# Ontology Development for Life Sciences

Presented by:
Anastasia Penkova & Yassir Makloul

# Table of Contents



# 01

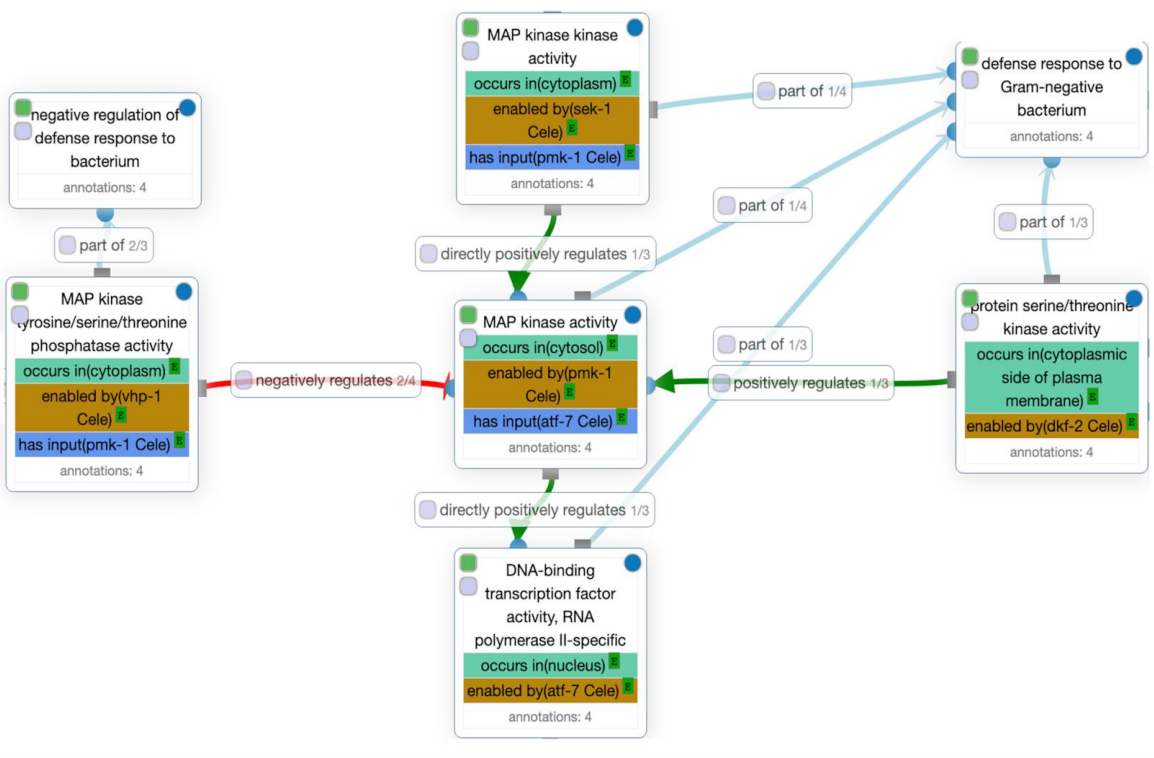
# Introduction

# Ontologies

***Def.:*** Structured representations of knowledge about a domain, defining concepts, relationships, and constraints within that domain.

***Languages used:***

- RDF
- OWL
- SPARQL (for queries)



# 02

# Motivation

# Why Do We Use Ontologies?

01

**Data Interoperability**

consistent and coherent data sharing in life sciences

02

**Integration**

unification of varied biological data for comprehensive synthesis

03

**Customization & Flexibility**

adaptable life science research applications

04

**Domain Knowledge Analysis**

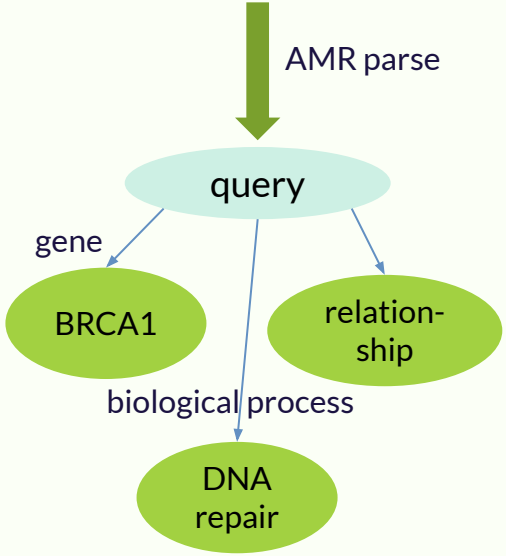
concise examination of life science information

# Why Do We Use Ontologies?

1. **Semantic Parse** of natural language query

User Question:

*"What processes is the BRCA1 gene involved in related to DNA repair?*




2. **Semantic Search** against a collection of gene ontology annotations

**GO:838201:** BRCA1 gene
The BRCA1 gene encodes for the BRCA1 protein or RNA molecule

is involved in

**GO:0006302:** Double-strand break repair
The BRCA1 protein plays a role in the process of DNA repair

is involved in

**GO:0006281:** DNA repair process
The BRCA1 protein plays a role in the process of DNA repair

3. Rank based on match score, return the most relevant gene **ontology annotations** and supporting documents

**System Response:**

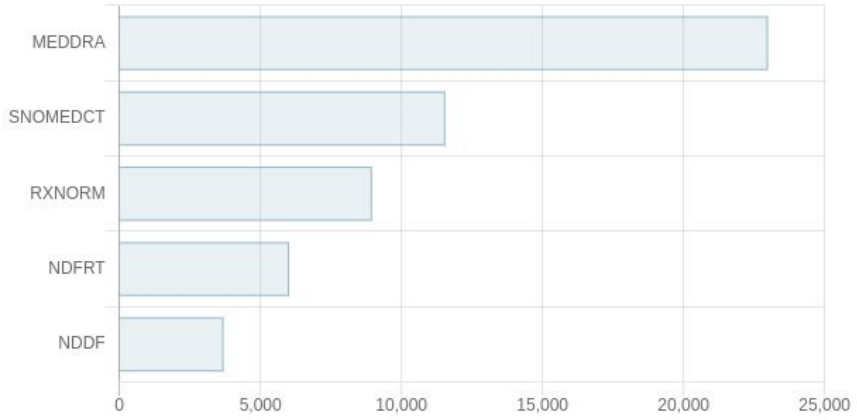
The system might respond with "BRCA1 is involved in the biological process of double-strand break repair (GO:0006302) and plays a role in homologous recombination (GO:0000724)

MATCHING PAPER

Click

# Ontologies in Bioportal

## Ontology visits (November 2023)




More

## Statistics

| | |
|---|---|
| Ontologies | 1,088 |
| Classes | 14,649,134 |
| Properties | 36,286 |
| Mappings | 81,457,436 |

