national center for ontological research
ncor

University at Buffalo
The State University of New York

Cybersecurity through an Ontological Lens

John Beverley

Assistant Professor, *University at Buffalo*
Co-Director, *National Center for Ontological Research*
Affiliate Faculty, *Institute of Artificial Intelligence and Data Science*

Outline

- Cybersecurity Landscape
- Proliferation of Knowledge Representation
- Addressing the Alignment Problem
- Design Pattern Practice

Outline

- Cybersecurity Landscape
- Proliferation of Knowledge Representation
- Addressing the Alignment Problem
- Design Pattern Practice

2012 LinkedIn Hack

- Passwords for nearly **6.5 million** users were stolen by Russian cybercriminals, in particular Yevgeniy Nikulin

- In 2016, LinkedIn discovered **100 million** email addresses and hashed passwords had also been stolen

- In 2021, a torrent file was made available with data for over **700 million** users, much of which could be traced to the linkedIn breach

Cause for Alarm

- You might be thinking "I don't really use LinkedIn, so I've no reason to worry..."
- I bet you reuse passwords...
- We create new accounts all the time...
- We reuse passwords all the time...

Top 25 most common passwords by year according to SplashData

Rank	2011[6]	2012[7]	2013[8]	2014[9]	2015[10]	2016[5]	2017[11]	2018[12]	2019[13]
1	password	password	123456	123456	123456	123456	123456	123456	123456
2	123456	123456	password	password	password	password	password	password	123456789
3	12345678	12345678	12345678	12345	12345678	12345	12345678	123456789	qwerty
4	qwerty	abc123	qwerty	12345678	qwerty	12345678	qwerty	12345678	password
5	abc123	qwerty	abc123	qwerty	12345	football	12345	12345	1234567
6	monkey	monkey	123456789	123456789	123456789	qwerty	123456789	111111	12345678
7	1234567	letmein	111111	1234	football	1234567890	letmein	1234567	12345
8	letmein	dragon	1234567	baseball	1234	1234567	1234567	sunshine	iloveyou
9	trustno1	111111	iloveyou	dragon	1234567	princess	football	qwerty	111111
10	dragon	baseball	adobe123[a]	football	baseball	1234	iloveyou	iloveyou	123123
11	baseball	iloveyou	123123	1234567	welcome	login	admin	princess	abc123
12	111111	trustno1	admin	monkey	1234567890	welcome	welcome	admin	qwerty123
13	iloveyou	1234567	1234567890	letmein	abc123	solo	monkey	welcome	1q2w3e4r
14	master	sunshine	letmein	abc123	111111	abc123	login	666666	admin
15	sunshine	master	photoshop[a]	111111	1qaz2wsx	admin	abc123	abc123	qwertyuiop
16	ashley	123123	1234	mustang	dragon	121212	starwars	football	654321
17	bailey	welcome	monkey	access	master	flower	123123	123123	555555
18	passw0rd	shadow	shadow	shadow	monkey	passw0rd	dragon	monkey	lovely
19	shadow	ashley	sunshine	master	letmein	dragon	passw0rd	654321	7777777
20	123123	football	12345	michael	login	sunshine	master	!@#$%^&*	welcome
21	654321	jesus	password1	superman	princess	master	hello	charlie	888888
22	superman	michael	princess	696969	qwertyuiop	hottie	freedom	aa123456	princess
23	qazwsx	ninja	azerty	123123	solo	loveme	whatever	donald	dragon
24	michael	mustang	trustno1	batman	passw0rd	zaq1zaq1	qazwsx	password1	password1
25	Football	password1	000000	trustno1	starwars	password1	trustno1	qwerty123	123qwe

Be honest, how many of you have used the password “password” or “123456” before?

Top 25 most common passwords by year according to SplashData

Rank	2011[6]	2012[7]	2013[8]	2014[9]	2015[10]	2016[5]	2017[11]	2018[12]	2019[13]
1	password	password	123456	123456	123456	123456	123456	123456	123456
2	123456	123456	password	password	password	password	password	password	123456789
3	12345678	[illegible]	12345678	12345	12345678	12345	[illegible]	123456789	qwerty
4	qwerty	abc123	qwerty	12345678	qwerty	12345678	qwerty	12345678	password
5	abc123	qwerty	abc123	qwerty	12345	football	12345	12345	1234567
6	monkey	monkey	123456789	123456789	123456789	qwerty	123456789	111111	12345678
7	1234567	letmein	111111	1234	football	1234567890	letmein	1234567	12345
8	letmein	dragon	1234567	baseball	1234	1234567	1234567	sunshine	iloveyou
9	trustno1	111111	iloveyou	dragon	1234567	princess	football	qwerty	111111
10	dragon	baseball	adobe123[a]	football	baseball	1234	iloveyou	iloveyou	123123
11	baseball	iloveyou	123123	1234567	welcome	login	admin	princess	abc123
12	111111	trustno1	admin	monkey	1234567890	welcome	welcome	admin	qwerty123
13	iloveyou	1234567	1234567890	letmein	abc123	solo	monkey	welcome	1q2w3e4r
14	master	sunshine	letmein	abc123	111111	abc123	login	666666	admin
15	sunshine	master	photoshop[a]	111111	1qaz2wsx	admin	abc123	abc123	qwertyuiop
16	ashley	123123	1234	mustang	dragon	121212	starwars	football	654321
17	bailey	welcome	monkey	access	master	flower	123123	123123	555555
18	passw0rd	shadow	shadow	shadow	monkey	passw0rd	dragon	monkey	lovely
19	shadow	ashley	sunshine	master	letmein	dragon	passw0rd	654321	7777777
20	123123	football	12345	michael	login	sunshine	master	!@#$%^&*	welcome
21	654321	jesus	password1	superman	princess	master	hello	charlie	888888
22	superman	michael	princess	696969	qwertyuiop	hottie	freedom	aa123456	princess
23	qazwsx	ninja	azerty	123123	solo	loveme	whatever	donald	dragon
24	michael	mustang	trustno1	batman	passw0rd	zaq1zaq1	qazwsx	password1	password1
25	Football	password1	000000	trustno1	starwars	password1	trustno1	qwerty123	123qwe

Be honest, how many of you have used the password “111111” before?

Top 25 most common passwords by year according to SplashData

Rank	2011[6]	2012[7]	2013[8]	2014[9]	2015[10]	2016[5]	2017[11]	2018[12]	2019[13]
1	password	password	123456	123456	123456	123456	123456	123456	123456
2	123456	123456	password	password	password	password	password	password	123456789
3	12345678	12345678	12345678	12345	12345678	12345	12345678	123456789	qwerty
4	qwerty	abc123	qwerty	12345678	qwerty	12345678	qwerty	12345678	password
5	abc123	qwerty	abc123	qwerty	12345	football	12345	[illegible]	1234567
6	monkey	monkey	[illegible]6789	123456789	123456789	qwerty	123456789	111111	12345678
7	1234567	letmein	111111	1234	football	1234567890	letmein	[illegible]7	12345
8	letmein	dragon	[illegible]7	baseball	1234	1234567	1234567	sunshine	[illegible]
9	trustno1	111111	iloveyou	dragon	1234567	princess	football	qwerty	111111
10	dragon	baseball	adobe123[a]	football	baseball	1234	iloveyou	iloveyou	[illegible]
11	baseball	iloveyou	123123	1234567	welcome	login	admin	princess	abc123
12	111111	trustno1	admin	monkey	1234567890	welcome	welcome	admin	qwerty123
13	iloveyou	1234567	1234567890	letmein	abc123	solo	monkey	welcome	1q2w3e4r
14	master	sunshine	letmein	abc123	111111	abc123	login	666666	admin
15	sunshine	master	photoshop[a]	111111	1qaz2wsx	admin	abc123	abc123	qwertyuiop
16	ashley	123123	1234	mustang	dragon	121212	starwars	football	654321
17	bailey	welcome	monkey	access	master	flower	123123	123123	555555
18	passw0rd	shadow	shadow	shadow	monkey	passw0rd	dragon	monkey	lovely
19	shadow	ashley	sunshine	master	letmein	dragon	passw0rd	654321	7777777
20	123123	football	12345	michael	login	sunshine	master	!@#$%^&*	welcome
21	654321	jesus	password1	superman	princess	master	hello	charlie	888888
22	superman	michael	princess	696969	qwertyuiop	hottie	freedom	aa123456	princess
23	qazwsx	ninja	azerty	123123	solo	loveme	whatever	donald	dragon
24	michael	mustang	trustno1	batman	passw0rd	zaq1zaq1	qazwsx	password1	password1
25	Football	password1	000000	trustno1	starwars	password1	trustno1	qwerty123	123qwe

Be honest, how many of you have used the password “dragon” before?

Top 25 most common passwords by year according to SplashData

Rank	2011[6]	2012[7]	2013[8]	2014[9]	2015[10]	2016[5]	2017[11]	2018[12]	2019[13]
1	password	password	123456	123456	123456	123456	123456	123456	123456
2	123456	123456	password	password	password	password	password	password	123456789
3	12345678	12345678	12345678	12345	12345678	12345	12345678	123456789	qwerty
4	qwerty	abc123	qwerty	12345678	qwerty	12345678	qwerty	12345678	password
5	abc123	qwerty	abc123	qwerty	12345	football	12345	12345	1234567
6	monkey	monkey	123456789	123456789	123456789	qwerty	123456789	111111	12345678
7	1234567	letmein	111111	1234	football	1234567890	letmein	1234567	12345
8	letmein	dragon	1234567	baseball	1234	1234567	1234567	sunshine	iloveyou
9	trustno1	111111	iloveyou	dragon	1234567	princess	football	qwerty	111111
10	dragon	baseball	adobe123[a]	football	baseball	1234	iloveyou	iloveyou	123123
11	baseball	iloveyou	123123	1234567	welcome	login	admin	princess	abc123
12	111111	trustno1	admin	monkey	1234567890	welcome	welcome	admin	qwerty123
13	iloveyou	1234567	1234567890	letmein	abc123	solo	monkey	welcome	1q2w3e4r
14	master	sunshine	letmein	abc123	111111	abc123	login	666666	admin
15	sunshine	master	photoshop[a]	111111	1qaz2wsx	admin	abc123	abc123	qwertyuiop
16	ashley	123123	1234	mustang	dragon	121212	starwars	football	654321
17	bailey	welcome	monkey	access	master	flower	123123	123123	555555
18	passw0rd	shadow	shadow	shadow	monkey	passw0rd	dragon	monkey	lovely
19	shadow	ashley	sunshine	master	letmein	dragon	passw0rd	654321	7777777
20	123123	football	12345	michael	login	sunshine	master	!@#$%^&*	welcome
21	654321	jesus	password1	superman	princess	master	hello	charlie	888888
22	superman	michael	princess	696969	qwertyuiop	hottie	freedom	aa123456	princess
23	qazwsx	ninja	azerty	123123	solo	loveme	whatever	donald	dragon
24	michael	mustang	trustno1	batman	passw0rd	zaq1zaq1	qazwsx	password1	passw0rd1
25	Football	password1	000000	trustno1	starwars	password1	trustno1	qwerty123	123qwe

Breadcrumbs

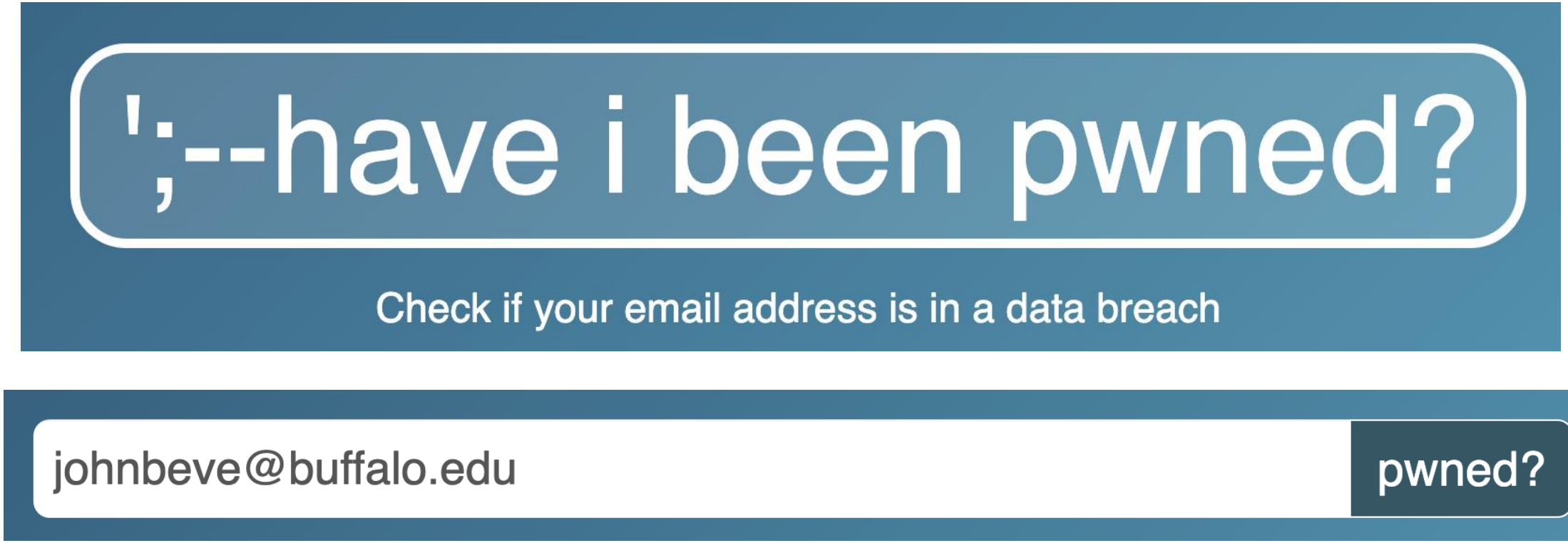
- Compromising a single account leads to information that can be used to compromise other accounts

- With your name and address, someone could call a credit card provider and – because they often ask for this information first – learn if you have a card

- Breadcrumbs lead to bread:
  - Date of Birth
  - Social Security Number
  - Security Questions and Answers

https://haveibeenpwned.com/

https://haveibeenpwned.com/



? ? ?

Why Yes I Have

- I've had this email address for over 10 years
- It's not a matter of if so much as a matter of when
- I'm surprised it's only appeared in 4 breaches

CafePress: In February 2019, the custom merchandise retailer CafePress suffered a data breach. The exposed data included 23 million unique email addresses with some records also containing names, physical addresses, phone numbers and passwords stored as SHA-1 hashes. The data was provided to HIBP by a source who requested it be attributed to "JimScott.Sec@protonmail.com".

Compromised data: Email addresses, Names, Passwords, Phone numbers, Physical addresses

Chegg: In April 2018, the textbook rental service Chegg suffered a data breach that impacted 40 million subscribers. The exposed data included email addresses, usernames, names and passwords stored as unsalted MD5 hashes. The data was provided to HIBP by a source who requested it be attributed to "JimScott.Sec@protonmail.com".

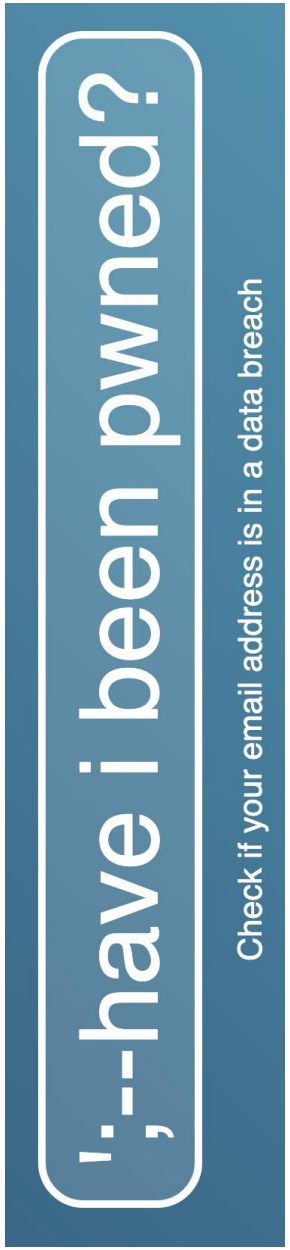
Compromised data: Email addresses, Names, Passwords, Usernames

Gravatar: In October 2020, a security researcher published a technique for scraping large volumes of data from Gravatar, the service for providing globally unique avatars . 167 million names, usernames and MD5 hashes of email addresses used to reference users' avatars were subsequently scraped and distributed within the hacking community. 114 million of the MD5 hashes were cracked and distributed alongside the source hash, thus disclosing the original email address and accompanying data. Following the impacted email addresses being searchable in HIBP, Gravatar release an FAQ detailing the incident.

Compromised data: Email addresses, Names, Usernames

MyFitnessPal: In February 2018, the diet and exercise service MyFitnessPal suffered a data breach. The incident exposed 144 million unique email addresses alongside usernames, IP addresses and passwords stored as SHA-1 and bcrypt hashes (the former for earlier accounts, the latter for newer accounts). In 2019, the data appeared listed for sale on a dark web marketplace (along with several other large breaches) and subsequently began circulating more broadly. The data was provided to HIBP by a source who requested it to be attributed to "BenjaminBlue@exploit.im".

Compromised data: Email addresses, IP addresses, Passwords, Usernames



lewispow@buffalo.edu pwned?

Oh no — pwned!

Pwned in 3 data breaches and found no pastes (subscribe to search sensitive breaches)

mpjens@gmail.com pwned?

