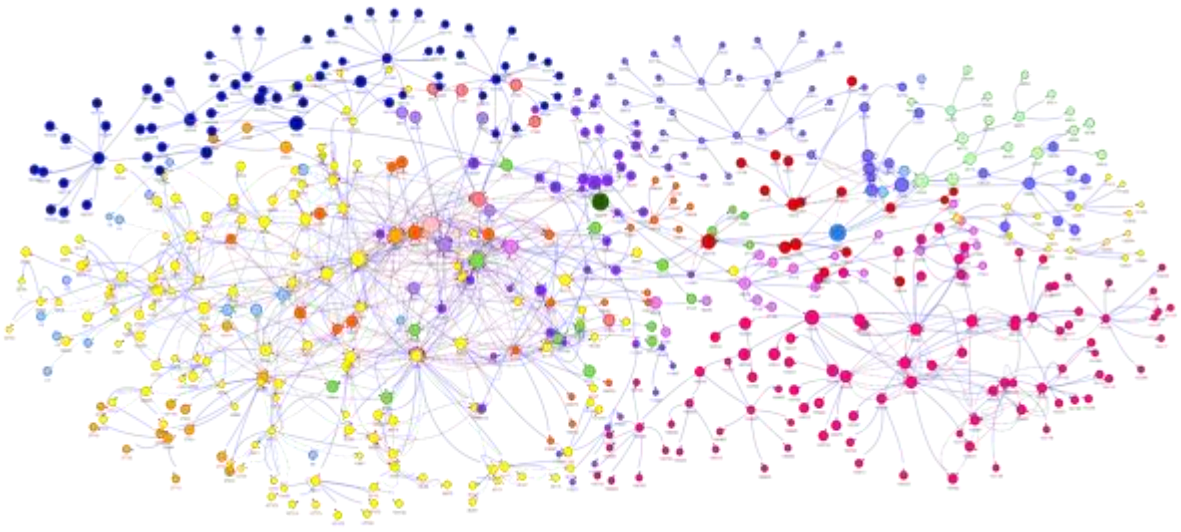
- POC *knowledge ontology*
- POC *cybersecurity domain ontology*: prevenzione e predittività
- POC *cybersecurity pragmatic ontology: services*

22

POC *knowledge ontology* (*upper level*)

23

POC cybersecurity domain ontology



24

POC *cybersecurity subdomain ontologies*

- financial
- automotive
- shipping

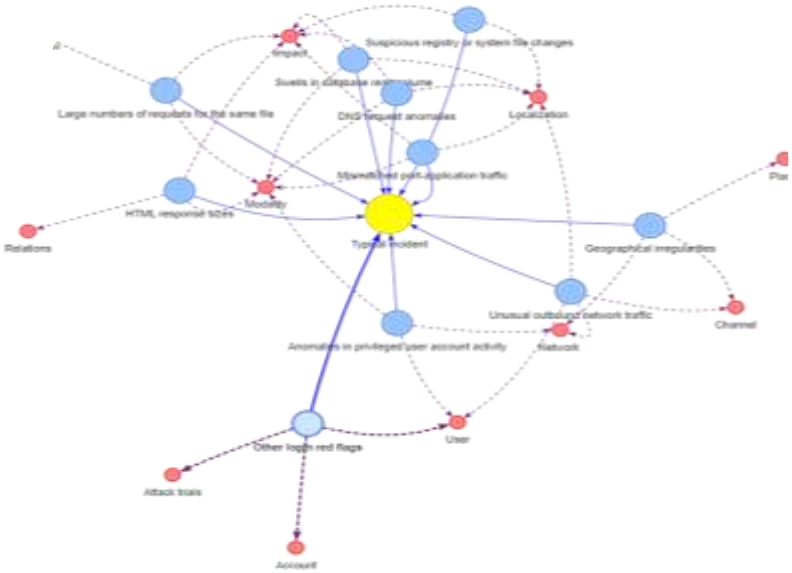
25

POC: typical incident

26

POC: *typical incident* IOC

- integrazione e filtro
- *clustering* di parametri



27

AI, cybersecurity, protezione dei dati: il triplo ambito

- ❑architetture per *defense systems*
- ❑il perimetro di insicurezza AI
- ❑AI e *adversarial attacks*
- ❑AI e *cybersecurity by design*

28

Norme, *compliance* e rischi

- Il regolamento EU AI Act e le norme corollarie di settori: sanità, *cybersecurity* e protezione dati (NIS e NIS 2, GDPR, *data governance*, servizi digitali, ecc.)
- Le *guidelines* sicurezza AI
- Le norme ISO: 42001e altro
- *Compliance*: soggetti implicati
- Le nozioni di etica e rischio nell'AI

Prospettive problematiche e vasti ambiti di ricerca e applicazione

La rilevanza della ricerca e dello sviluppo e le *joint ventures*

GRAZIE PER L'ATTENZIONE!