# 03

# The General Development Pipeline

# I.

Scope

Reuse of Ontologies

databases for bioinformatic ontologies:

| Bioinformatic Databases | Short Description |
|---|---|
| OLS Ontology Search [53] | Ontology Lookup Service |
| OBO Library [79] | Open Biological and Biomedical Ontology |
| OMIM [57] | Public database of bibliographic information about human genes and genetic disorders. |
| BioPortal [80] | Repository of biomedical ontologies |
| AberOWL [81] | Ontology repository, semantic search engine |
| OntoBee [82] | A linked ontology data server to support ontology term dereferencing, linkage, query and integration |
| DiseaseCard [83] [84] [85] | Web-based tool for the collaborative integration of genetic and medical information |
| MalaCards [86] [87] | Integrated compendium for human diseases and their annotation |
| GeneCard [88] | Human Gene Database |
| DISEASES [89] | Text mining and data integration of dis-ease–gene associations |
| SIGNOR [90] | SIGNaling Network Open Resource Database of causal relationships between biological entities |
| KEGG [91] | Kyoto Encyclopedia of Genes and Genomes. Knowledge base for systematic analysis of gene functions, linking genomic information |
| MENTHA [92] | Resource for browsing integrated proteininteraction networks |
| PhosphositePlus [93] | Knowledge base dedicated to mammalian post-translational modifications (PTMs) |
| PhosphoELM [94] | Database of phosphorylation sites—update |
| UniProtKB [95] | Universal protein resource |
| HGMD [96] | Human Gene Mutation Database |

"An ontology you want to develop" and its general subfields and scope

Terminology

Class Hierarchy

Slots/ Properties

Facets/ Values

Instances

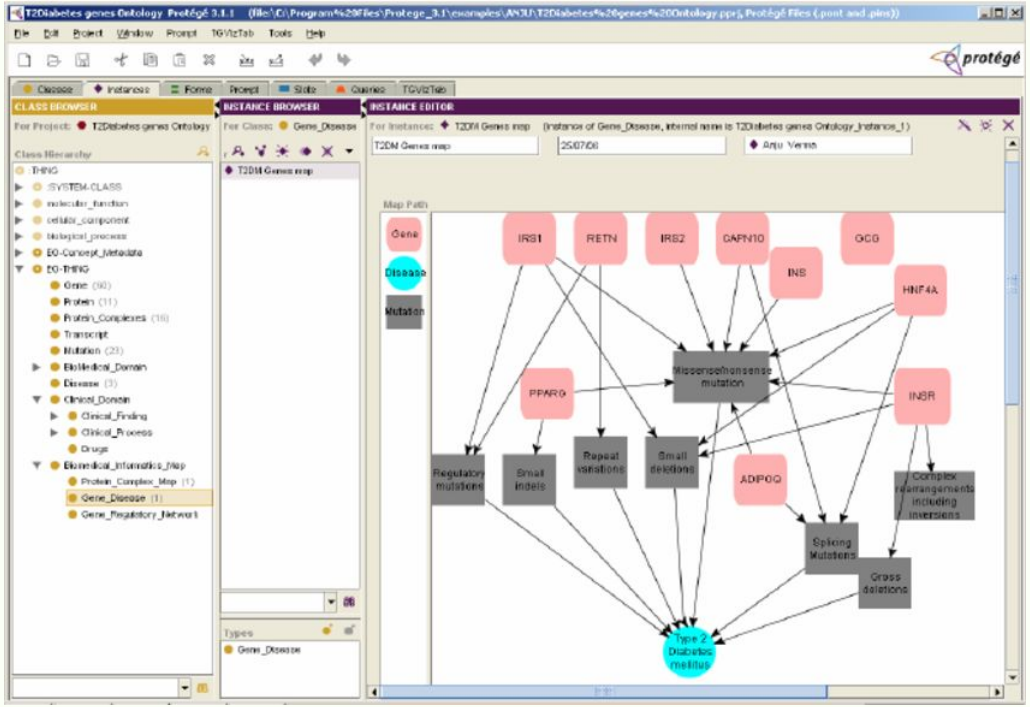
# II.

Terminology | Class Hierarchy | Slots/ Properties

| term | uri | synonyms | Relation to others | Onto Ref (reuse) |
|---|---|---|---|---|
| Cellular process | http://purl.obolibrary.org/obo/GO_0009987 | • Cell operation<br>• Cellular activity | Cellular metabolic process | GO |
| … | … | … | … | … |

Scope | Reuse of Ontologies | Facets/ Values | Instances

# III.




Facets/ Values/ constraints

Instances

- cardinality
- value restrictions
- default values

Specific instance to a class

Scope

Reuse of Ontologies

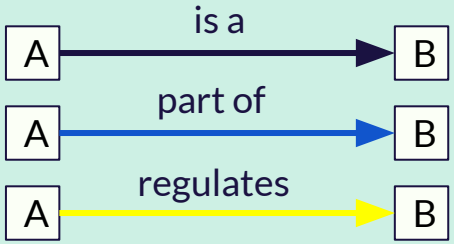
Terminology

Class Hierarchy

Slots/ Properties

Gene products:
- Molecular functions
- Biological processes
- Cellular components

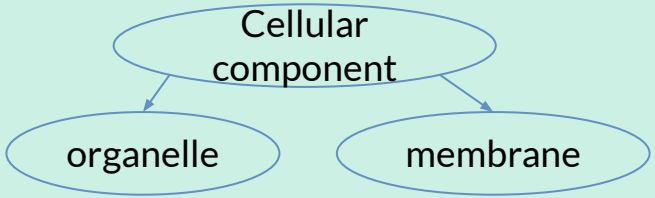
**GO: 0008150**
Biological process




Protein TP53

Scope | Reuse of Ontologies | Terminology | Class Hierarchy | Slots/ Properties | Facets/ Values | Instances

- Sequence Ontology
- Cell Ontology
- Protein Ontology
etc.




| GO_ID | No. of Genes | p-value |
|---|---|---|
| 00488 | 58 | 3.7E-12 |
| 00326 | 63 | 2.9E-10 |
| 00078 | 13 | 1.9E-4 |

# 04
# Perspectives

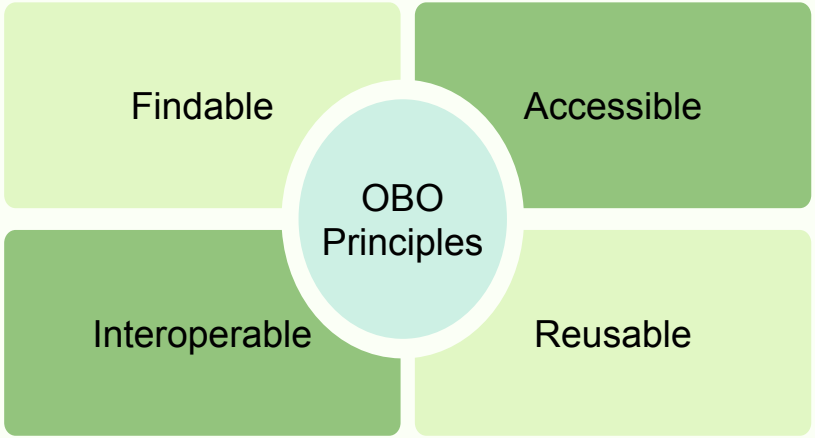
# 4.1.