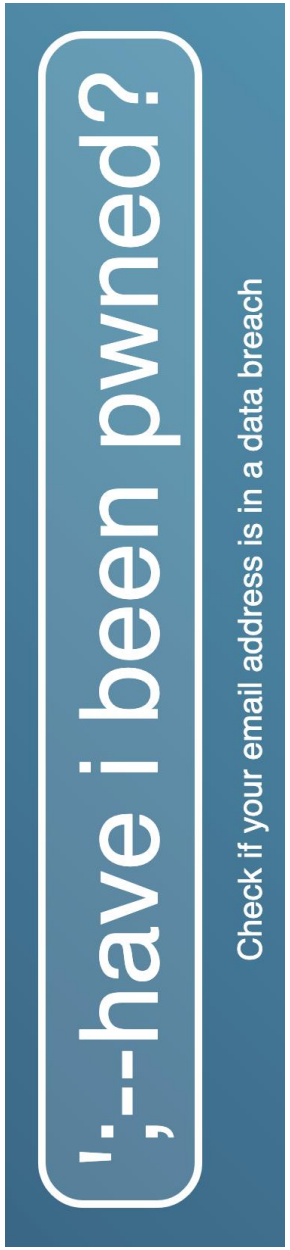
Oh no — pwned!

Pwned in 3 data breaches and found no pastes (subscribe to search sensitive breaches)

dbraun2@buffalo.edu pwned?

Oh no — pwned!

Pwned in 4 data breaches and found no pastes (subscribe to search sensitive breaches)



gdecolle@buffalo.edu pwned?

Good news — no pwnage found!

No breached accounts and no pastes (subscribe to search sensitive breaches)

mcnulty5@buffalo.edu pwned?

Good news — no pwnage found!

No breached accounts and no pastes (subscribe to search sensitive breaches)

cheung47@buffalo.edu pwned?

Good news — no pwnage found!

No breached accounts and no pastes (subscribe to search sensitive breaches)

Group Exercise

- Because I know you're not going to be able to resist, navigate to the "Have I been pwned?" website and explore whether your email accounts have been compromised

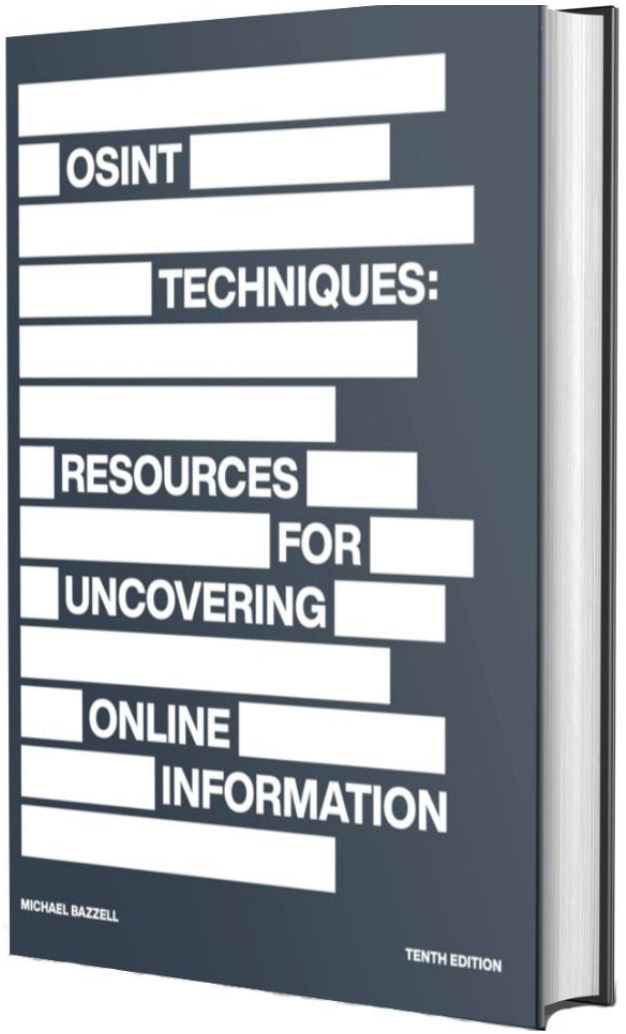
- You're welcome to discuss in small groups

Keep in mind, we'll be discussing as a class

Open Source Intelligence OSINT

- OSINT is intelligence that can be gathered from publicly available sources

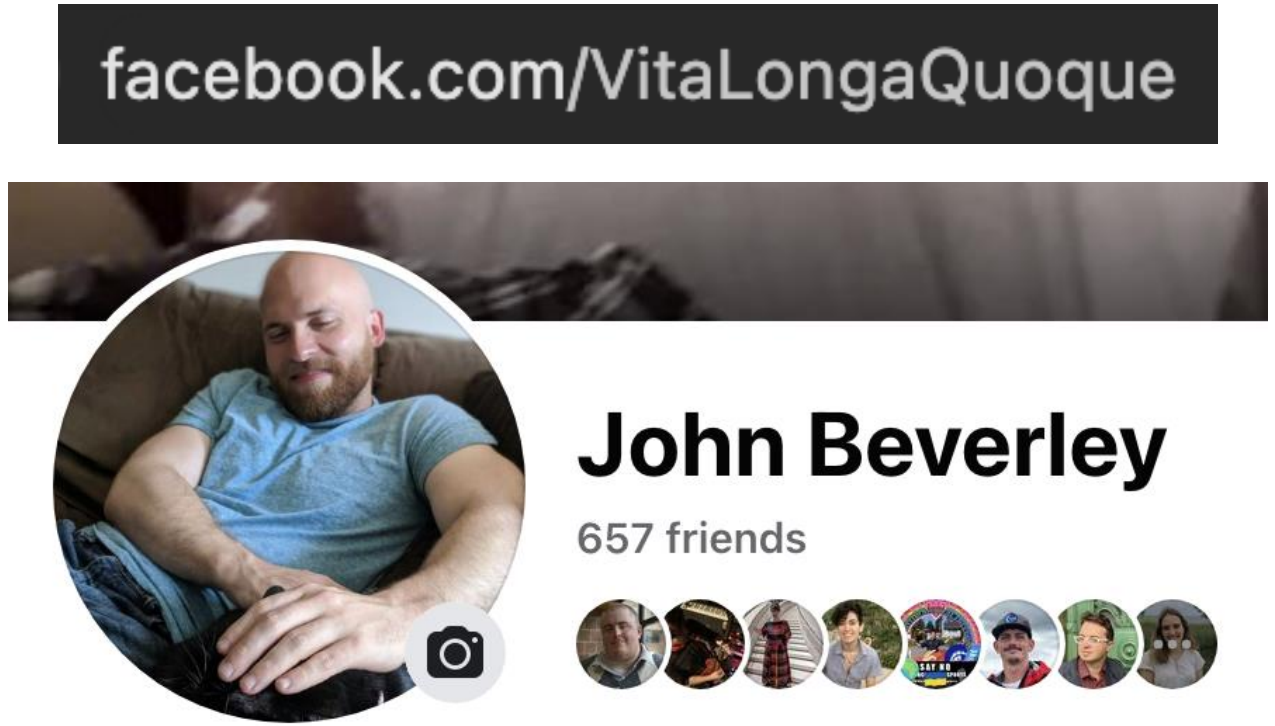
- For example, social and print media, public government data, professional publications, commercial data, etc.



Tools

Tools

- Search Engines
- Facebook
- Twitter
- Instagram
- LinkedIn
- Communities
- Email Addresses
- Usernames
- Names
- Addresses
- Telephone Numbers
- Maps
- Documents
- Pastes
- Images
- Videos
- Domains
- IP Addresses
- Business & Government
- Vehicles
- Virtual Currencies
- Breaches & Leaks
- Live Audio Streams
- Live Video Streams
- APIs



Tools

- Search Engines
- Facebook
- Twitter
- Instagram
- LinkedIn
- Communities
- Email Addresses
- Usernames
- Names
- Addresses
- Telephone Numbers
- Maps
- Documents
- Pastes
- Images
- Videos
- Domains
- IP Addresses
- Business & Government
- Vehicles
- Virtual Currencies
- Breaches & Leaks
- Live Audio Streams
- Live Video Streams
- APIs

Facebook Search Tool

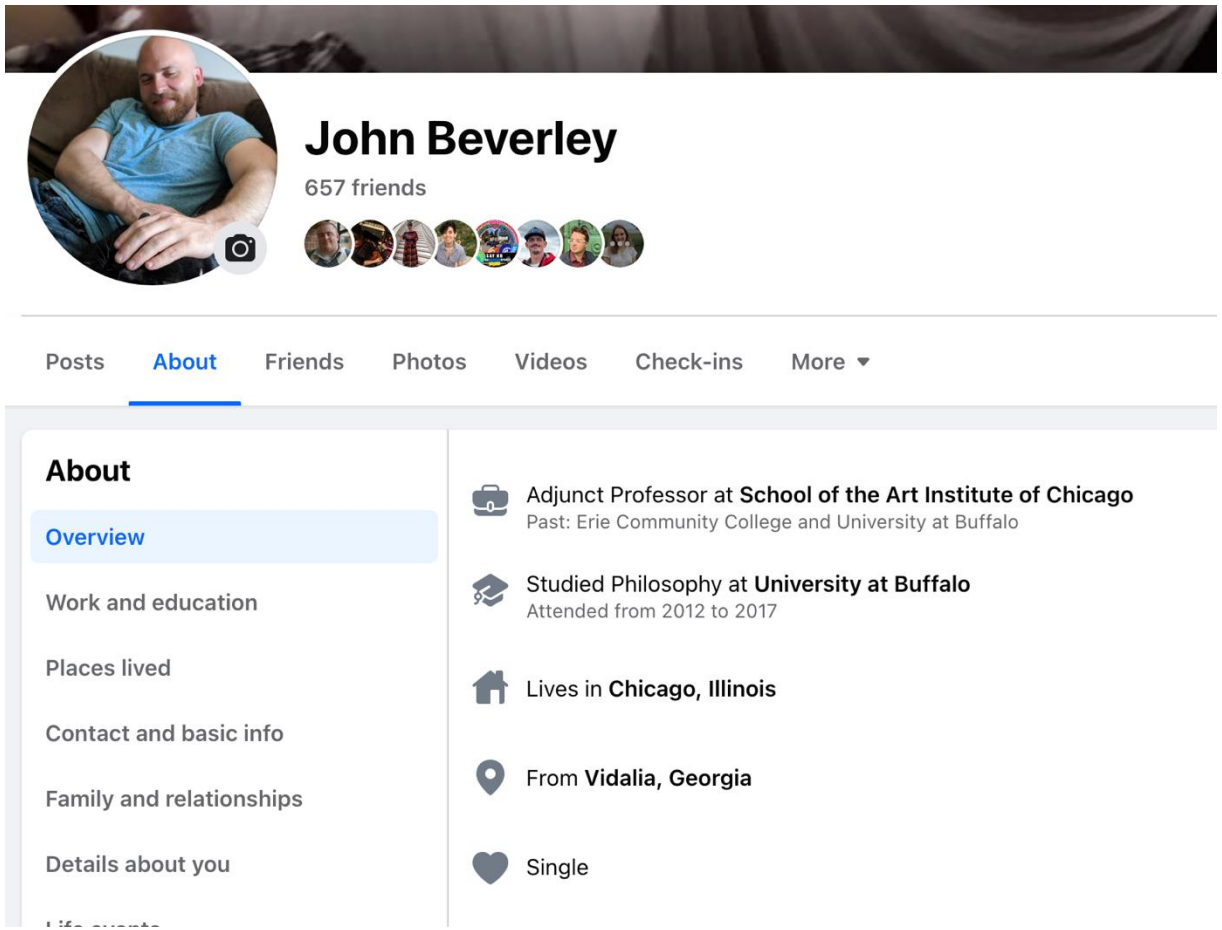
VitaLongaQuoque	Populate All
VitaLongaQuoque	Timeline
VitaLongaQuoque	About
VitaLongaQuoque	Employment
VitaLongaQuoque	Education
VitaLongaQuoque	Locations
VitaLongaQuoque	Contact Info
VitaLongaQuoque	Basic Info
VitaLongaQuoque	Relationships
VitaLongaQuoque	Family
VitaLongaQuoque	Biography
VitaLongaQuoque	Life Events
VitaLongaQuoque	Friends
VitaLongaQuoque	Following
VitaLongaQuoque	Photos
VitaLongaQuoque	Photos Albums
VitaLongaQuoque	Videos
VitaLongaQuoque	Reels
VitaLongaQuoque	Check-ins
VitaLongaQuoque	Visits
VitaLongaQuoque	Recent Check-ins
VitaLongaQuoque	Sports

Tools

- Search Engines
- Facebook
- Twitter
- Instagram
- LinkedIn
- Communities
- Email Addresses
- Usernames
- Names
- Addresses
- Telephone Numbers
- Maps
- Documents
- Pastes
- Images
- Videos
- Domains
- IP Addresses
- Business & Government
- Vehicles
- Virtual Currencies
- Breaches & Leaks
- Live Audio Streams
- Live Video Streams
- APIs

Facebook Search Tool

VitaLongaQuoque	Populate All
VitaLongaQuoque	Timeline
VitaLongaQuoque	About
VitaLongaQuoque	Employment
VitaLongaQuoque	Education
VitaLongaQuoque	Locations
VitaLongaQuoque	Contact Info
VitaLongaQuoque	Basic Info
VitaLongaQuoque	Relationships
VitaLongaQuoque	Family
VitaLongaQuoque	Biography
VitaLongaQuoque	Life Events
VitaLongaQuoque	Friends
VitaLongaQuoque	Following
VitaLongaQuoque	Photos
VitaLongaQuoque	Photos Albums
VitaLongaQuoque	Videos
VitaLongaQuoque	Reels
VitaLongaQuoque	Check-ins
VitaLongaQuoque	Visits
VitaLongaQuoque	Recent Check-ins
VitaLongaQuoque	Sports

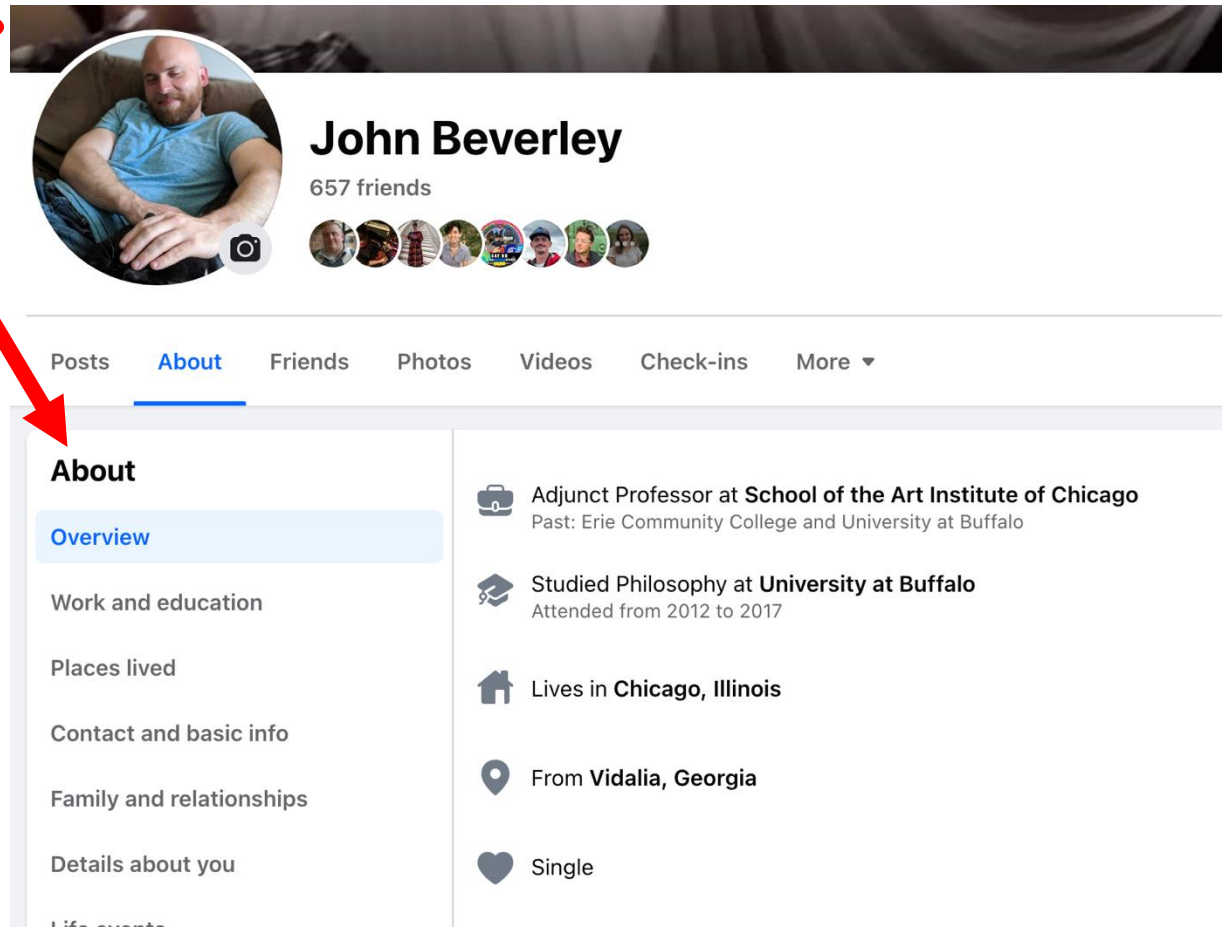


Tools

- Search Engines
- Facebook
- Twitter
- Instagram
- LinkedIn
- Communities
- Email Addresses
- Usernames
- Names
- Addresses
- Telephone Numbers
- Maps
- Documents
- Pastes
- Images
- Videos
- Domains
- IP Addresses
- Business & Government
- Vehicles
- Virtual Currencies
- Breaches & Leaks
- Live Audio Streams
- Live Video Streams
- APIs