## Ontology Development Kit:

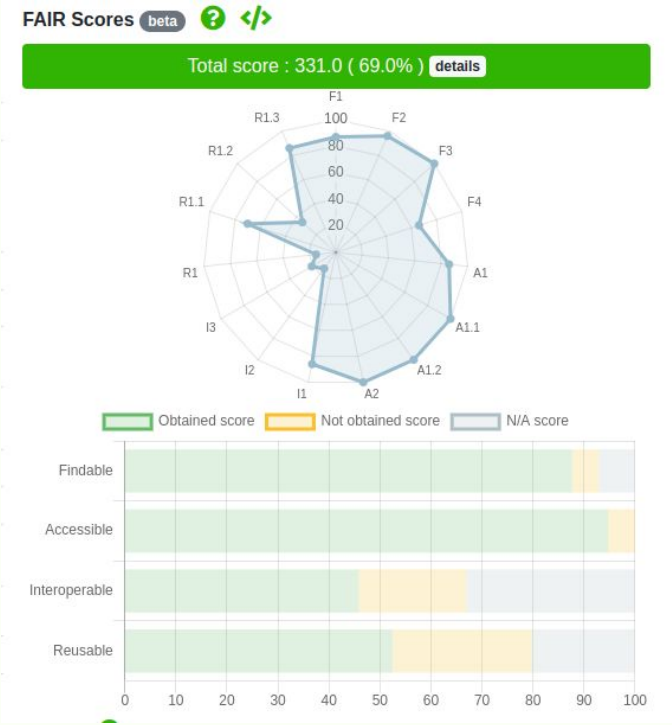
a toolkit for building, maintaining and standardizing biomedical ontologies.

# Open Biomedical Ontologies

## FAIR principles




## FAIR Score



## Ontology development challenges

- Variety of formats used for Ontology files: RDF/XML, OBO Flatfile, etc.
- Repositories inconsistencies when multiple ontologies use different repositories.

# Ontology Development Kit

## Purpose & Domain

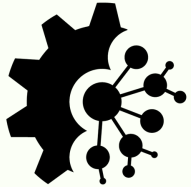
Tool to manage ontology life cycle, incl. over 70 biomedical ontologies

## Components

- Set of executable ontology workflows
- Toolbox to execute workflows

## Standardization

Customizable git repo with files and scripts for releases, tests, importing terms from other ontologies



odk

ONTOLOGY DEVELOPMENT KIT

## Toolbox Delivery

As a Docker image for convenience and centralization

## Automation

Automated generation of ontology release versions

## CI & Quality Control

Ensure the integrity and standard compliance of ontologies. Automated run of tests for validation

# 4.2.

# Towards a Core Ontology for Hierarchies of Hypotheses in Invasion Biology

# Motivation

## Challenges

- Complexity of hypotheses
- Inconsistency of results across studies

01

02

## Strategy

Merge method to assemble a core ontology from related ontologies

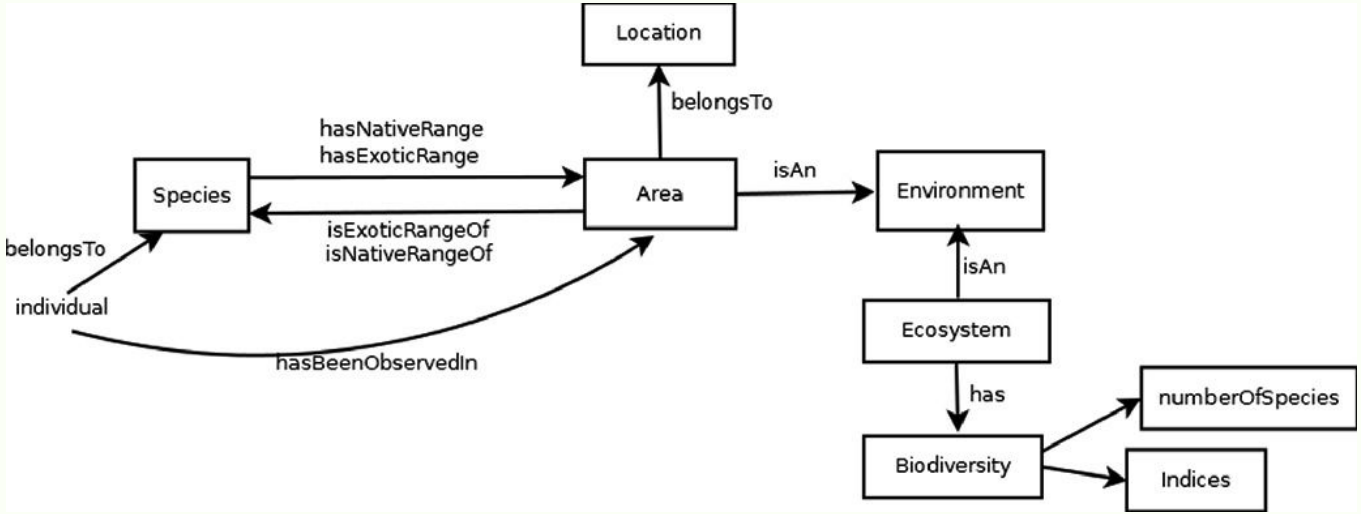
## Outcome

Major hypotheses in invasion biology

03

# The hierarchy of Hypotheses (HoH)

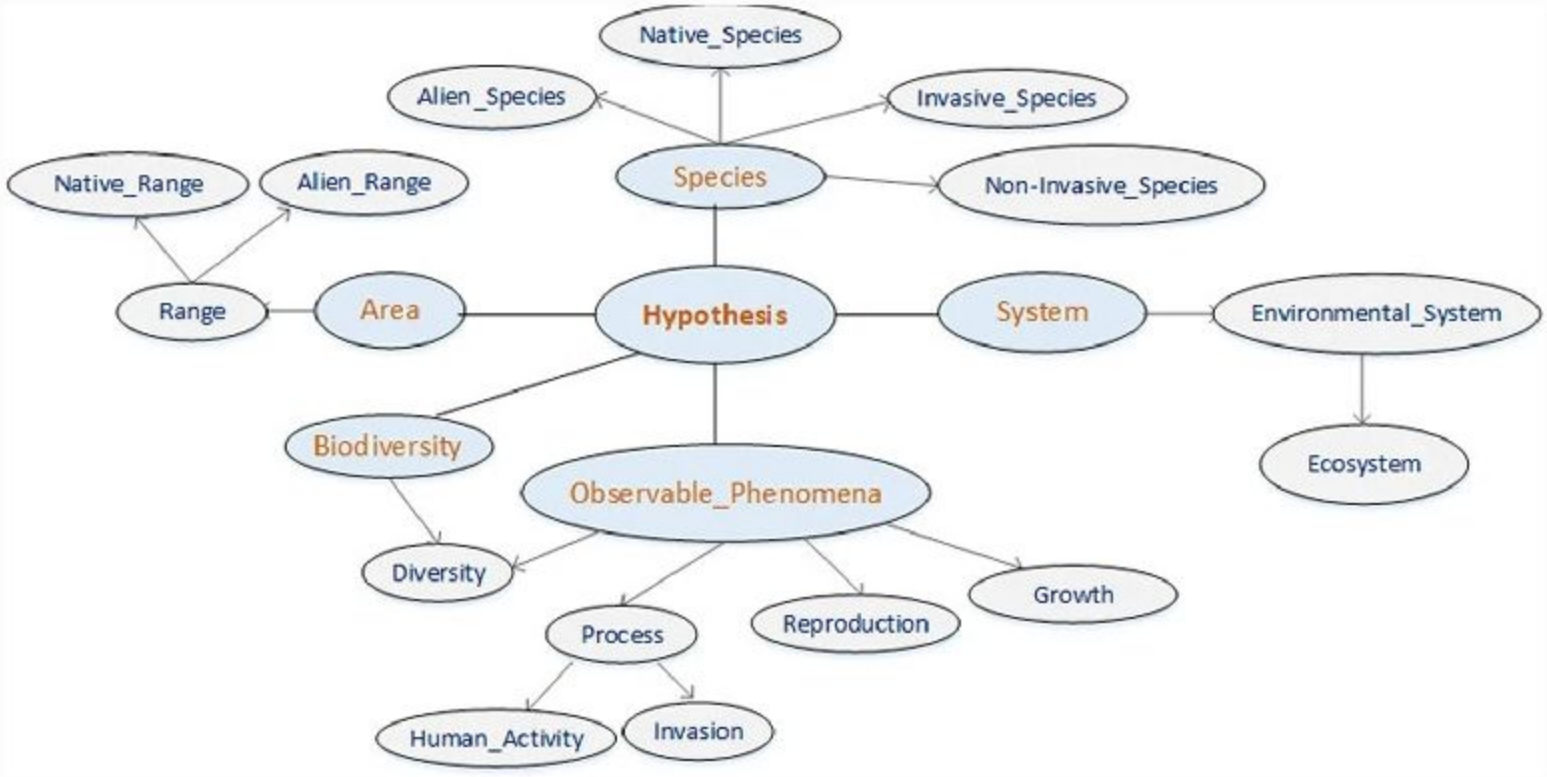
H01
H02
H03
H04
H05
H06
H07
H08
H09
H10
H11



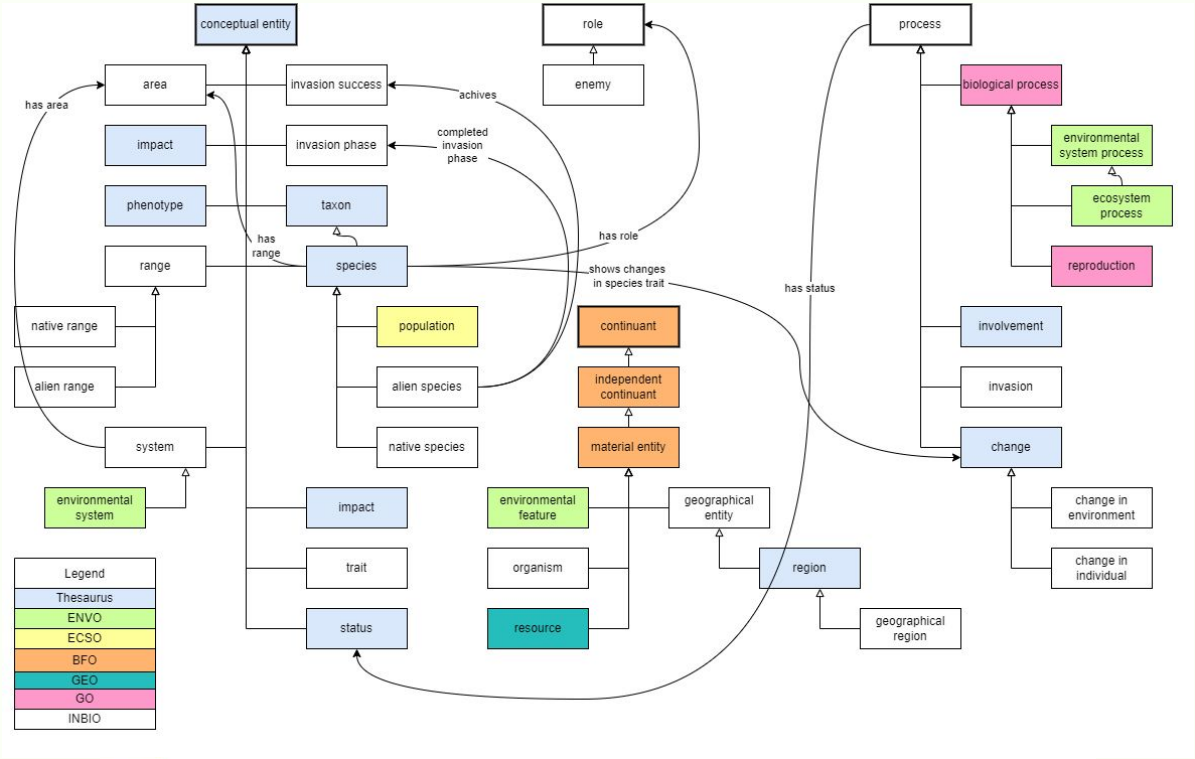

Challenges:

- Terms can exist in different ontologies with different definitions and IRIs
- Complexity of hypotheses

# 6 core concepts



# Invasion Biology Ontology (InBio)




process[C]

**IRI:** https://w3id.org/inbio#_000025

See NCIt:C29862

A process is an entity that exists in time by occurring or happening, has temporal parts and always involves and depends on some entity during the time it occurs.