Facebook Search Tool

VitaLongaQuoque	Populate All
VitaLongaQuoque	Timeline
VitaLongaQuoque	About
VitaLongaQuoque	Employment
VitaLongaQuoque	Education
VitaLongaQuoque	Locations
VitaLongaQuoque	Contact Info
VitaLongaQuoque	Basic Info
VitaLongaQuoque	Relationships
VitaLongaQuoque	Family
VitaLongaQuoque	Biography
VitaLongaQuoque	Life Events
VitaLongaQuoque	Friends
VitaLongaQuoque	Following
VitaLongaQuoque	Photos
VitaLongaQuoque	Photos Albums
VitaLongaQuoque	Videos
VitaLongaQuoque	Reels
VitaLongaQuoque	Check-ins
VitaLongaQuoque	Visits
VitaLongaQuoque	Recent Check-ins
VitaLongaQuoque	Sports

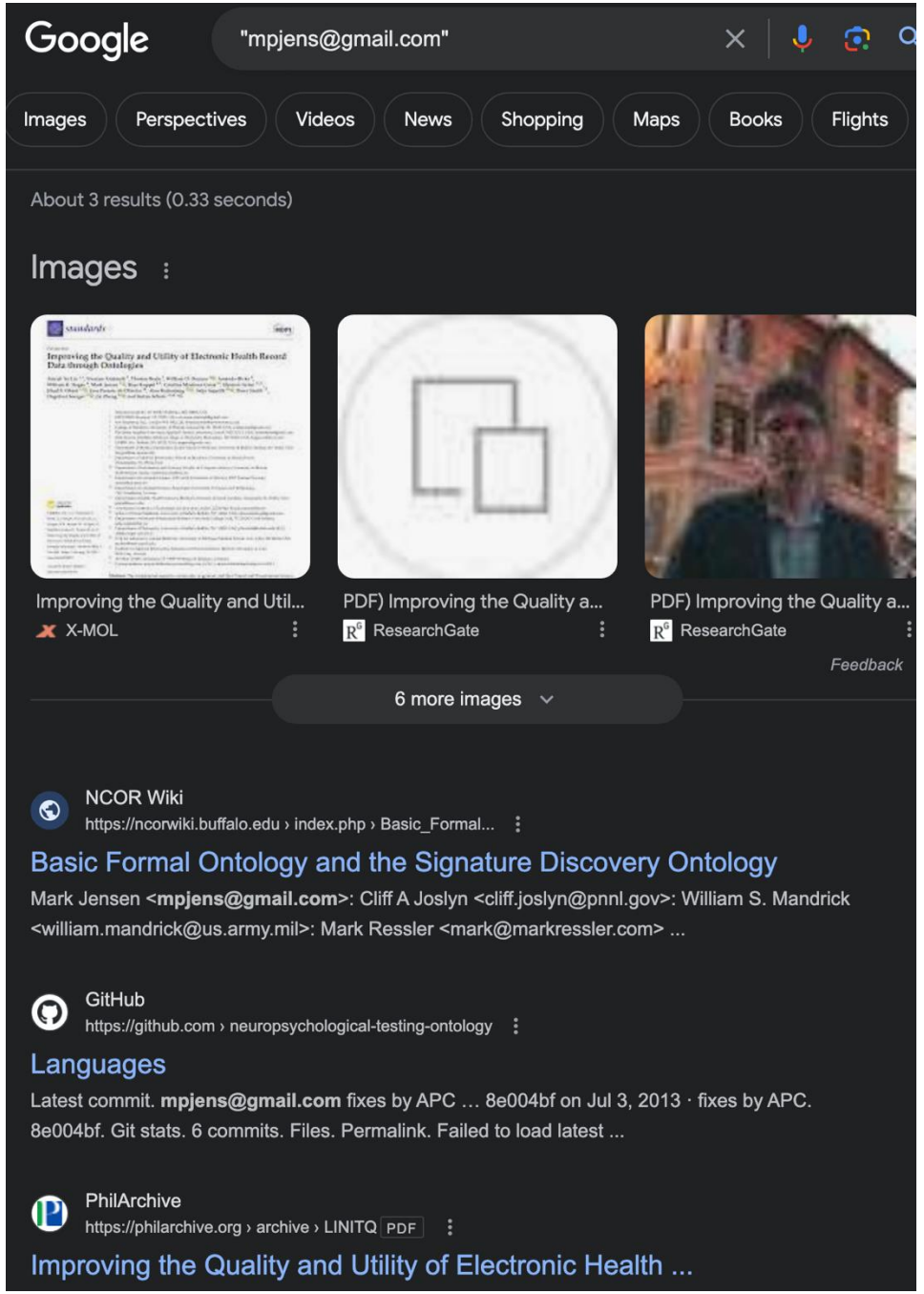


Tools

- Search Engines
- Facebook
- Twitter
- Instagram
- LinkedIn
- Communities
- Email Addresses
- Usernames
- Names
- Addresses
- Telephone Numbers
- Maps
- Documents
- Pastes
- Images
- Videos
- Domains
- IP Addresses
- Business & Government
- Vehicles
- Virtual Currencies
- Breaches & Leaks
- Live Audio Streams
- Live Video Streams
- APIs

Email Addresses Search Tool

mpjens@gmail.com	Populate All
mpjens@gmail.com	Google
mpjens@gmail.com	Bing
mpjens@gmail.com	Yandex
mpjens@gmail.com	Emailrep
mpjens@gmail.com	PingUtil
mpjens@gmail.com	Gravatar

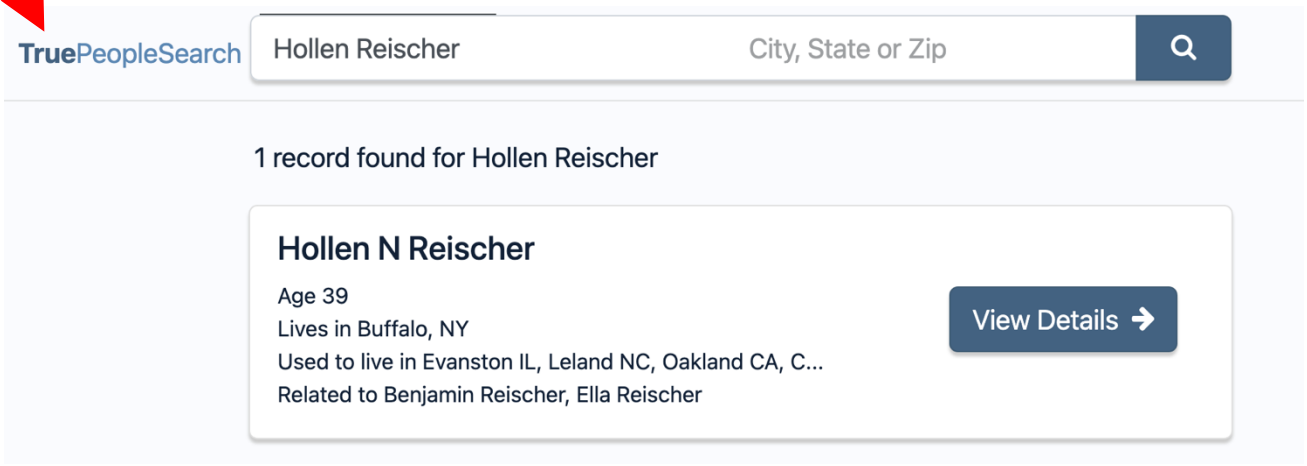


Tools

- Search Engines
- Facebook
- Twitter
- Instagram
- LinkedIn
- Communities
- Email Addresses
- Usernames
- Names
- Addresses
- Telephone Numbers
- Maps
- Documents
- Pastes
- Images
- Videos
- Domains
- IP Addresses
- Business & Government
- Vehicles
- Virtual Currencies
- Breaches & Leaks
- Live Audio Streams
- Live Video Streams
- APIs

Names Search Tool

Hollen	Reischer	Populate All
Hollen	Reischer	TruePeople
Hollen	Reischer	FastPeople
Hollen	Reischer	Nuwber
Hollen	Reischer	XLEK
Hollen	Reischer	FamilyTreeNow
Hollen	Reischer	Intelius
Hollen	Reischer	Radaris
Hollen	Reischer	CyberBackground
Hollen	Reischer	Spytox
Hollen	Reischer	SearchPeople
Hollen	Reischer	Spokeo
Hollen	Reischer	AdvBackground
Hollen	Reischer	Yasni
Hollen	Reischer	Zabasearch
Hollen	Reischer	PeopleSearchNow
Hollen	Reischer	WebMii
Hollen	Reischer	SocialSearcher
Hollen	Reischer	TruthFinder
Hollen	Reischer	PeopleByName
Hollen	Reischer	White Pages
Hollen	Reischer	Thats Them
Hollen	Reischer	ClustrMaps



Group Exercise

- Because I know you're not going to be able to resist, navigate to the "OSINT" website and explore what sorts of information is available about you online

- You're welcome to discuss in small groups

Keep in mind, we'll be discussing as a class

Outline

- Cybersecurity Landscape
- Proliferation of Knowledge Representation
- Addressing the Alignment Problem
- Design Pattern Practice

Need for Speed

- Hackers need 15 hours to compromise a system; defenders need 200-300 days to discover a breach

- Identifying vulnerabilities and potential exploits in software and hardware early and often is of paramount importance

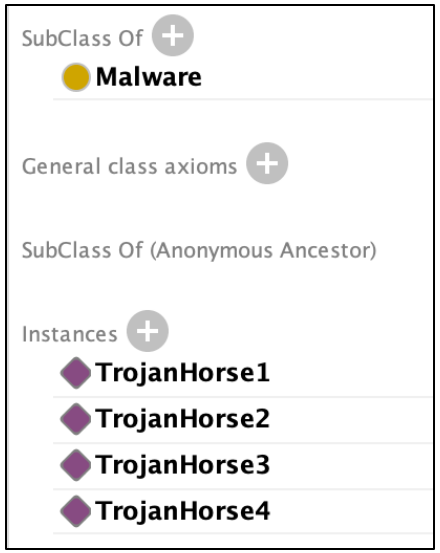
- The variety and complexity of cybersecurity data strikes me as analogous to the human genome project...

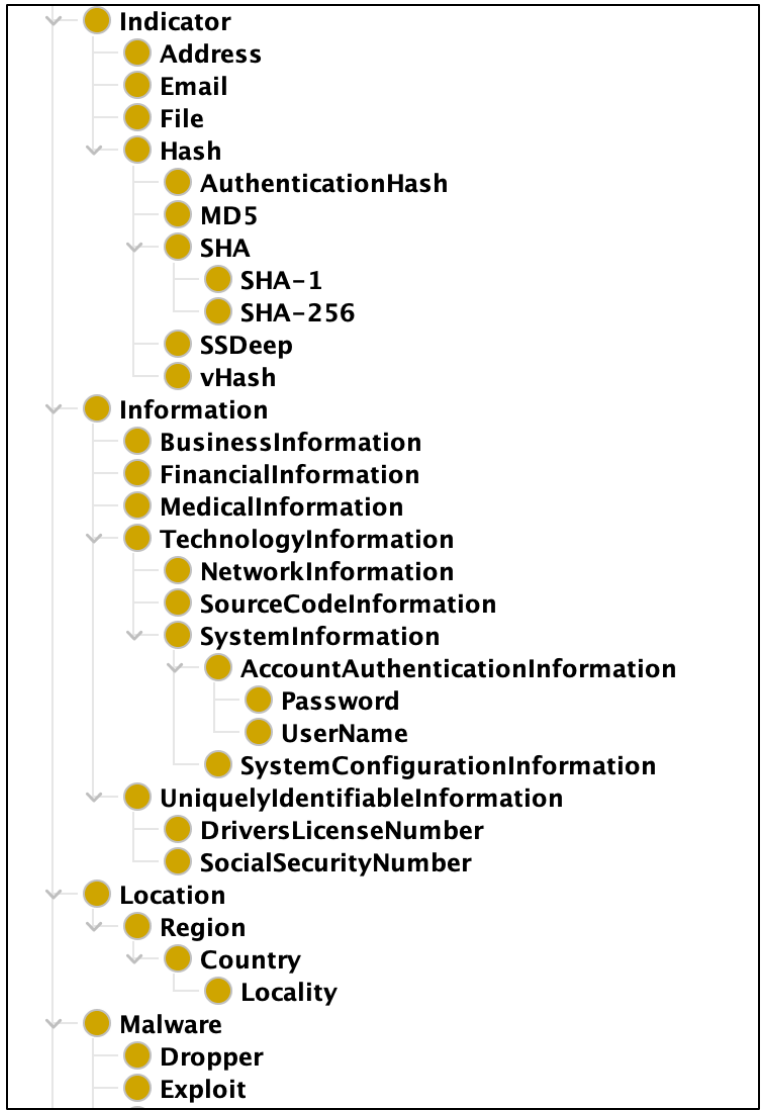
Proliferation of Ontologies

- When developed correctly, ontologies provide **common vocabularies** with **common semantics** across **multiple domains**
- The success of the Gene Ontology led to a proliferation of ontologies developed by subject-matter experts, computer scientists, and logicians
- Almost **none** of which were developed in coordination
- The result was **massive incompatibility** of terms and relations, confusion, in-fighting, name-calling, etc.

The Malware Ontology (MalOnt)

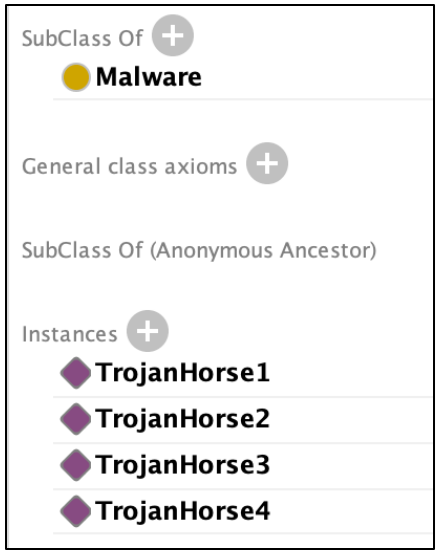
- Developed by Nidhi Ristogi at RIT
- Designed to provide design patterns for common malware attacks and defense postures...
- I guess...it's hard to say since there's no description of the scope in the ontology

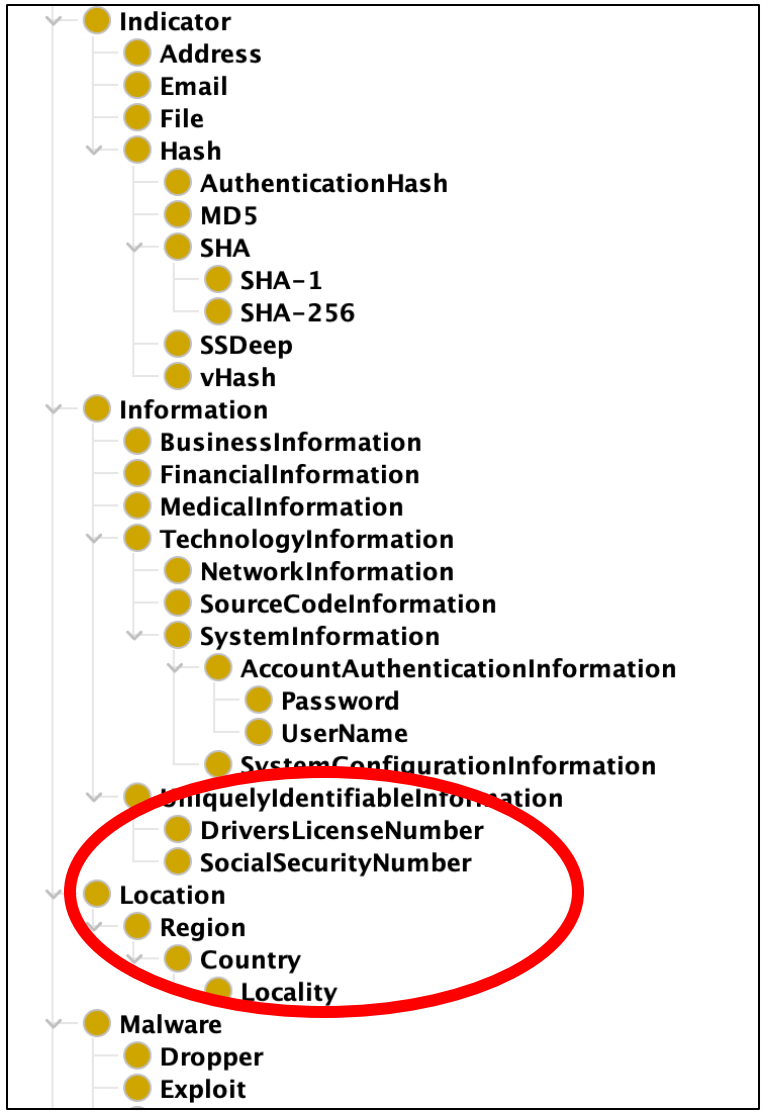




The Malware Ontology (MalOnt)