**has sub-classes**

allocation [C], biological process [C], change [C], community dynamics [C], human activity [C], invasion [C], involvement [C]

Click

# Invasion Biology Ontology (InBio) metrics

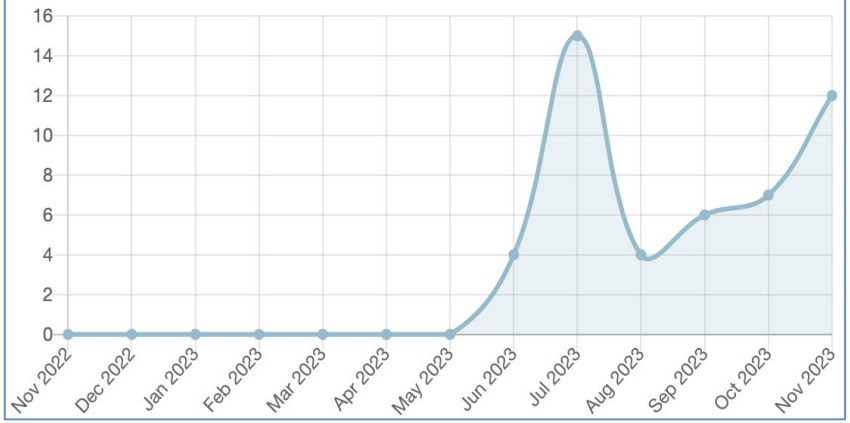
## Submissions

| Version | Released | Uploaded | Downloads |
|---|---|---|---|
| 1.1 (Parsed, Indexed, Metrics, Annotator) | 02/02/2023 | 02/02/2023 | OWL \| CSV \| RDF/XML \| Diff |
| 1.0.0 (Archived) | 07/15/2022 | 07/15/2022 | OWL |

## Metrics

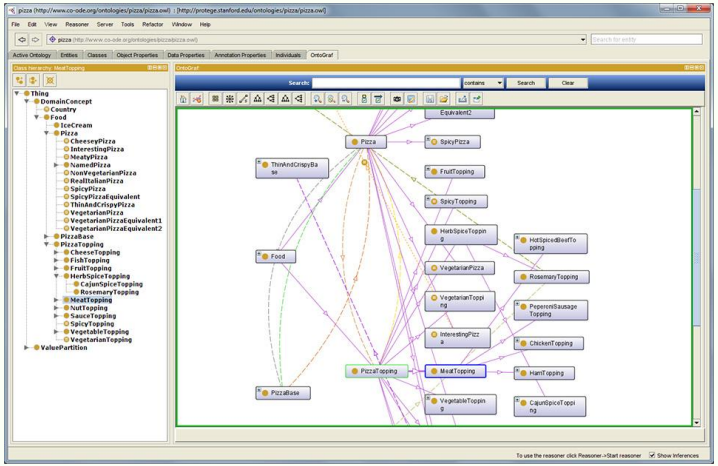
| | |
|---|---|
| Classes | 458 |
| Individuals | 18 |
| Properties | 98 |
| Maximum depth | 30 |
| Maximum number of children | 18 |
| Average number of children | 2 |
| Classes with a single child | 103 |
| Classes with more than 25 children | 0 |
| Classes with no definition | 120 |

## Visits



# 05
# Concept

## Reproduce the INBIO Ontology




Reproduce and visualize the INBIO Ontology using Protege

## ODK toolbox, for OntoDev workflows

odk

ONTOLOGY DEVELOPMENT KIT

Seed our repo, through the ODK toolbox, for our ontology repository

# 06
# Conclusion

# Conclusion

- Challenges persist in extracting implicit knowledge from diverse sources.
- Automation in ontology creation and evolution, alongside mapping current ontologies, presents future research directions.

# THANK YOU FOR YOUR ATTENTION

QUESTIONS?

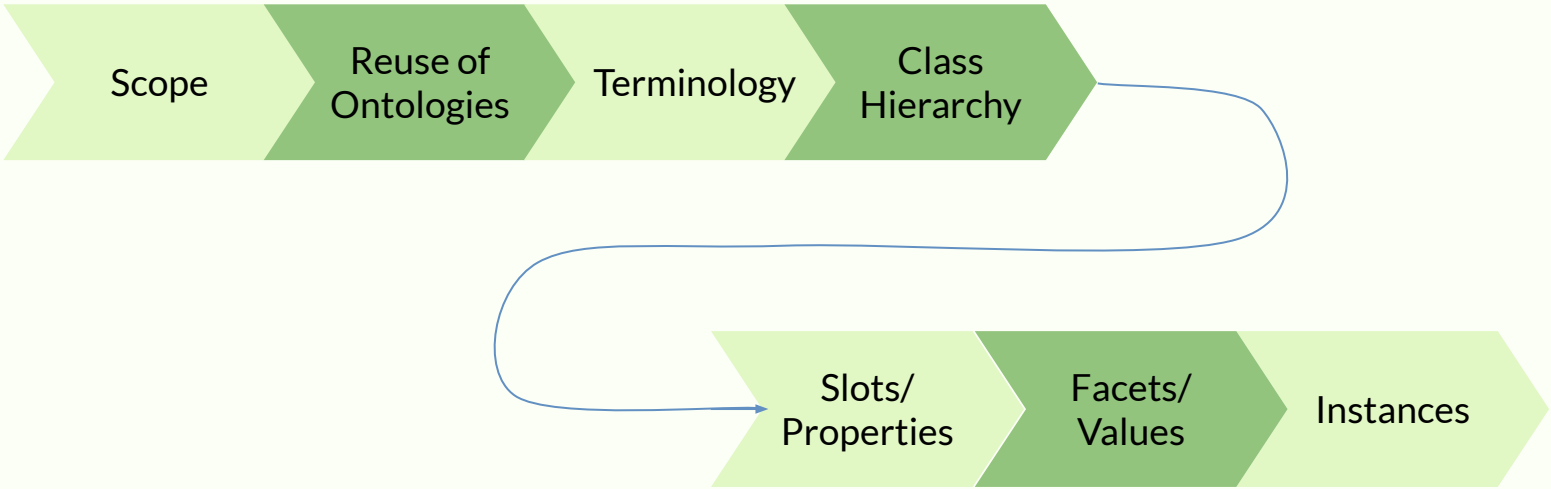
# References:


- **“Ontology Development Kit: a toolkit for building, maintaining and standardizing biomedical ontologies” (Matentzoglu et al., 2022)** https://academic.oup.com/database/article/doi/10.1093/database/baac087/6754192
- **“Towards a Core Ontology Hierarchies of Hypotheses in Invasion Biology” (Algergawy et al., 2020)** https://link.springer.com/chapter/10.1007/978-3-030-62327-2_1
- **Ontologies: An Overview (Selen Parlar, 2019)** https://medium.com/analytics-vidhya/ontologies-an-overview-b23ccc7e976
- **VOWL: Visual Notation for OWL Ontologies** http://vowl.visualdataweb.org/
- **Ontology Development 101: A Guide to Creating Your First Ontology (Natalya Noy, 2000)** https://protege.stanford.edu/publications/ontology_development/ontology101.pdf

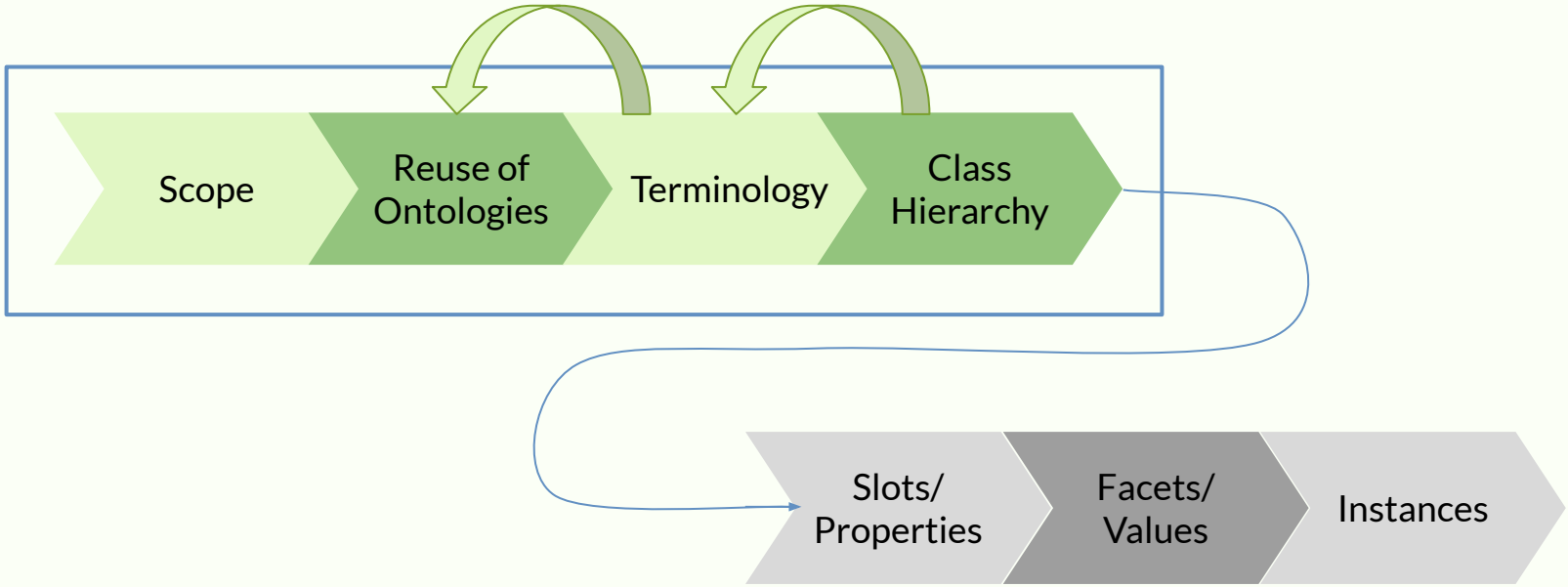
# Ontology Development for Life Sciences: Outcome

Presented by:
Anastasia Penkova & Yassir Makloul

# Ontology Development Pipeline



# Ontology Development Pipeline

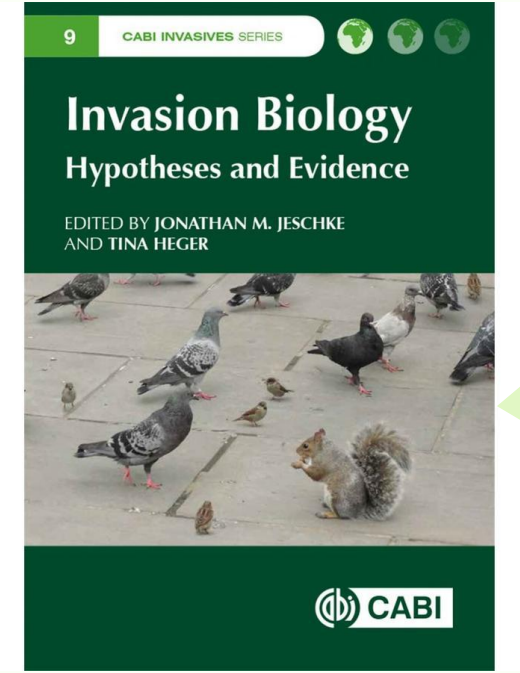


# Step 1: Define the scope/domain

## Invasion Biology

A research area dealing with the translocation, establishment, spread, impact and management of species outside of their native ranges, where they are called non-native or alien species

# Step 1: Define the domain




**H01 – Enemy release hypothesis**

**H02 – Biotic resistance hypothesis**

**H03 – Evolution of increased competitive ability**

**H04 – Shifting defense hypothesis**

**H05 – Phenotypic plasticity hypothesis**

**H06 – Darwin's naturalization hypothesis**

**H07 – Island susceptibility hypothesis**

**H08 – Limiting similarity hypothesis**

**H09 – Propagule pressure hypothesis**

**H10 – Disturbance hypothesis**

**H11 – Invasional meltdown hypothesis**

# Step 1: Define the domain

## Hierarchy-of-Hypotheses

**H01 – Enemy release hypothesis**

**H02 – Biotic resistance hypothesis**

**H03 – Evolution of increased competitive ability**

**H04 – Shifting defense hypothesis**

**H05 – Phenotypic plasticity hypothesis**

**H06 – Darwin's naturalization hypothesis**

**H07 – Island susceptibility hypothesis**

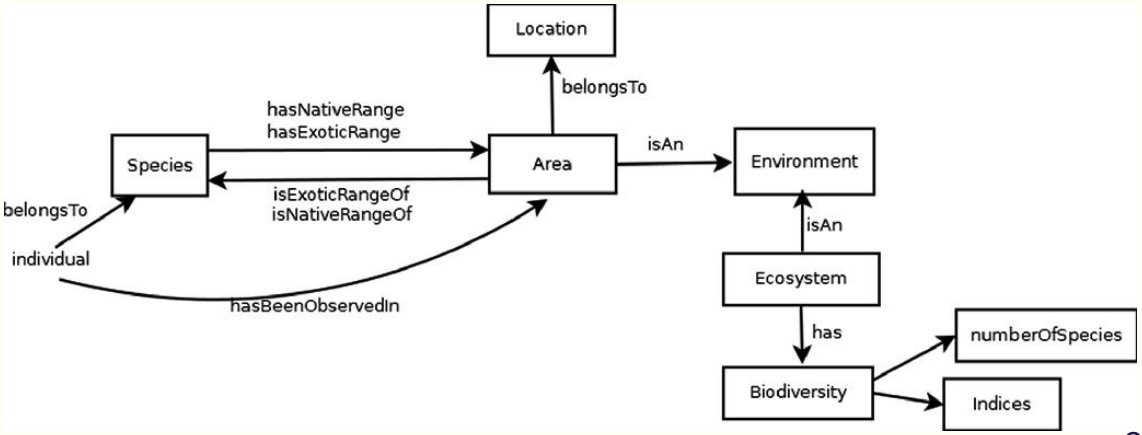
**H08 – Limiting similarity hypothesis**

**H09 – Propagule pressure hypothesis**

**H10 – Disturbance hypothesis**

**H11 – Invasional meltdown hypothesis**

*"An ecosystem with high biodiversity is more resistant against non-native species than an ecosystem with lower biodiversity"*