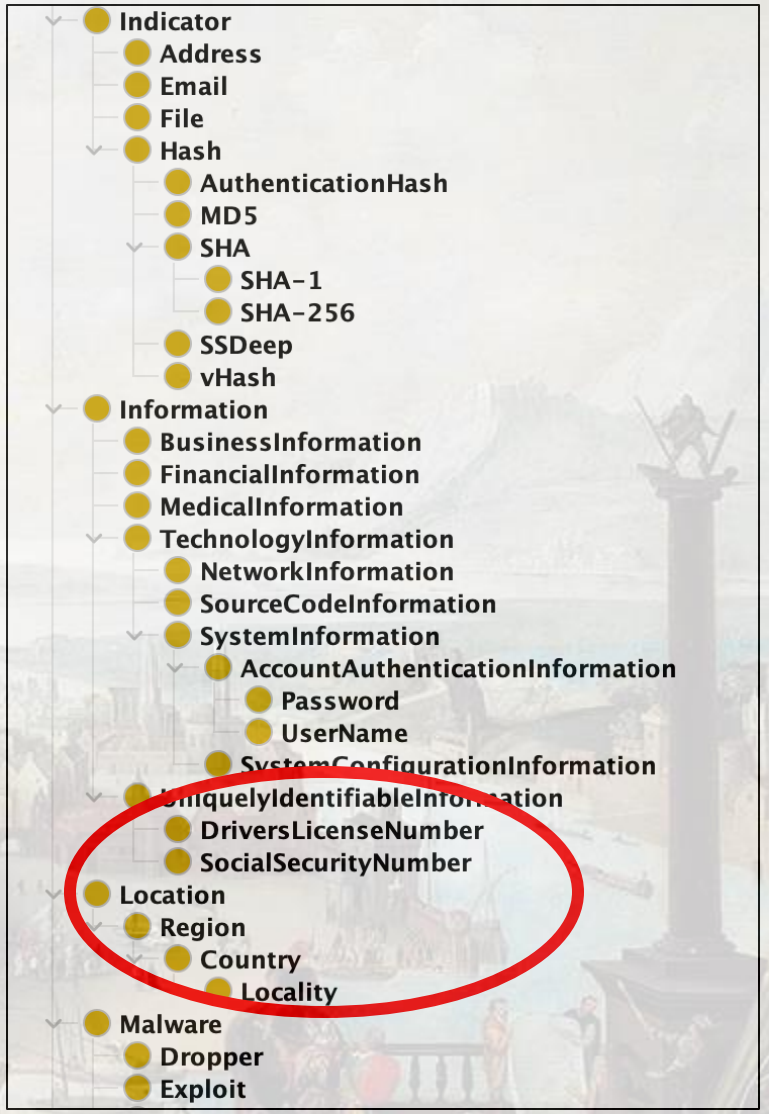
- Developed by Nidhi Ristogi at RIT
- Designed to provide design patterns for common malware attacks and defense postures...
- I guess...it's hard to say since there's no description of the scope in the ontology

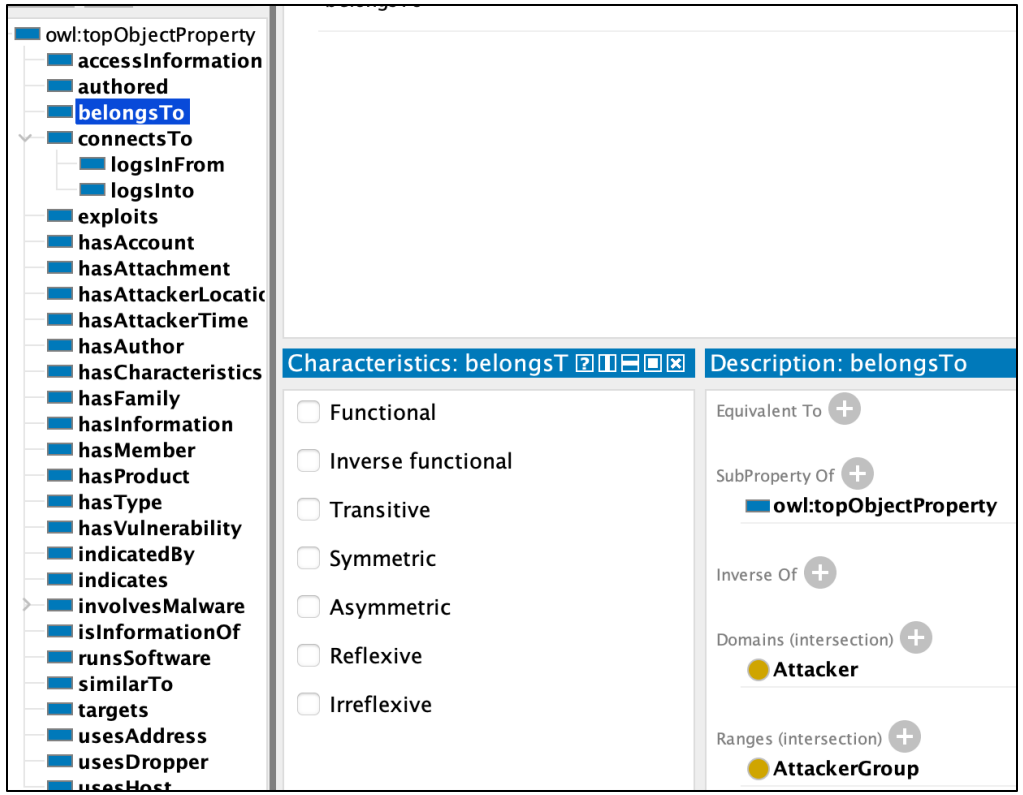


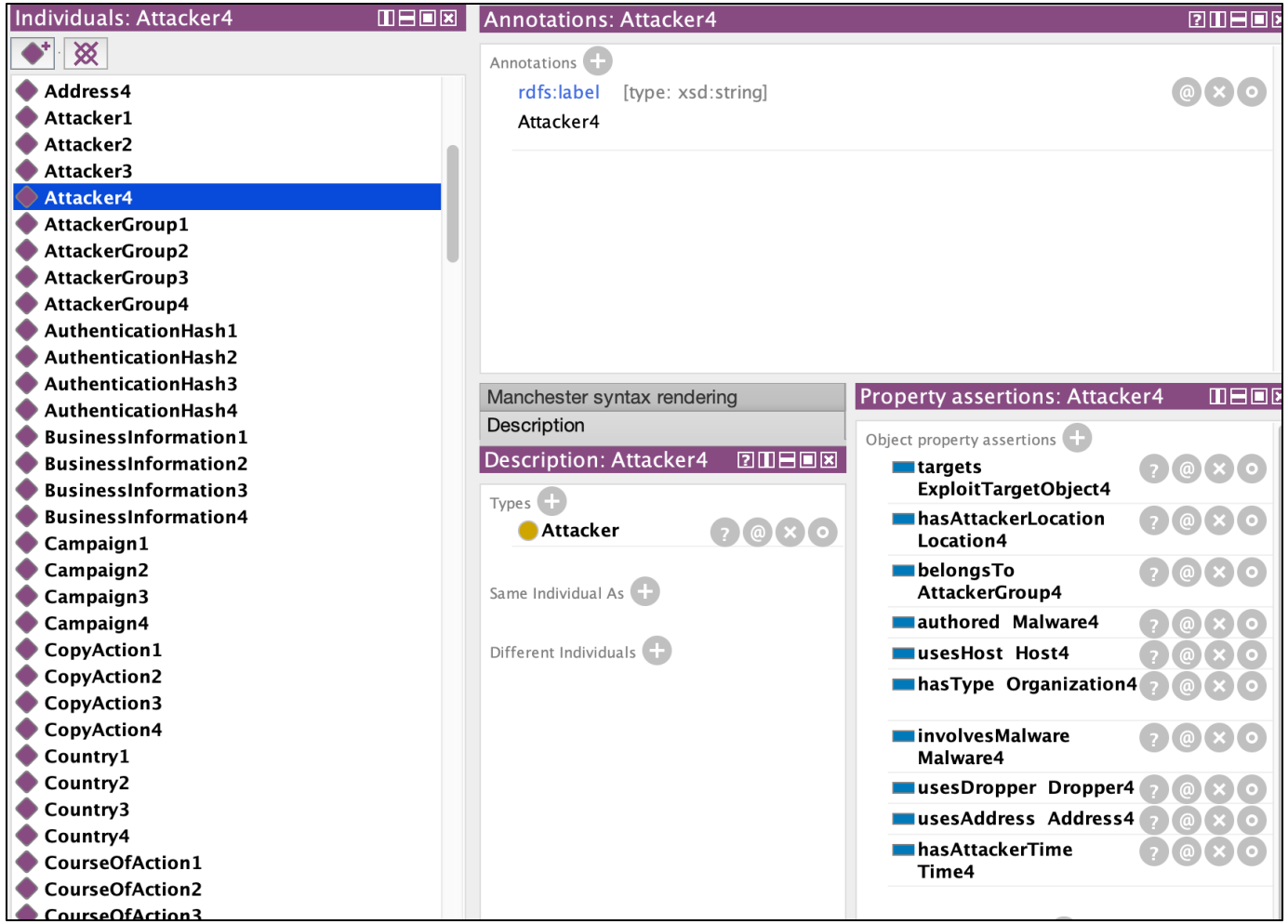


The Malware Ontology (MalOnt)

- Developed by Nidhi Ristogi at RIT
- Designed to provide design patterns for common malware attacks and defense postures...
- I guess...it's hard to say since there's no description of the scope in the ontology

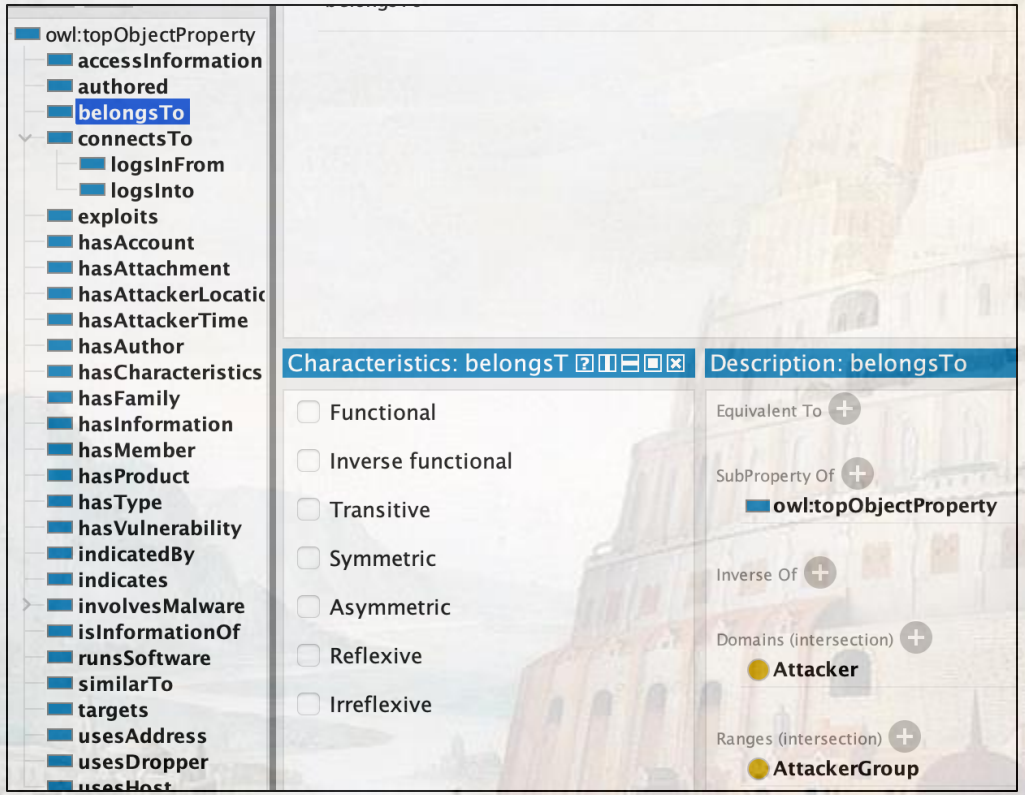


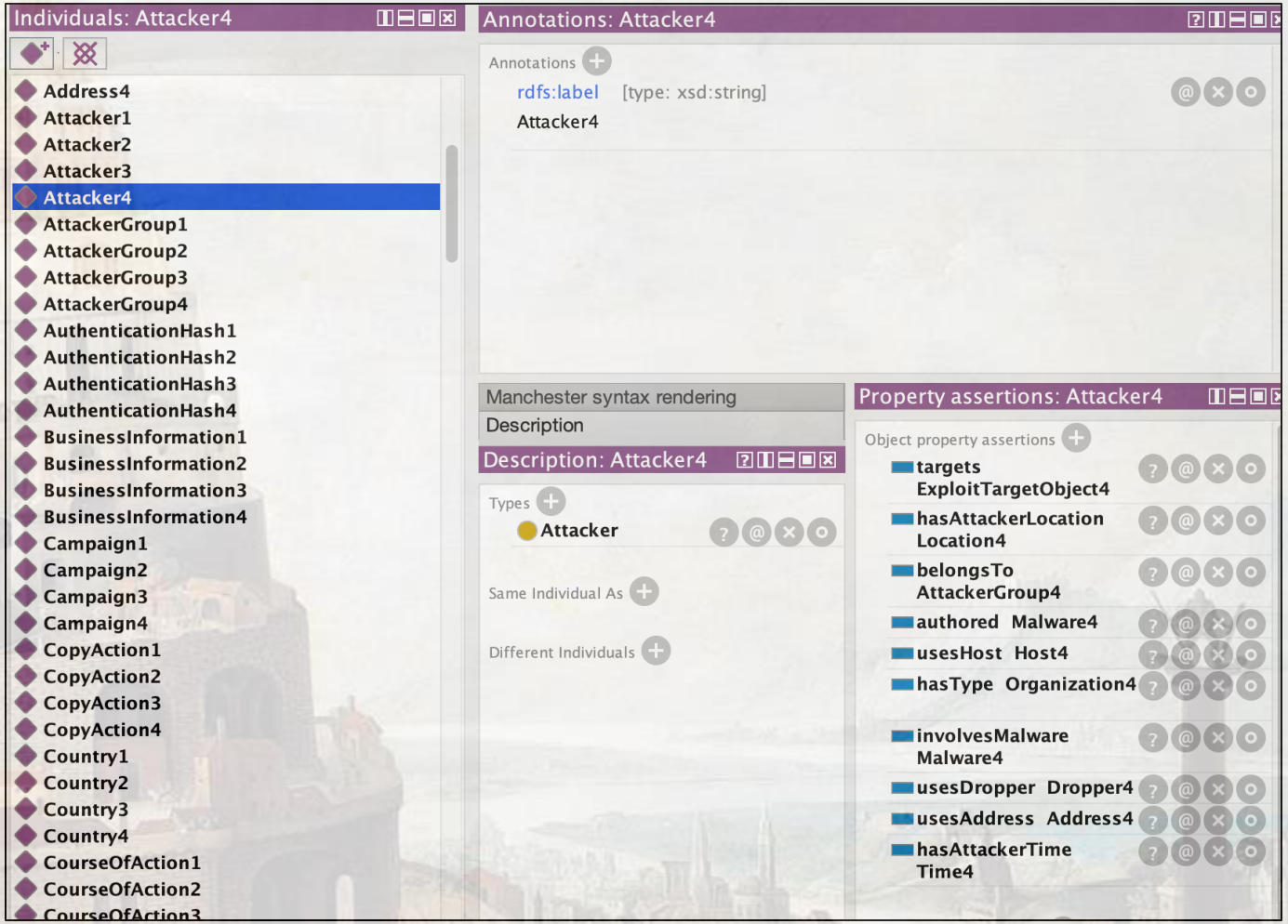




No Formal Properties

No Justification or Descriptions





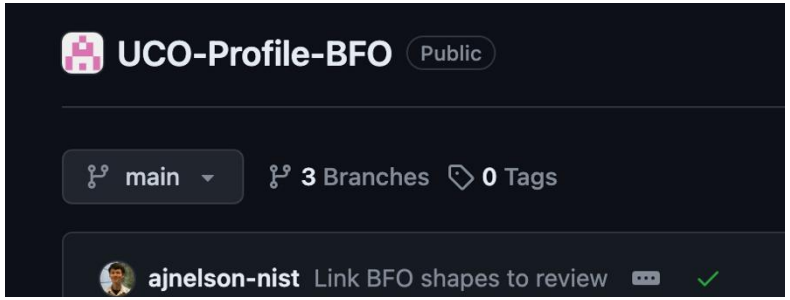
No Formal Properties

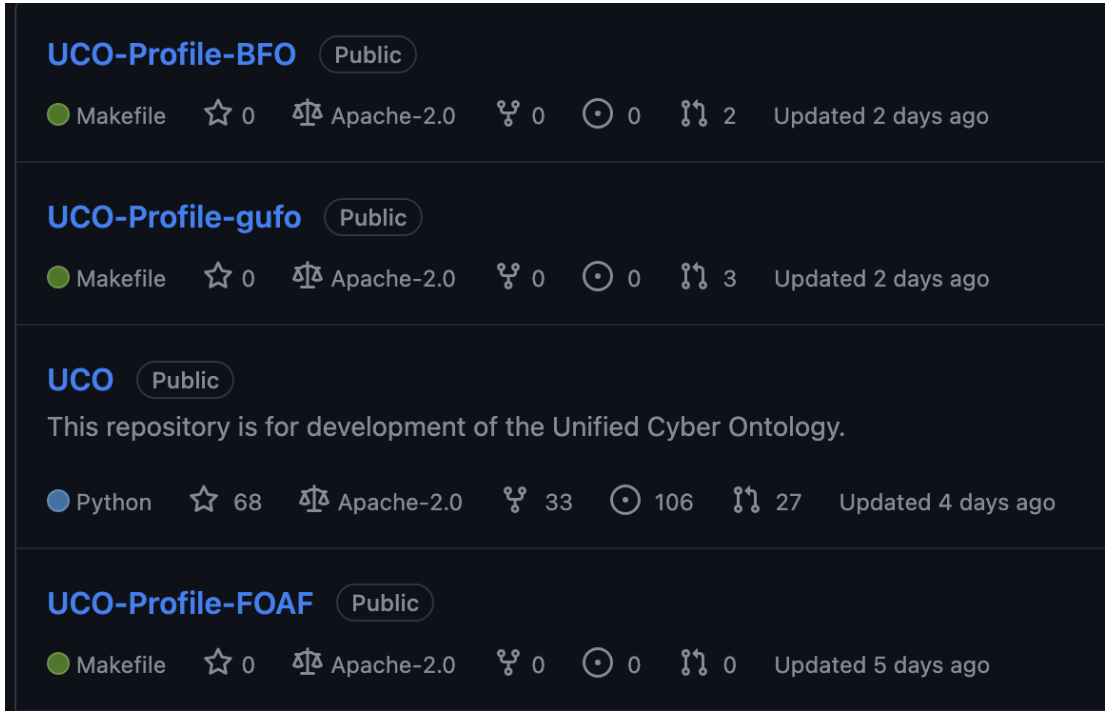
No Justification or Descriptions

Unified Cyber Ontology (UCO)