# Step 2: Define entities from hypotheses

| MODEL | CODE | CRM | DOCS | INTEGRATE | MARKETING | REASON | FINAL 🏆 | COST | SPEED |
|---|---|---|---|---|---|---|---|---|---|
| GPT-4 v1/0314 ☁️ | 80 | 88 | 98 | 52 | 88 | 50 | 76 | 7.19 € | 1.26 rps |
| GPT-4 v2/0613 ☁️ | 80 | 83 | 95 | 52 | 88 | 50 | 74 | 7.19 € | 2.07 rps |
| GPT-4 Turbo v3/1106-preview ☁️ | 60 | 75 | 98 | 52 | 88 | 62 | 72 | 2.52 € | 0.68 rps |
| GPT-3.5 v2/0613 ☁️ | 62 | 79 | 73 | 75 | 81 | 48 | 70 | 0.35 € | 1.39 rps |
| GPT-3.5 v3/1106 ☁️ | 62 | 68 | 71 | 63 | 78 | 59 | 67 | 0.24 € | 2.29 rps |
| GPT-3.5-instruct 0914 ☁️ | 44 | 90 | 69 | 60 | 88 | 32 | 64 | 0.36 € | 2.12 rps |
| GPT-3.5 v1/0301 ☁️ | 44 | 75 | 67 | 67 | 82 | 24 | 60 | 0.36 € | 3.93 rps |
| Mistral 7B OpenChat-3.5 f16 ✅ | 46 | 72 | 72 | 49 | 88 | 31 | 60 | 0.51 € | 2.14 rps |
| Starling 7B-alpha f16 ⚠️ | 51 | 66 | 67 | 45 | 88 | 36 | 59 | 0.61 € | 1.80 rps |
| Mistral 7B Instruct f16 ✅ | 32 | 68 | 68 | 44 | 74 | 36 | 54 | 0.58 € | 1.89 rps |
| Mistral 7B OpenOrca f16 ✅ | 42 | 57 | 76 | 21 | 78 | 26 | 50 | 0.43 € | 2.55 rps |
| Llama2 13B Vicuna-1.5 f16 🦙 | 36 | 37 | 44 | 39 | 82 | 38 | 46 | 1.02 € | 1.07 rps |
| Llama2 13B Hermes f16 🦙 | 38 | 23 | 30 | 61 | 60 | 43 | 42 | 1.03 € | 1.06 rps |
| Llama2 13B Hermes b8 🦙 | 32 | 24 | 29 | 61 | 60 | 43 | 42 | 4.94 € | 0.22 rps |
| Llama2 13B Puffin f16 🦙 | 37 | 12 | 33 | 33 | 56 | 41 | 36 | 4.89 € | 0.22 rps |
| Llama2 13B Puffin b8 🦙 | 37 | 9 | 34 | 31 | 56 | 39 | 34 | 8.65 € | 0.13 rps |
| Llama2 13B chat f16 🦙 | 15 | 38 | 17 | 30 | 75 | 8 | 30 | 0.76 € | 1.43 rps |
| Llama2 13B chat b8 🦙 | 15 | 38 | 15 | 30 | 75 | 6 | 30 | 3.35 € | 0.33 rps |
| Orca 2 13B f16 ⚠️ | 15 | 22 | 32 | 22 | 67 | 19 | 29 | 0.99 € | 1.11 rps |
| Llama2 7B chat f16 🦙 | 20 | 33 | 14 | 27 | 50 | 20 | 27 | 0.59 € | 1.86 rps |
| Mistral 7B Zephyr-β f16 ✅ | 23 | 34 | 27 | 44 | 29 | 4 | 27 | 0.51 € | 2.14 rps |
| Mistral 7B Notus-v1 f16 ⚠️ | 16 | 43 | 6 | 41 | 48 | 4 | 26 | 0.80 € | 1.37 rps |
| Mistral 7B f16 ✅ | 0 | 4 | 20 | 42 | 52 | 12 | 22 | 0.93 € | 1.17 rps |
| Orca 2 7B f16 ⚠️ | 13 | 0 | 22 | 18 | 52 | 4 | 18 | 0.81 € | 1.34 rps |
| Llama2 7B f16 🦙 | 0 | 2 | 5 | 2 | 28 | 2 | 7 | 1.01 € | 1.08 rps |

# Step 2: Define entities from hypotheses

```
openai-env > start.py
```

```python
from openai import OpenAI
client = OpenAI()

completion = client.chat.completions.create(
  model="gpt-3.5-turbo",
  messages=[
    {"role": "system", "content": "You will be provided with 11 hypotheses. "+
                                  "These hypotheses are necessary for development of ontology 'Invasive Biology'." + "
                                  "Your task is to distinguish entities and find the relationship 'is a'"},
    {"role": "user", "content": "H01 – Enemy release hypothesis: The absence of enemies in the exotic range is a cause of invasion success." +
                                "H02 – Biotic resistance hypothesis:An ecosystem with high biodiversity is more resistant against non-native species than an ecosyste
                                "H03 – Evolution of increased competitive ability: After having been released from natural enemies, non-native species will allocate
                                " reproduction (this re-allocation is due to genetic changes), which makes them more competitive." +
                                "H04 – Shifting defense hypothesis: After having been released from natural specialist enemies, non-native species will allocate more
                                " defenses against generalist enemies and less energy in expensive defenses against specialist enemies (this re-allocation is due to
                                "this way will be invested in growth and/or reproduction, which makes the non-native species more competitive."+
                                "H05 – Phenotypic plasticity hypothesis: Invasive species are more phenotypically plastic than non-invasive or native ones."+
                                "H06 – Darwin's naturalization hypothesis: The invasion success of non-native species is higher in areas that are poor in closely "+
                                "related species than in areas that are rich in closely related species."+
                                "H07 – Island susceptibility hypothesis: Non-native species are more likely to become established and have major ecological impacts o
                                "H08 – Limiting similarity hypothesis: The invasion success of non-native species is high if they strongly differ from native species
                                "H09 – Propagule pressure hypothesis: A high propagule pressure (a composite measure consisting of the number of individuals introduc
                                "H10 – Disturbance hypothesis: The invasion success of non-native species is higher in highly disturbed than in relatively undisturbe
                                "H11 – Invasional meltdown hypothesis:The presence of non-native species in an ecosystem facilitates invasion by additional species,
  ]
)

print(completion.choices[0].message)
```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL 1: bash

```
(openai-env) Anastasias-MacBook-Air:openai-env bronnika$ python3 start.py
ChatCompletionMessage(content="Entities:\n1. Enemy\n2. Exotic range\n3. Invasion success\n4. Ecosystem\n5. Biodiversity\n6. Non-native species\n7. Natural enemies\n8. Energy\n9. Growth\n10. Reproduction\n11. Genetic changes\n12. Defense\n13. Specialist enemies\n14. Generalist enemies\n15. Phenotypic plasticity\n16. Native species\n17. Darwin's naturalization hypothesis\n18. Closely related species\n19. Areas\n20. Islands\n21. Continents\n22. Ecological impacts\n23. Limiting similarity\n24. Propagule pressure\n25. Introduction events\n26. Disturbance\n27. Undisturbed ecosystems\n28. Invasional meltdown\n29. Additional species\n30. Survival\n31. Facilitates invasion\n\nRelationship 'is a':\n1. Exotic range is a cause of invasion success.\n2. Ecosystem with high biodiversity is more resistant against non-native species.\n3. Non-native species will allocate more energy in growth and/or reproduction.\n4. Non-native species will allocate more energy in cheap defenses against generalist enemies and less energy in expensive defenses against specialist enemies.\n5. Invasive species are more phenotypically plastic than non-invasive or native ones.\n6. Invasion success of non-native species is higher in areas that are poor in closely related species.\n7. Non-native species are more likely to become established and have major ecological impacts on islands than on continents.\n8. Invasion success of non-native species is high if they strongly differ from native species, and it is low if they are similar to native species.\n9. High propagule pressure is a cause of invasion success.\n10. Invasion success of non-native species is higher in highly disturbed ecosystems.\n11. The presence of non-native species in an ecosystem facilitates invasion by additional species.", role='assistant', function_call=None, tool_calls=None)
(openai-env) Anastasias-MacBook-Air:openai-env bronnika$
```