- Designed to integrate heterogeneous cyber datasets and provide a foundation for mapping cyber data to Semantic Web resources, e.g. Linked Open Data, DBPedia, Yago Knowledge Base

- **Also mapping to top-level ontologies**

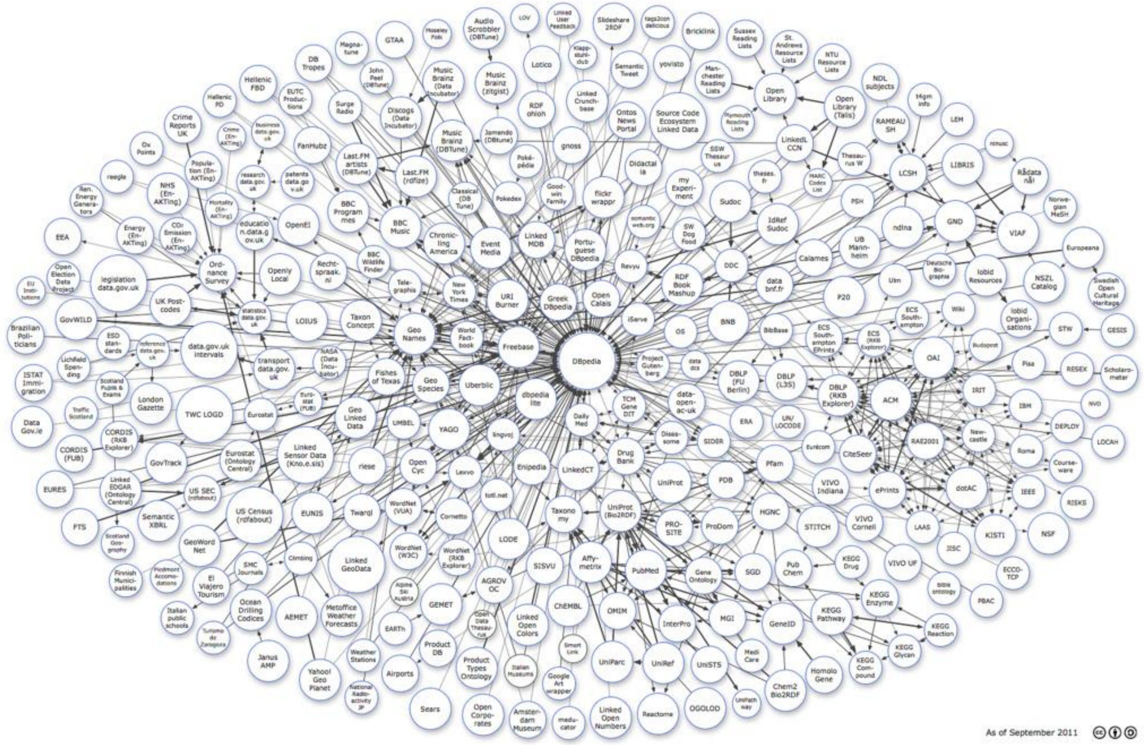




Unified Cyber Ontology (UCO)

- Designed to integrate heterogeneous cyber datasets and provide a foundation for mapping cyber data to Semantic Web resources, e.g. Linked Open Data, DBPedia, Yago Knowledge Base

- However, **semantic web resources vary widely in their coherence**.



Yago Knowledge Base

- Yago extracts content from Wikipedia's crowd sourced pages, which is then automatically annotated with ontology terms

- Illustrative examples:
  - An elephants' graveyard is a mammal
  - A Lockheed Martin A2100 is a person
  - The Berneuse Mountain is a cable car

- UCO coordinates cyber data with data from other domains but the results cannot be trusted

D3FEND™

- Knowledge graph of cybersecurity countermeasure techniques

- Project aims to **standardize defensive cybersecurity technology vocabulary** in a semantic graph representing computer system components and relations to defensive and offensive cyber techniques

MITRE

https://github.com/d3fend

D3FEND™

Detection, Denial, and Disruption Framework Empowering Network Defense

Foundations

- The D3FEND knowledge graph **defines** key terms and relations concerning cybersecurity countermeasures

- Definitions are based on **over 500** countermeasure patents from the U.S. Patent Office between 2001-2018

- Graph supports **queries** connecting cybersecurity countermeasures to **offensive tactics**, **techniques**, and **procedures**

DEFEND™

A knowledge graph of cybersecurity countermeasures
0.9.2-BETA-3

Harden

Application Hardening
- Dead Code Elimination (1)
- Exception Handler Pointer Validation (1)
- Process Segment Execution Prevention (2)
- Segment Address Offset Randomization (2)
- Stack Frame Canary Verification (2)
- Pointer Authentication (2)

Credential Hardening
- Certificate Pinning (2)
- Multi-factor Authentication (1)
- One-time Password (1)
- Strong Password Policy (1)

Message Hardening
- Message Authentication (2)
- Message Encryption (1)
- Transfer Agent Authentication (3)

Platform Hardening
- Disk Encryption (1)
- Driver Load Integrity Checking (2)
- TPM Boot Integrity (3)

Detect

File Analysis
- Dynamic Analysis (2)
- Emulated File Analysis (1)
- File Content Rules (4)
- File Hashing (1)

Identifier Analysis
- Homoglyph Detection (2)
- URL Analysis (2)

Message Analysis
- Sender MTA Reputation Analysis (1)
- Sender Reputation Analysis (1)

Network Traffic Analysis
- Administrative Network Activity Analysis (3)
- Certificate Analysis (1)
- Active Certificate Analysis (1)
- Passive Certificate Analysis (2)
- Client-server Payload Profiling (1)
- DNS Traffic Analysis (5)
- File Carving (1)
- IPC Traffic Analysis (6)
- Network Traffic Community Deviation (1)
- Per Host Download-Upload Ratio Analysis (1)
- Protocol Metadata Anomaly Detection (3)
- Remote Terminal Session Detection (3)
- RPC Traffic Analysis (7)
- Connection Attempt Analysis (1)
- Inbound Session

Platform Monitoring
- Firmware Verification (3)
- Operating System Monitoring (2)
- Endpoint Health Beacon (1)
- Input Device Analysis (2)
- Memory Boundary Tracking (1)
- Scheduled Job Analysis (3)
- System Daemon Monitoring (3)
- System File Analysis (3)
- Service Binary Verification (1)
- Firmware Behavior Analysis (2)
- Firmware Embedded Monitoring Code (2)

Process Analysis
- Database Query String Analysis (1)
- File Access Pattern Analysis (1)
- Indirect Branch Call Analysis (1)
- Process Code Segment Verification (7)
- Process Self-Modification Detection (1)
- Process Spawn Analysis (17)
- Process Lineage Analysis (13)
- Script Execution Analysis (1)
- Shadow Stack Comparisons (1)
- System Call Analysis (4)

User Behavior Analysis
- Authentication Event Thresholding (6)
- Authorization Event Thresholding (4)
- Job Function Access Pattern Analysis (1)
- Resource Access Pattern Analysis (5)
- User Data Transfer Analysis (2)
- User Geolocation Logon Pattern Analysis (2)
- Web Session Activity Analysis (4)
- Session Duration Analysis (2)

Isolate

Execution Isolation
- Hardware-based Process Isolation (3)
- Mandatory Access Control (2)
- Executable Denylisting (2)
- Executable Allowlisting (2)

Network Isolation
- Broadcast Domain Isolation (2)
- Encrypted Tunnels (1)
- Inbound Traffic Filtering (9)
- Outbound Traffic Filtering (1)
- DNS Allowlisting (1)
- DNS Denylisting (1)
- Forward Resolution Domain Denylisting (1)
- Hierarchical Domain Denylisting (1)
- Homoglyph Denylisting (1)
- Forward Resolution IP Denylisting (1)
- Reverse Resolution IP Denylisting (1)

Deceive

Decoy Environment
- Connected Honeynet (1)
- Integrated Honeynet (1)
- Standalone Honeynet (1)

Decoy Object
- Decoy File (4)
- Decoy Network Resource (4)
- Decoy Persona (2)
- Decoy Public Release (1)
- Decoy Session Token (1)
- Decoy User Credential (3)

Evict

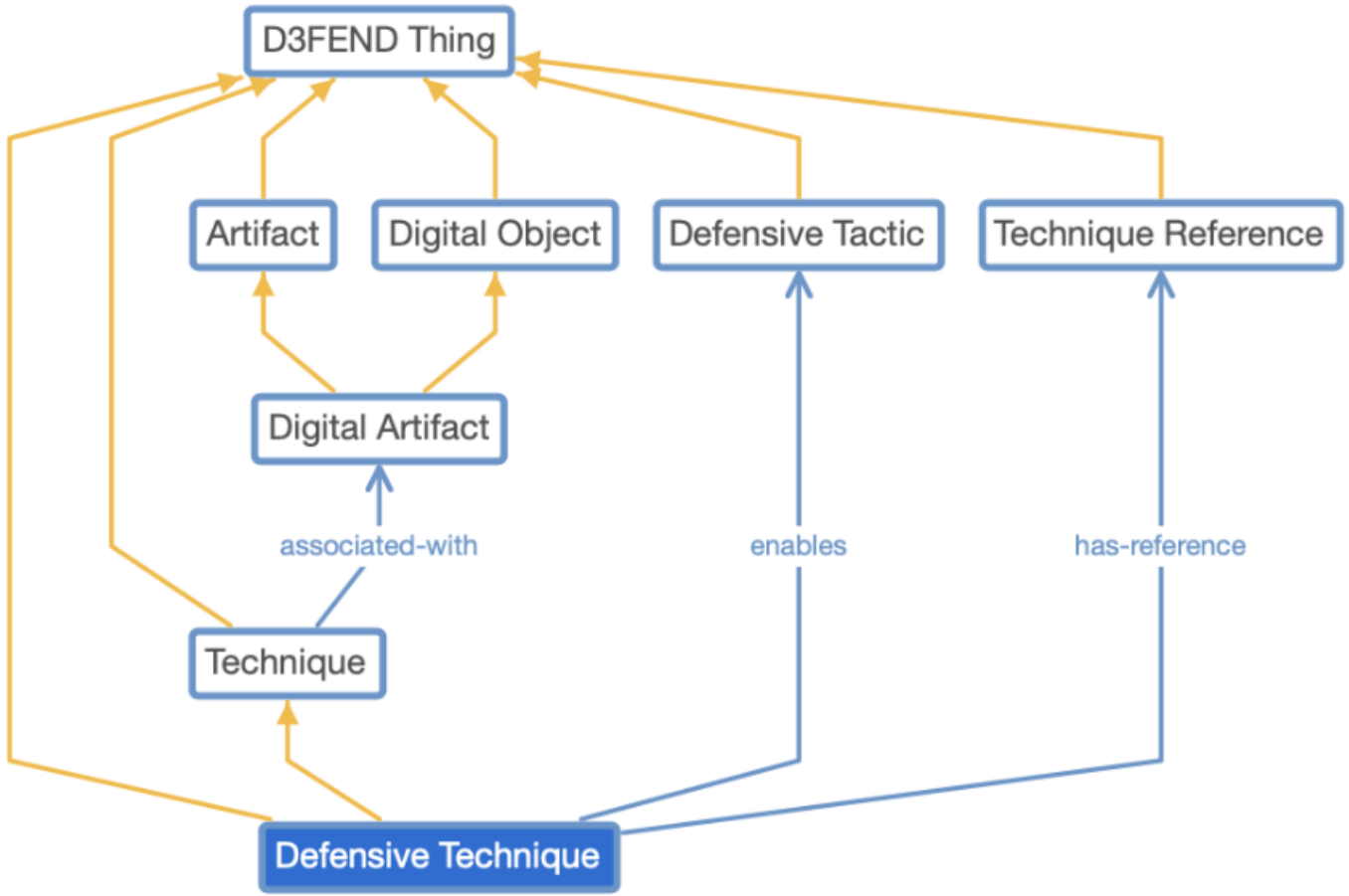
Credential Eviction
- Account Locking (2)
- Authentication Cache Invalidation (2)

Process Eviction
- Process Termination (3)

Metrics	
Axiom	22394
Logical axiom count	7543
Declaration axioms count	4155
Class count	2986
Object property count	211
Data property count	41
Individual count	886
Annotation Property count	37

Class axioms	
SubClassOf	4480
EquivalentClasses	0
DisjointClasses	9
GCI count	0
Hidden GCI Count	0

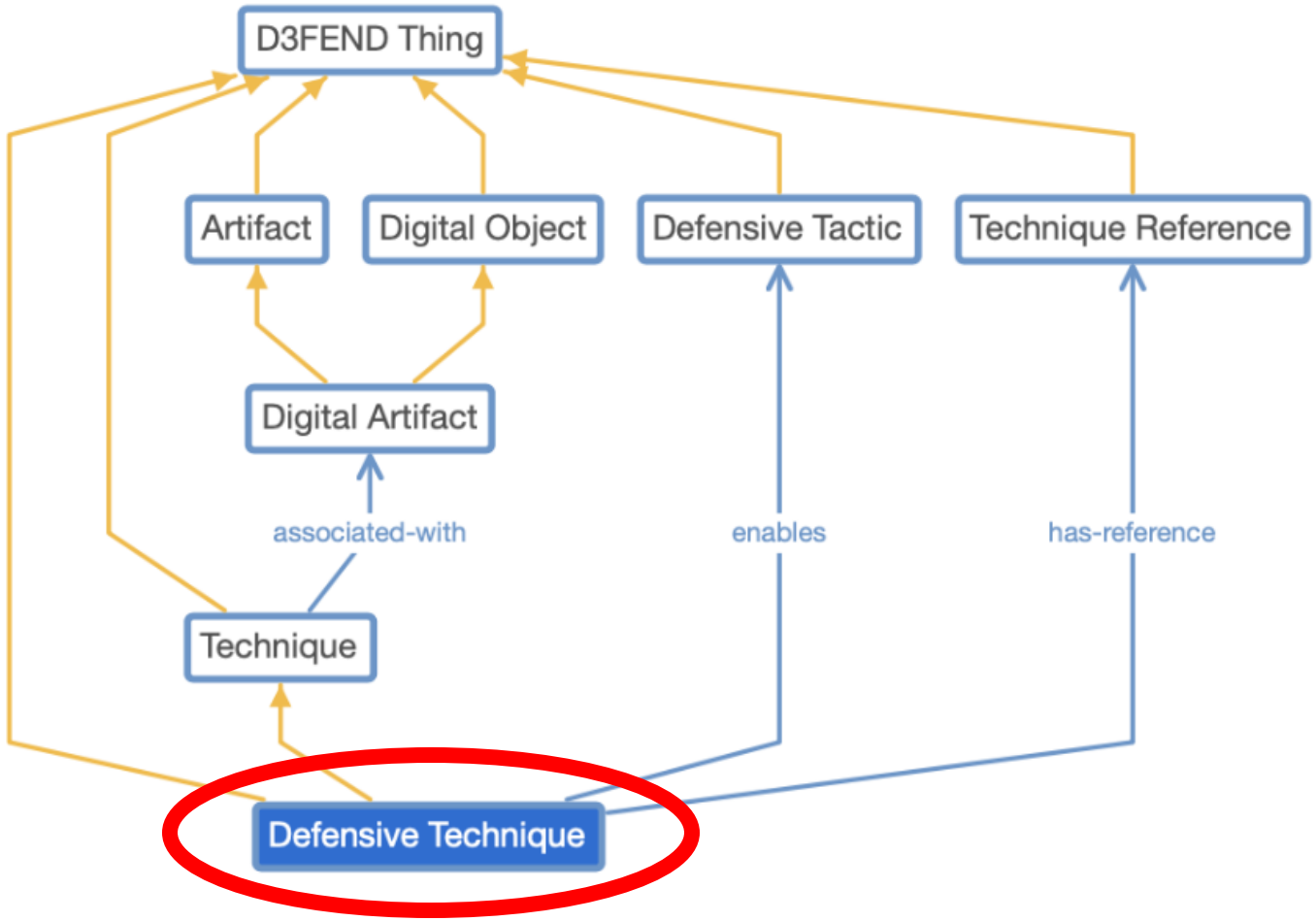
Object property axioms	
SubObjectPropertyOf	257
EquivalentObjectProperties	0
InverseObjectProperties	41
DisjointObjectProperties	0
FunctionalObjectProperty	0
InverseFunctionalObjectProperty	0
TransitiveObjectProperty	5
SymmetricObjectProperty	0
AsymmetricObjectProperty	0
ReflexiveObjectProperty	0
IrrefexiveObjectProperty	0
ObjectPropertyDomain	10
ObjectPropertyRange	17
SubPropertyChainOf	1

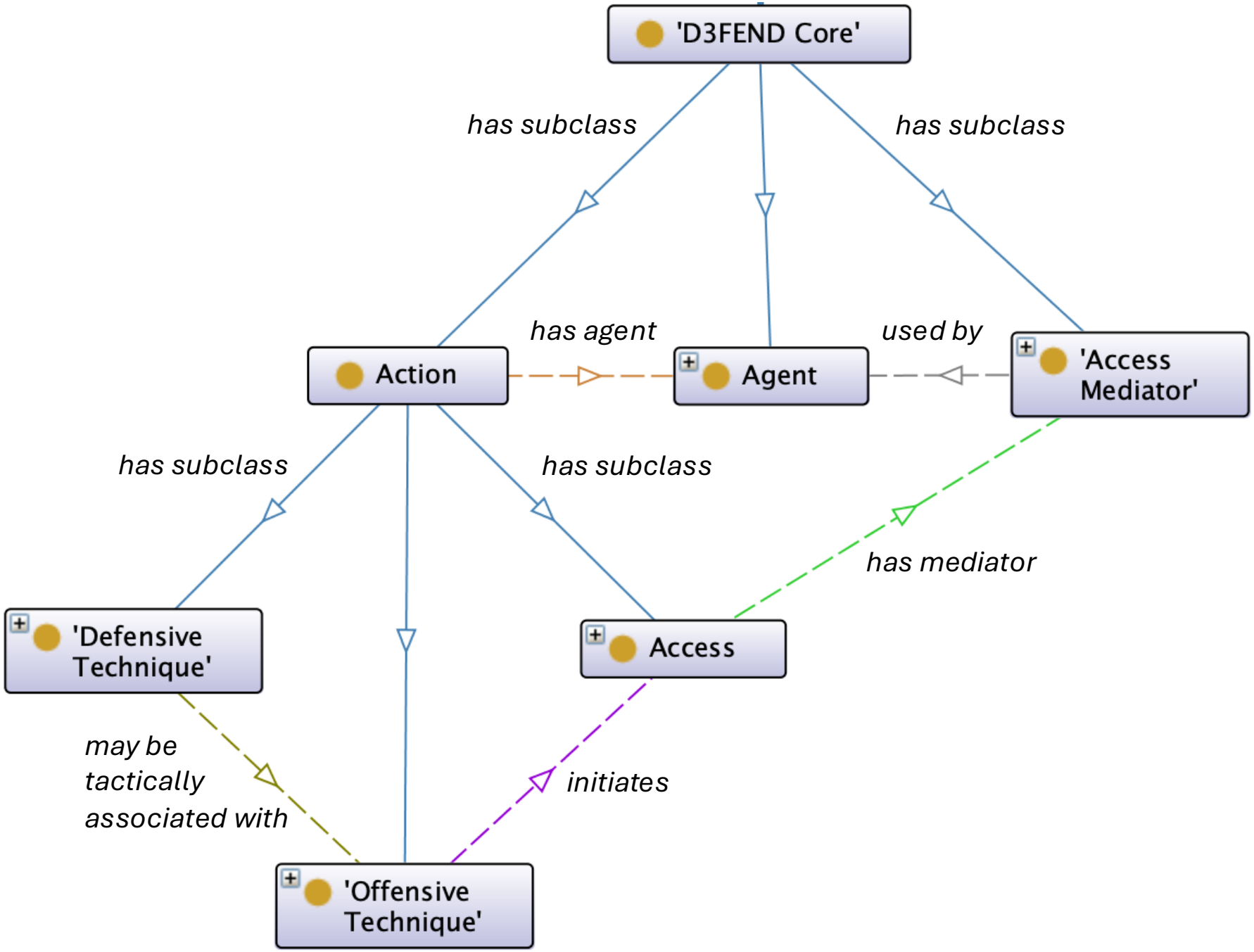


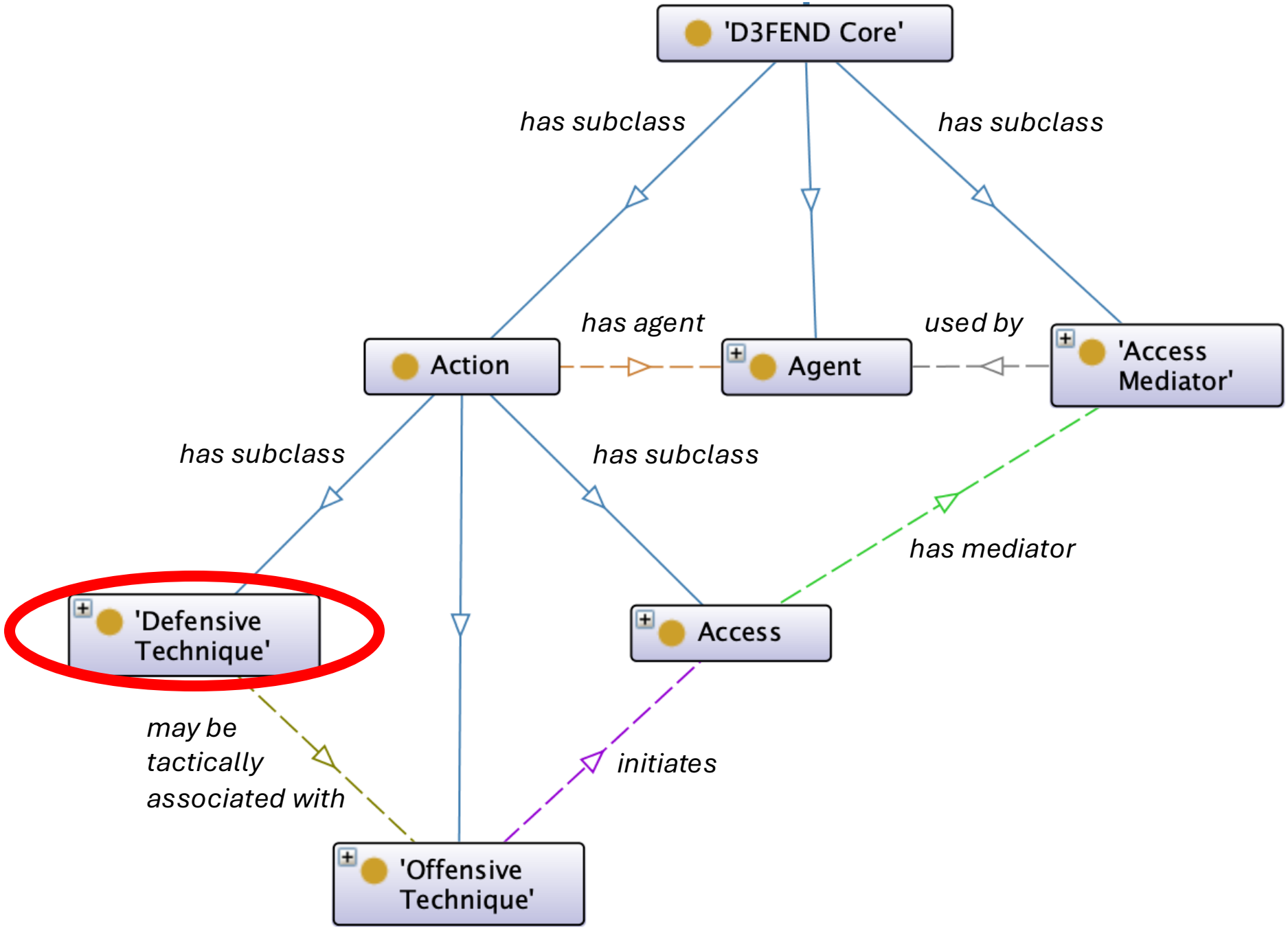
Metrics	
Axiom	22394
Logical axiom count	7543
Declaration axioms count	4155
Class count	2986
Object property count	211
Data property count	41
Individual count	886
Annotation Property count	37

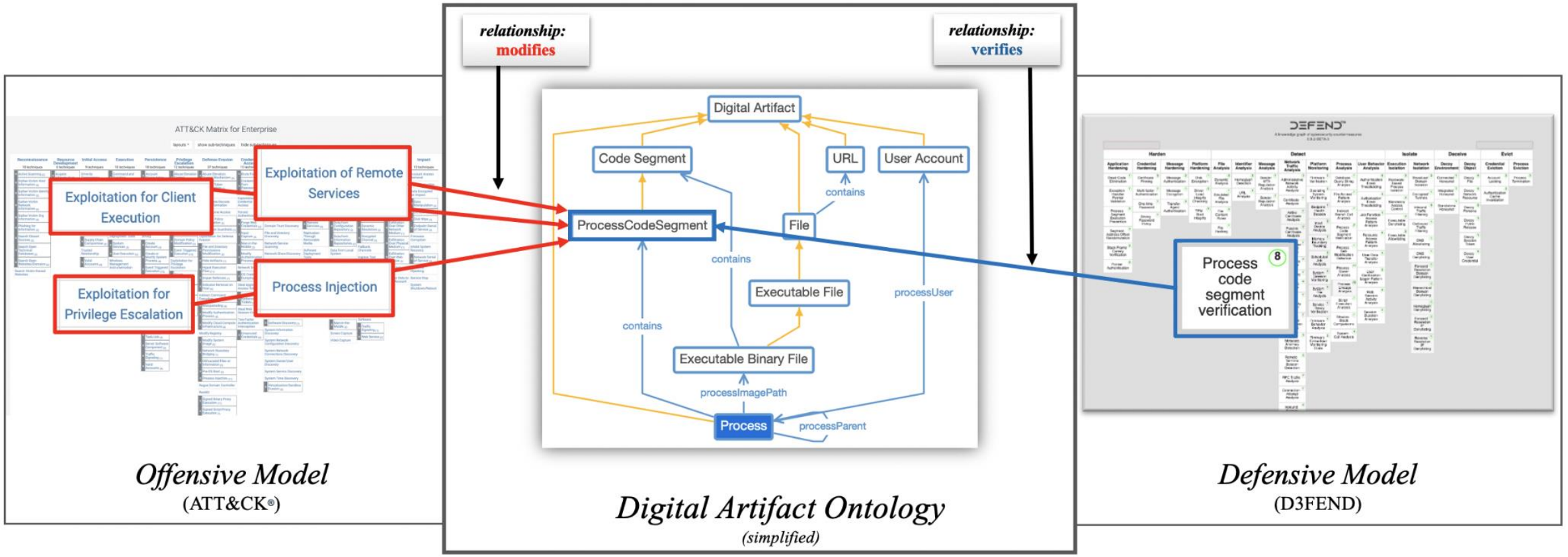
Class axioms	
SubClassOf	4480
EquivalentClasses	0
DisjointClasses	9
GCI count	0
Hidden GCI Count	0

Object property axioms	
SubObjectPropertyOf	257
EquivalentObjectProperties	0
InverseObjectProperties	41
DisjointObjectProperties	0
FunctionalObjectProperty	0
InverseFunctionalObjectProperty	0
TransitiveObjectProperty	5
SymmetricObjectProperty	0
AsymmetricObjectProperty	0
ReflexiveObjectProperty	0
IrrefexiveObjectProperty	0
ObjectPropertyDomain	10
ObjectPropertyRange	17
SubPropertyChainOf	1




'D3FEND Core'
has subclass
has subclass
has agent
used by
Action
Agent
'Access Mediator'
has subclass
has subclass
has mediator
'Defensive Technique'
Access
may be tactically associated with
initiates
'Offensive Technique'


'D3FEND Core'
has subclass
has subclass
has agent
used by
Action
Agent
'Access Mediator'
has subclass
has subclass
has mediator
'Defensive Technique'
Access
may be tactically associated with
initiates
'Offensive Technique'



Toward a Knowledge Graph of Cybersecurity Countermeasures

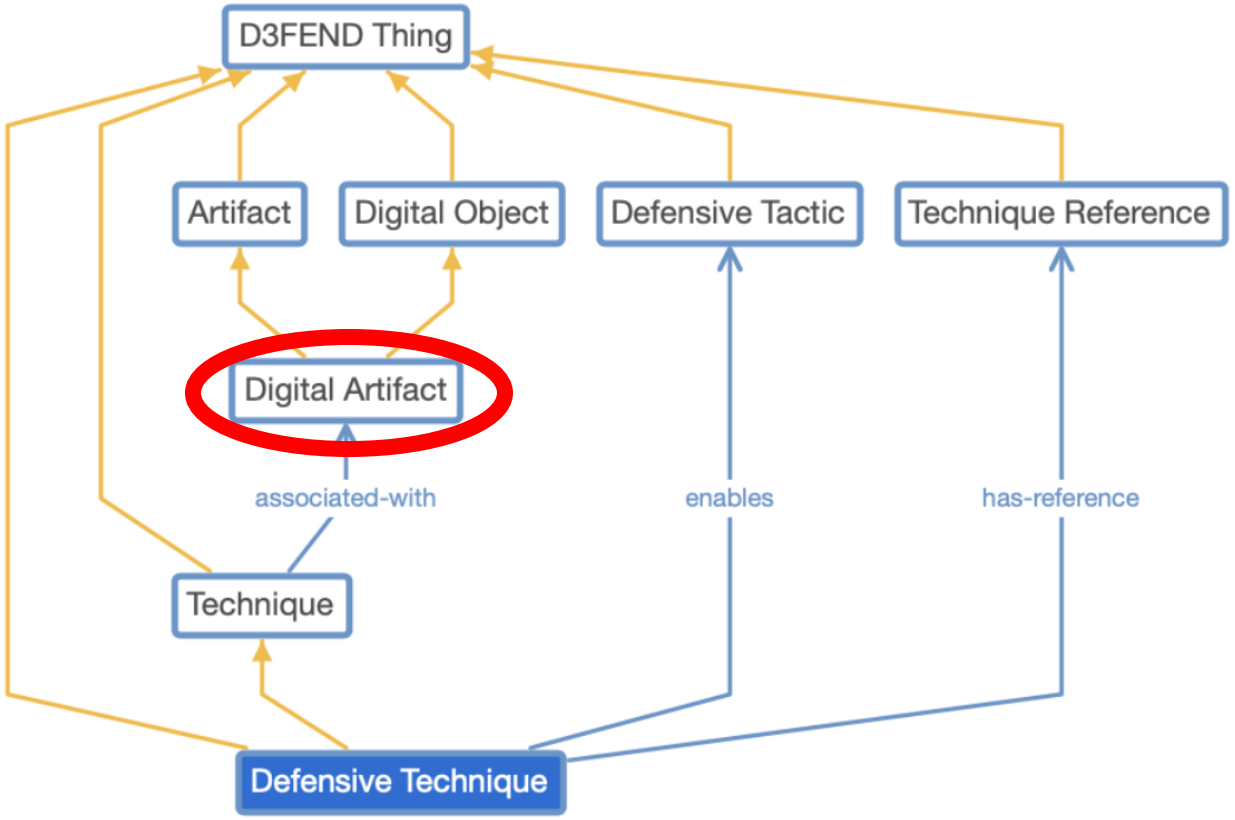
Peter E. Kaloroumakis
The MITRE Corporation
Annapolis Junction, MD
pk@mitre.org

Michael J. Smith
The MITRE Corporation
Annapolis Junction, MD
smithmj@mitre.org

D3FEND

Let's go spelunking...

- D3FEND's longer-term goals are to (1) create a sustain- able knowledge framework for characterizing and relating cybersecurity countermeasure technology; and (2) accelerate knowledge discovery and acquisition efforts required to keep pace with technological changes in the cyber domain.



Data Inventory

Quibble: Definition is about a process, but label suggests actual inventory

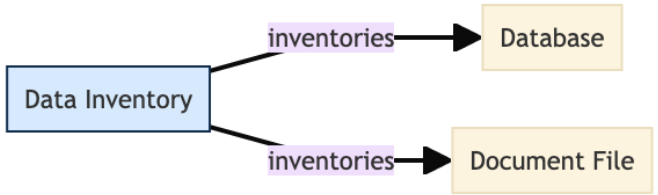
Definition

Data inventorying identifies and records the schemas, formats, volumes, and locations of data stored and used on the organization's architecture.

Synonyms: *Data Discovery* , and *Data Inventorying* .

Digital Artifact Relationships:

This defensive technique is related to specific digital artifacts. Click the artifact node for more information.



Data Inventory

Relationships section suggests data inventorying is a defensive technique…

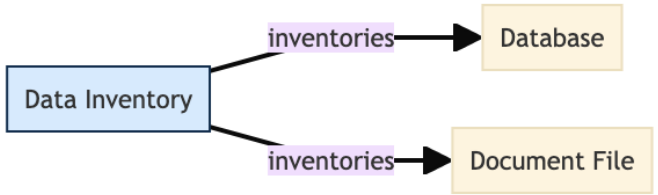
Definition

Data inventorying identifies and records the schemas, formats, volumes, and locations of data stored and used on the organization's architecture.

Synonyms: *Data Discovery* , and *Data Inventorying* .

Digital Artifact Relationships:

This defensive technique is related to specific digital artifacts. Click the artifact node for more information.



Data Inventory

Let's continue our adventure...

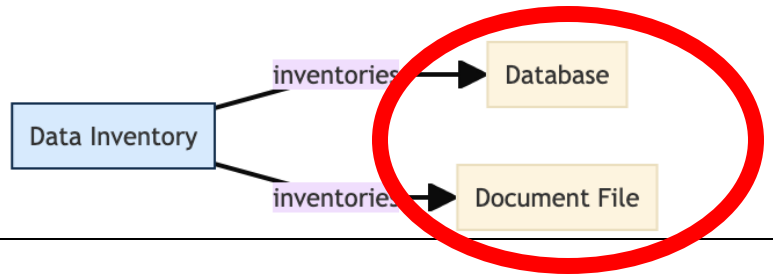
Definition

Data inventorying identifies and records the schemas, formats, volumes, and locations of data stored and used on the organization's architecture.

Synonyms: *Data Discovery* , and *Data Inventorying* .

Digital Artifact Relationships:

This defensive technique is related to specific digital artifacts. Click the artifact node for more information.



Database

- Check it out...

name	Database
definition	A database is an organized collection of data, generally stored and accessed electronically from a computer system. Where databases are more complex they are often developed using formal design and modeling techniques.
defined by	http://dbpedia.org/resource/Database
see also	http://dbpedia.org/resource/Database

Document File

- Describing continuant-like entities rather than processes?

name	Document File
definition	A document is a written, drawn, presented or recorded representation of thoughts. An electronic document file is usually used to describe a primarily textual file, along with its structure and design, such as fonts, colors and additional images.
see also	http://dbpedia.org/resource/Document

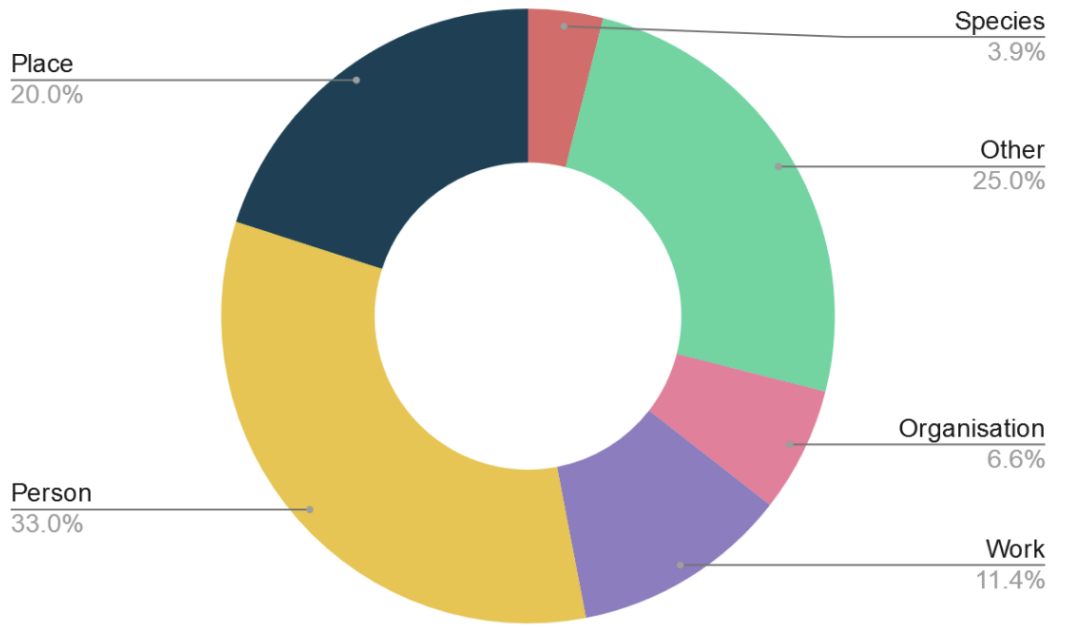
Document File

- Let's continue spelunking...

name	Document File
definition	A document is a written, drawn, presented or recorded representation of thoughts. An electronic document file is usually used to describe a primarily textual file, along with its structure and design, such as fonts, colors and additional images.
see also	http://dbpedia.org/resource/Document

DBpedia

Class	Instances
Resource (overall)	4,828,418
Place	967,491
Person	1,592,912
Work	552,115
Species	190,369
Organisation	317,867
Other	1,207,664



Instances per Class

About: Criticism

An Entity of Type: company, from Named Graph: http://dbpedia.org, within Data Space: dbpedia.org

Criticism is the construction of a judgement about the negative qualities of someone or something. Criticism can range from impromptu comments to a written detailed response. Criticism falls into several overlapping types including "theoretical, practical, impressionistic, affective, prescriptive, or descriptive". The term "brickbat" is sometimes used to mean "an unfavourable criticism, unkind remark or sharp put-down". The term originated in the 17th century, derived from the practice of throwing bricks as projectiles at a person who was disapproved of.

The Importance of Interoperability in Ontology: Case Study on DBpedia

September 7, 2023

Author: Carter-Beau Benson

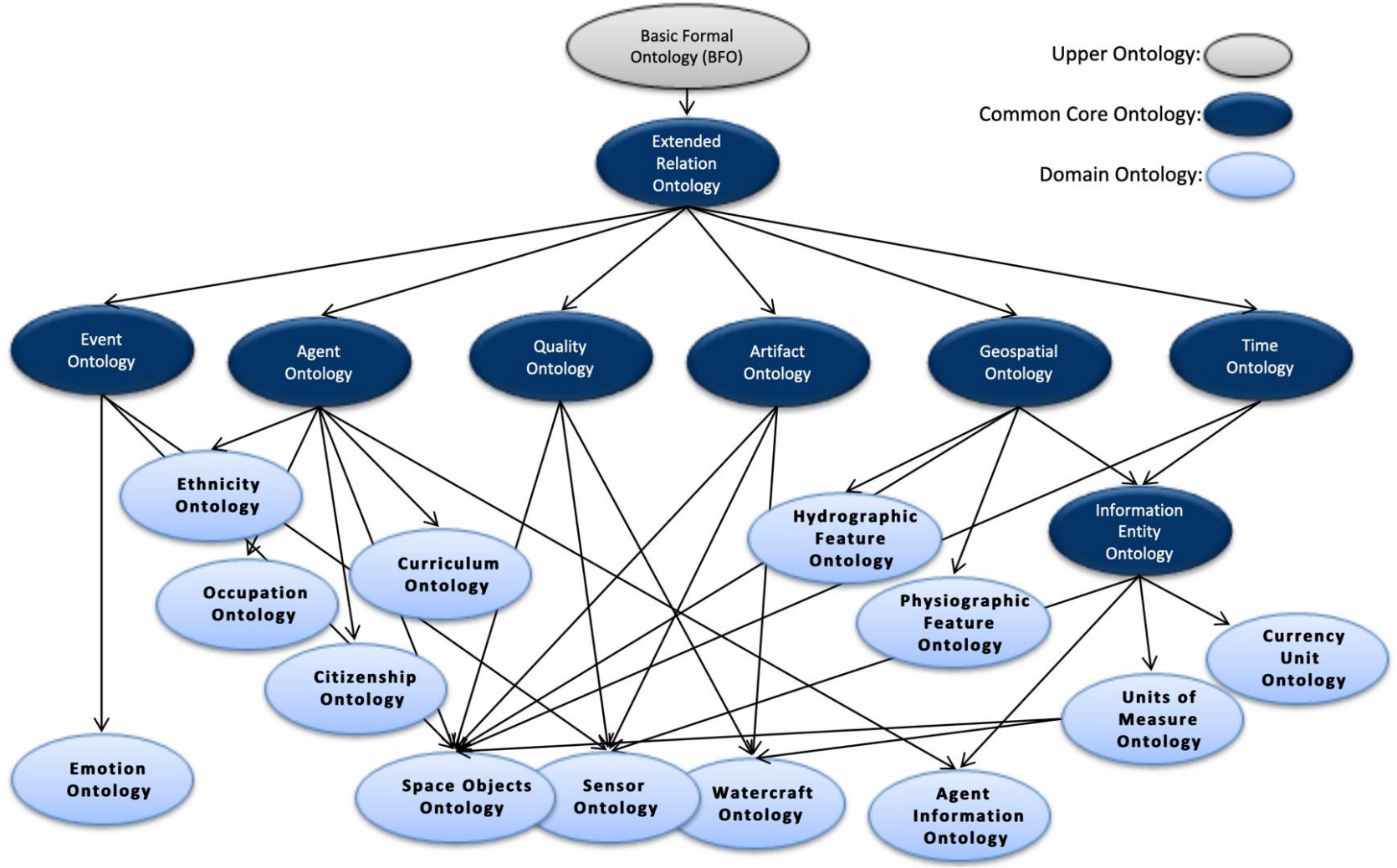
Semantic interoperability streamlines our ability to process and analyze vast data sets and creates a unified and efficient approach to understanding the information. This article sheds light on the significance of interoperability and reveals the potential pitfalls of a crowd-sourced resource like ***DBpedia***, with a particular focus on how sports players are connected to teams across different sports.

dbo:wikiPageWikiLink	• dbc:Criticism • dbr:Criticism_of_science • dbr:Critique_of_Pure_Reason • dbr:Analysis • dbr:Theatre_criticism • dbr:Gianni_Vattimo • dbr:Criticism_of_religion • dbr:Critique • dbc:Philosophical_methodology • dbr:Ad_hominem • dbr:Critical_Theory • dbr:Literary_criticism • dbr:H._L._Mencken • dbc:Literary_concepts • dbr:Art_criticism • dbr:Film_criticism • dbr:Immanuel_Kant • dbr:Self-criticism • dbr:Social_criticism

Outline

- Cybersecurity Landscape
- Proliferation of Knowledge Representation
- Addressing the Alignment Problem
- Design Pattern Practice

The growth of interest in D3FEND **is paralled by** the growth of interest the Common Core Ontologies



The growth of interest in D3FEND **is paralled by** the growth of interest the Common Core Ontologies

MEMORANDUM FOR CHIEF DIGITAL AND ARTIFICIAL INTELLIGENCE OFFICER COUNCIL MEMBERS
INTELLIGENCE COMMUNITY CHIEF DATA OFFICER COUNCIL MEMBERS

SUBJECT: Baseline Standards for Formal Ontology within the Department of Defense and the Intelligence Community

In April 2023, the Chief Digital and Artificial Intelligence Officer Council and the Intelligence Community Chief Data Officer Council chartered the joint Department of Defense (DoD) and Intelligence Community (IC) Ontology Working Group (DIOWG). It was tasked with developing coordinated ontologies to set the agreed definitions and standard necessary to make data machine understandable. Based on the DIOWG's recommendations, both Councils direct the use of three baselines: Top-Level Ontology, Basic Formal Ontology, and Common Core Ontology. These will set the baseline standards for formal DoD and IC ontology.

By aligning the DoD and IC ontologies to a common set of top and mid-level standards, the combined enterprise will realize significant gains in data interoperability, federated search and discovery, decreased analytic timelines, and better cost efficiency. This common approach to data ontology is key to deriving value from shared data assets at speed and scale. The DIOWG has provided additional background information on these international ontological standards in Attachment A.

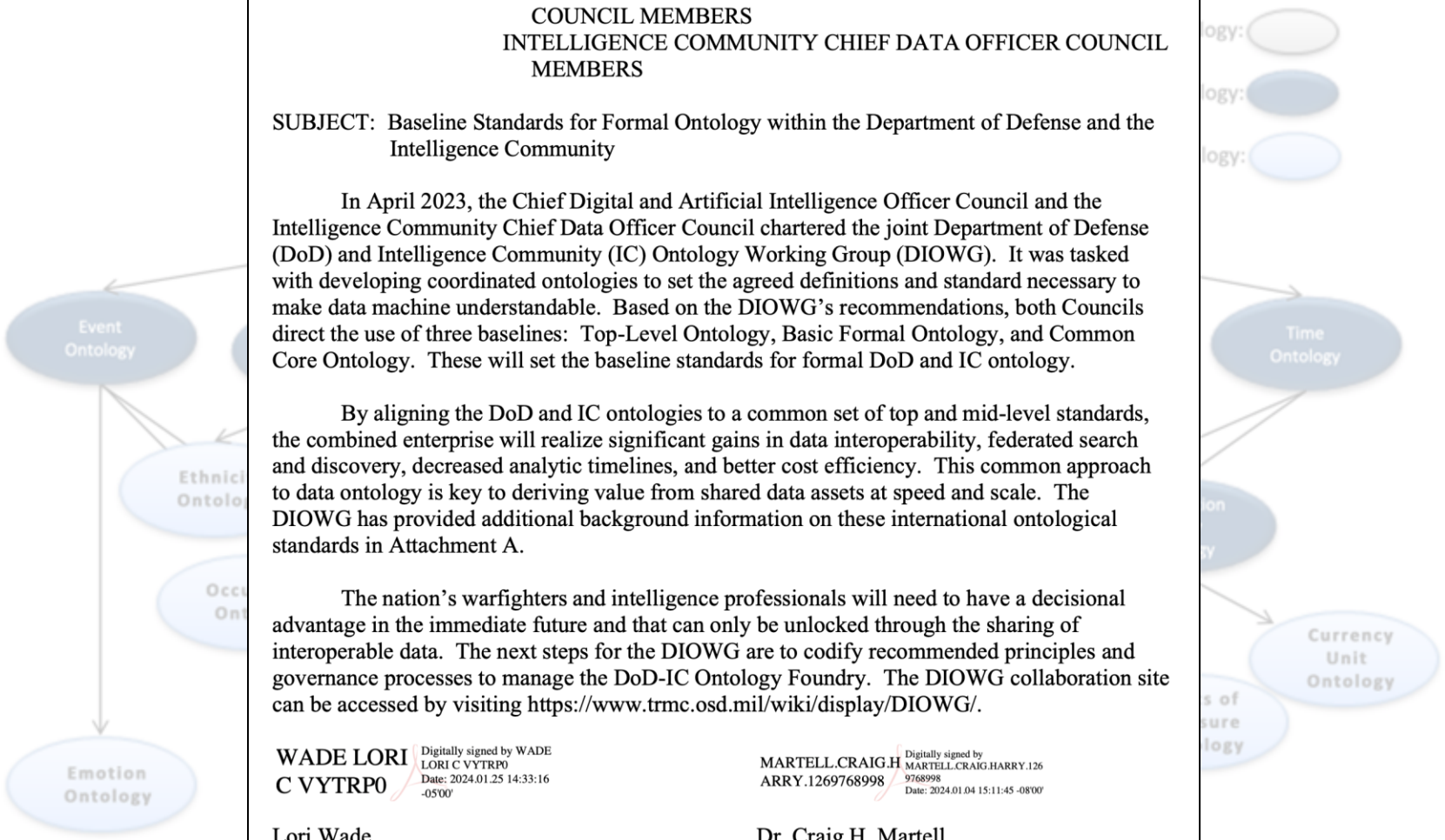
The nation's warfighters and intelligence professionals will need to have a decisional advantage in the immediate future and that can only be unlocked through the sharing of interoperable data. The next steps for the DIOWG are to codify recommended principles and governance processes to manage the DoD-IC Ontology Foundry. The DIOWG collaboration site can be accessed by visiting https://www.trmc.osd.mil/wiki/display/DIOWG/.

WADE LORI C VYTRP0 Digitally signed by WADE LORI C VYTRP0 Date: 2024.01.25 14:33:16 -05'00'

MARTELL.CRAIG.HARRY.1269768998 Digitally signed by MARTELL.CRAIG.HARRY.1269768998 Date: 2024.01.04 15:11:45 -08'00'

Lori Wade
Intelligence Community Chief Data Officer
Office of the Director of National Intelligence

Dr. Craig H. Martell
Chief Digital and Artificial Intelligence Officer
Department of Defense



The growth of interest in D3FEND **is paralled by** the growth of interest the Common Core Ontologies

The Common Core Ontologies

Mark JENSEN [a,b,1], Giacomo DE COLLE [b,c], Sean KINDYA [c], Cameron MORE [b,d], Alexander P. COX [b,d], and John BEVERLEY [b,c,e]

[a] *U.S. Customs and Border Protection*
[b] *National Center for Ontological Research*
[c] *University at Buffalo*
[d] *CUBRC, Inc.*
[e] *Institute for Artificial Intelligence and Data Science*

Abstract. The Common Core Ontologies (CCO) are designed as a mid-level ontology suite that extends the Basic Formal Ontology. In 2017, CUBRC, Inc. made CCO openly available. CCO has since been increasingly adopted by a broad group of users and applications and is proposed as the first standard mid-level ontology. Despite these successes, documentation of the contents and design patterns of the CCO has been comparatively minimal. This paper is a step toward providing enhanced documentation for the mid-level ontology suite through a discussion of the contents of the eleven ontologies that collectively comprise the Common Core Ontology suite.

Middle Architecture Criteria

John BEVERLEY [a,b,c,1], Giacomo DE COLLE [a,b], Mark JENSEN [b,d], Carter BENSON [a,b], and Barry SMITH [a,b,c]

[a] *Department of Philosophy, University at Buffalo*
[b] *National Center for Ontological Research, University at Buffalo*
[c] *Institute for Artificial Intelligence and Data Science, University at Buffalo*
[d] *U.S. Customs and Border Protection*

#	Project Leveraging CCO	Project Description
1	Additive Manufacturing and Maintenance Operations Ontology	Ontology for additive manufacturing and maintenance operations.
2	Aditive Manufacturing Ontology	AMOntology describes entities that capture knowledge about characteristics of computational models for AM processes.
3	AM-CDM-Ontology-Map	Mapping of Additive Manufacturing Common Data Model (AM-CDM) to CCO.
4	Bioregistry	Community-driven meta-registry of life science databases, ontologies, and other resources.
5	BWMD Domain Ontology	Represents domain-specific information for BWMD applications.
6	BWMD Mid-Level Ontology	A mid-level ontology for BWMD applications.
7	BWMD Original Non-Modularized Ontology	The original, non-modularized version of the BWMD ontology.
8	Cognitive Process Ontology	Consists of terms representing cognitive processes – kinds of mental processes – used by intelligence analysts.
9	Cyber Information Ontology	Represents cyber information, reliability, warrant, and transmission protocols.
10	Rare-Earth Elements (REE) Ontology	Formalizes the interactions among the members of various magmatic, hydrothermal, basinal, regolith, and supergene subsystems.
11	Regulatory Basis for Informed Consent (RUBRIC)	Ontology representing informed consent concepts.
12	Sciumo Tech	Technology projects by Sciumo.
13	Solid Data Interoperability Panel	Panel investigating data interoperability across applications while enabling secure collaboration and query using intuitive data boundaries.
14	Space Domain Ontologies	Ontologies representing phenomena relevant to the domain of space.
15	Standard Galactic	Standard Galactic project on data interoperability.
60	Sustainable Development and Climate (SDC) Ontology	Formalizes the sustainability impacts of implementations of actions, plans, strategies, and policies.
61	The KGC Knowledge Graph	Knowledge Graph of the Knowledge Graph Conference.
62	The MatWerk ontology	Represents comprehensive concepts in material work.
63	The PMD Core Ontology	Core ontology for PMD applications.
64	Thin-film solar cell ontology	Represents concepts and relationships in thin-film solar cell technology.
65	Typedb-schema-BFO	Upper ontology, with mid- and domain-level ontologies ported to Grakn schemae.
66	User Profile Ontology	Provides a standardized extensible semantics for representing information about a person's profile.
67	Virtual Materials Marketplace (VIMMP) Ontologies	Represents comprehensive concepts and relationships in the virtual materials marketplace.
68	Virus Infectious Disease Ontology	Represents epidemiology, classification, pathogenesis, and treatment of terms used by virologists, i.e. virus, prion, satellite, viroid, etc.
69	Wind Turbine Project	Represents wind turbine blade materials based on key performance parameters.

IEEE SA STANDARDS ASSOCIATION

IEEE

Active PAR

P3195

Standard for Requirements for a Mid-Level Ontology and Extensions

IEEE SA STANDARDS ASSOCIATION

IEEE

Active PAR

P3195.1

Standard for Common Core Ontology (CCO)

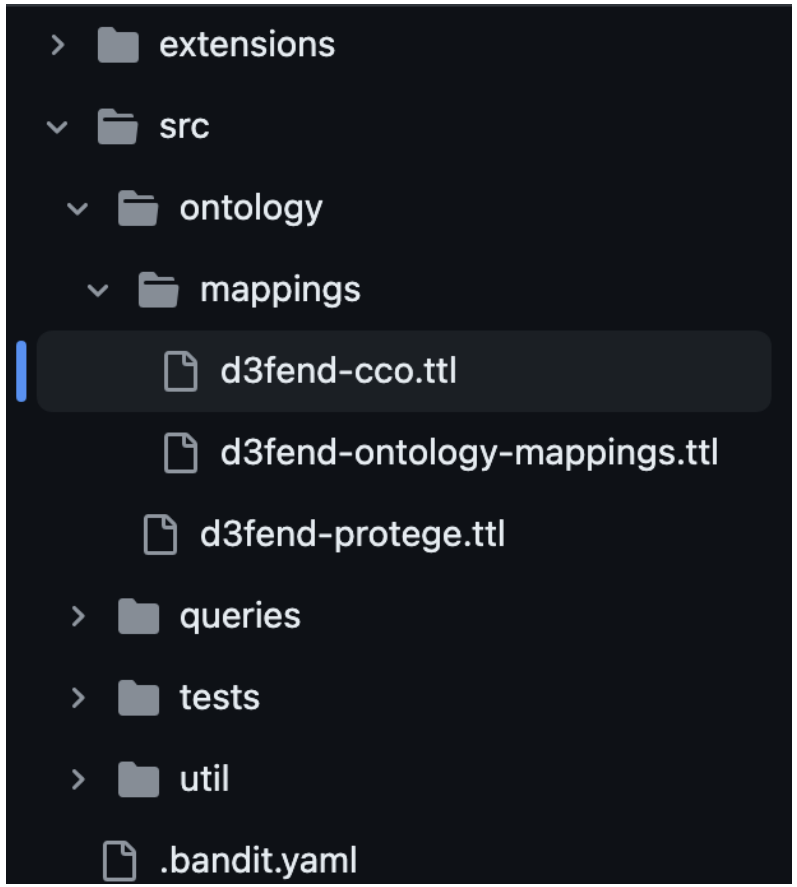
Harmonize D3FEND & CCO

- Strong motivation to **align the impressive work** found in the D3FEND knowledge graph with the **impressive work** found in CCO

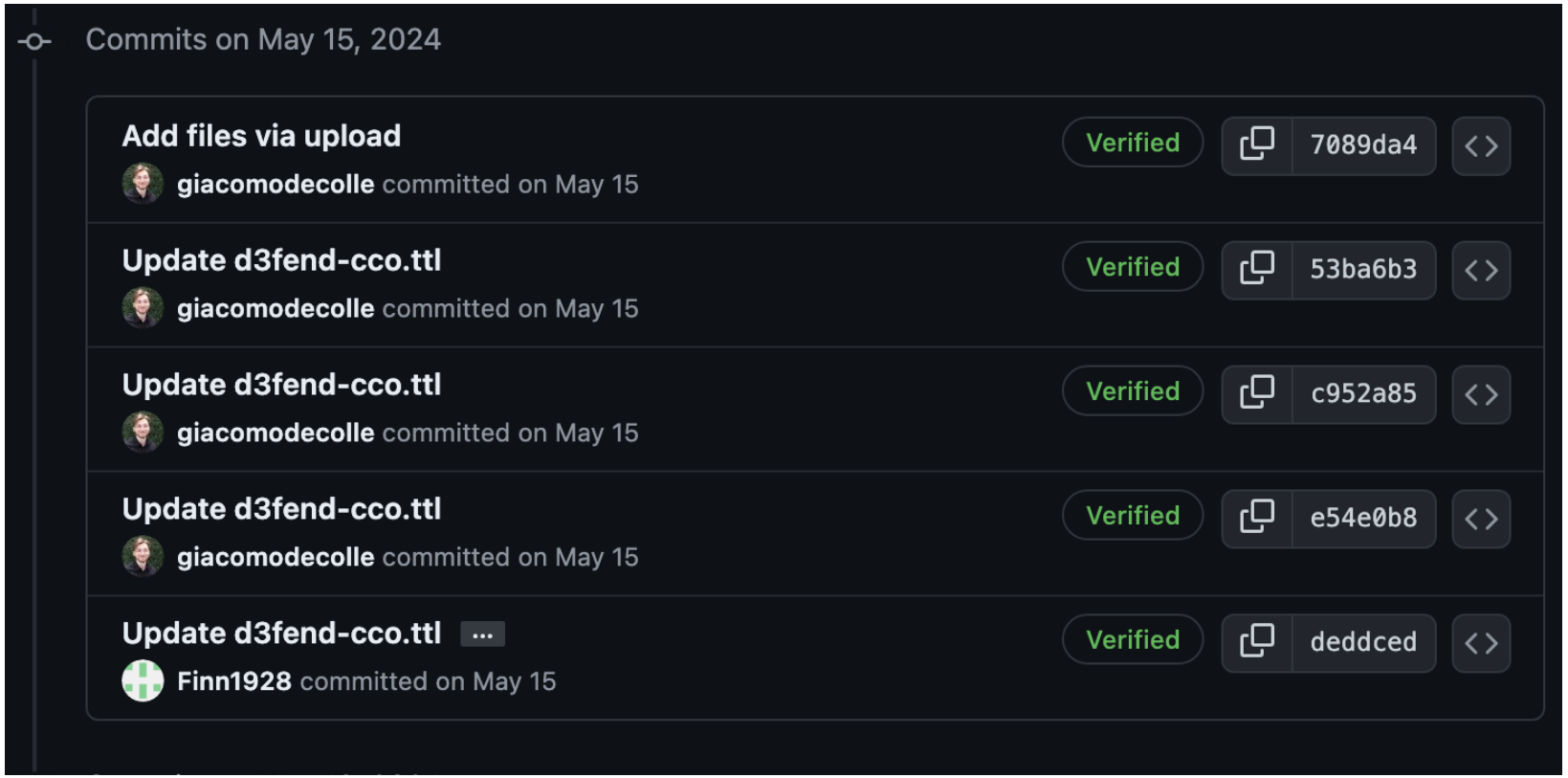
- All in the interest of **increasing semantic interoperability**

https://github.com/d3fend

Waste No Time



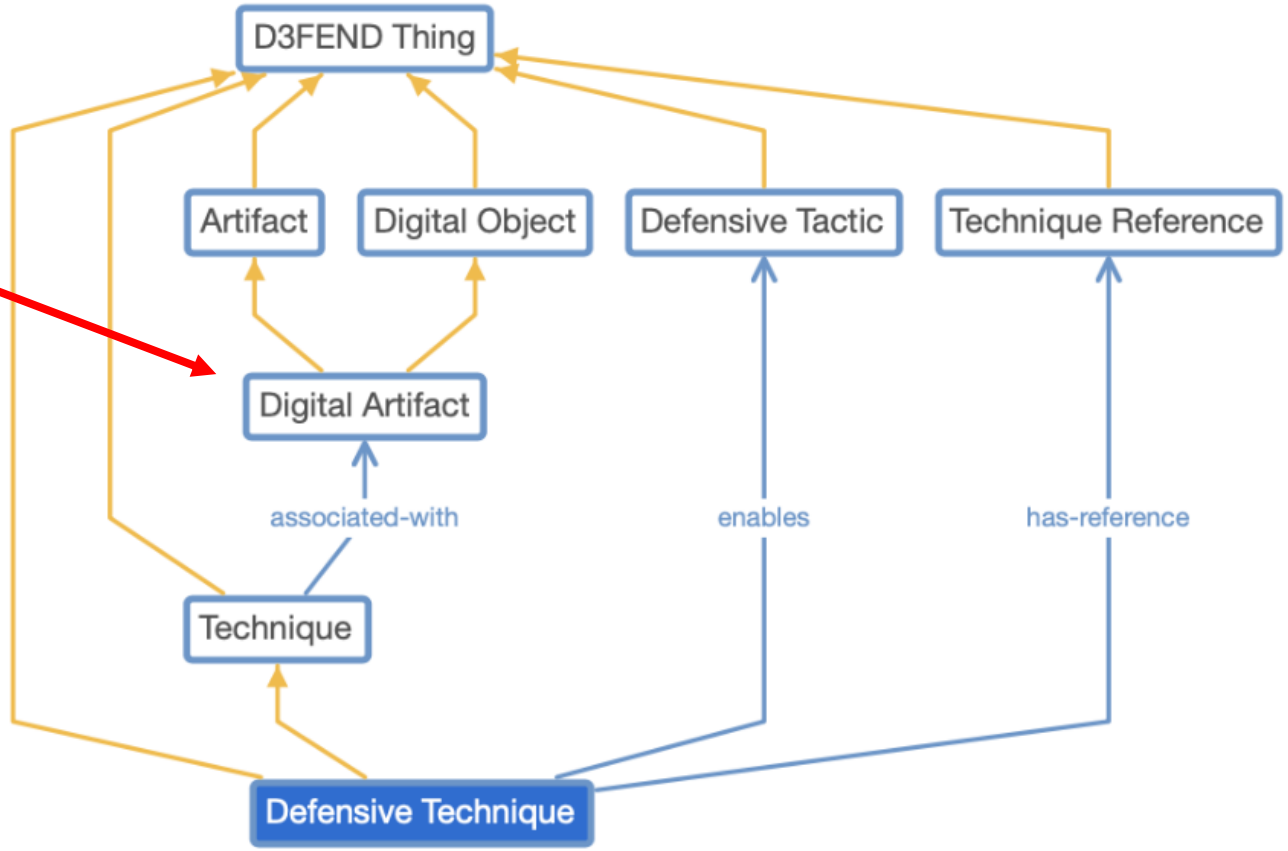
D3FEND team already started mapping



NCOR members already started assisting

- entity
  - continuant
    - **Digital Artifact**
      - **Digital Information**
      - **Digital Information Bearer**
    - generically dependent continuant
      - Information Content Entity
        - Descriptive Information Content Entity
        - Designative Information Content Entity
        - **Digital Information**
        - Directive Information Content Entity
        - Ideology
        - Religion
        - Representational Information Content Entity
    - independent continuant
    - Resource
    - specifically dependent continuant
  - occurrent
- External Threat Model Thing
  - ATTACK Thing
  - CAPEC Thing
  - SPARTA Thing

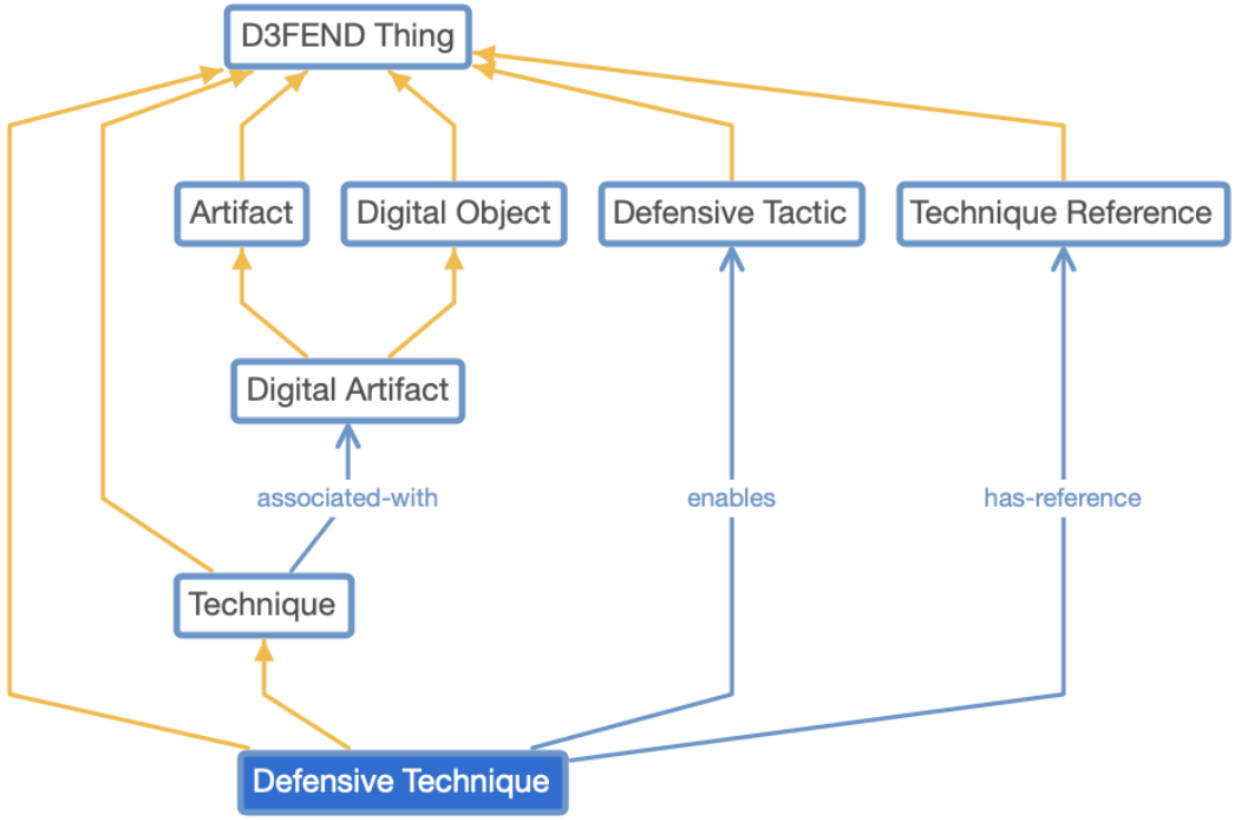
Preliminary D3CO Hierarchy



Core D3FEND Design Pattern

D3FEND

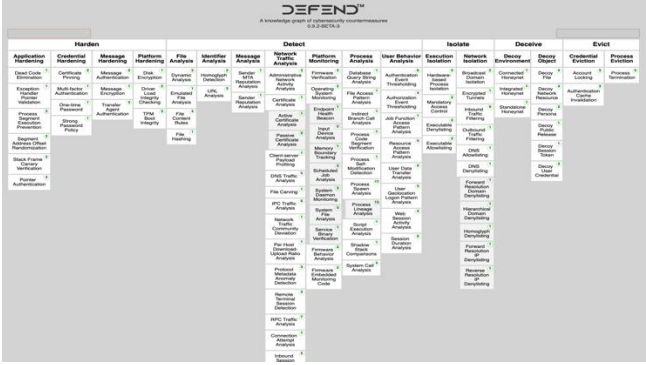
- D3FEND's longer-term goals are to (1) create a sustain- able knowledge framework for characterizing and relating cybersecurity countermeasure technology; and (2) accelerate knowledge discovery and acquisition efforts required to keep pace with technological changes in the cyber domain.



Outline

- Cybersecurity Landscape
- Proliferation of Knowledge Representation
- Addressing the Alignment Problem
- Design Pattern Practice

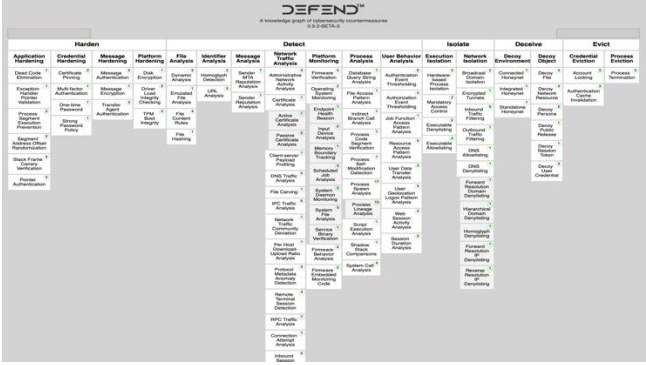
Competency Question



SELECT AND MODEL THE FOLLOWING

Information is encrypted on a solid-state drive.

Competency Question



SELECT AND MODEL THE FOLLOWING

Information is encrypted on a solid-state drive.

List:

- **Material entities involved in the encryption of a hard drive**
- **Qualities of those material entities**
- **Processes they participate in**
- **Information borne by the material entities**

Readings

- Cybersecurity Knowledge Graphs
- A Common Core-Based Cyber Ontology for Support of Cross-Domain Situational Awareness
- Cybonto: Towards Human Cognitive Digital Twins for Cybersecurity