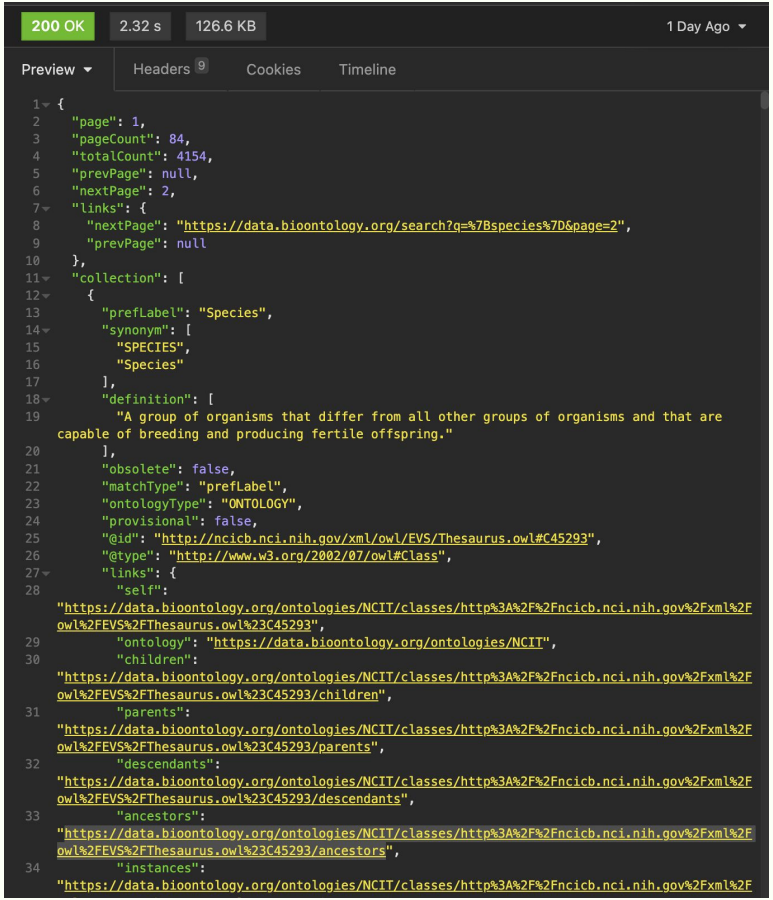
# Result of LLM: Entities

- Enemy release
- Exotic range
- Invasion success
- Biotic resistance
- Ecosystem
- Biodiversity
- Non-native species
- Genetic changes
- Energy allocation
- Growth
- Reproduction
- Defense mechanisms
- Generalist enemies
- Phenotypic plasticity
- Native species
- Darwin's naturalization
- Richness of closely related species
- Islands
- Continental areas
- Ecological impacts
- Limiting similarity
- Propagule pressure
- Introduction events
- Disturbance
- Undisturbed ecosystems
- Invasional meltdown

# Step 3: Reuse ontologies

```python
# Check if the request was successful (status code 200)
if response.status_code == 200:
    data = response.json()
    if "collection" in data and len(data["collection"]) >= 3:
        outcomes[value] =[]
        for outcome in data["collection"][:3]:
            prefLabel = outcome.get("prefLabel", "")
            definition = outcome.get("definition", "")
            id = outcome.get("@id", "")
            ontology_url = outcome["links"].get("ontology", "")
            synonyms = outcome.get("synonym", "")
            children = outcome["links"].get("children", "")
            parents = outcome["links"].get("parents", "")
            descendants = outcome["links"].get("descendants", "")
            ancestors = outcome["links"].get("ancestors", "")

            outcomes[value].append({
                "prefLabel": prefLabel,
                "definition": definition,
                "id": id,
                "ontology_url": ontology_url,
                "synonyms": synonyms,
                "children": children,
                "parents": parents,
                "descendants": descendants,
                "ancestors": ancestors
            })
```



# Step 4: Structure the results in csv files

| term | prefLabel | synonym | definition | Ontology |
|---|---|---|---|---|
| Enemy release | Release | ['CMNP', 'レレアーセ', 'Release'] | [] | IOBC |
| Enemy release | enemy | ['natural enemy'] | ['An organism that is a predator, consumer, parasite, parasitoid or pathogen of another organism.'] | INBIO |
| Enemy release | Enemy Swim Lake | [] | ['A lake in the glacial lake region of northeastern South Dakota located in Day County.'] | GAZ |
| Enemy release | Release | ['Release', 'Released', 'Releasing'] | ['To make available; set free.'] | NCIT |
| Enemy release | Conventional release and prolonged-release | ['Conventional release and extended-release', 'Conventio | ['A release characteristic where a dose form displays a rate and time of release of the active substance(s) in the dose form based on their intrinsic propert | SNOMEDCT |
| Enemy release | Release | [] | [] | AURA |
| Enemy release | Release | [] | [] | SYN |
| Enemy release | Release | [] | ['Once the viral envelope has separated from the cell membrane Influenza virus particles are actively released to complete the budding process. HA (hem | HINO |
| Enemy release | release | [] | ['<p>Change the status of an object representing an Act so it is no longer "held", i.e., allow action to occur. For an HL7 Act, the state transitions per the HL | HL7 |
| Enemy release | release | [] | [] | SWEET |
| Enemy release | percent release | ['% release'] | [] | RCD |
| Enemy release | Serotonin release | [] | [] | LOINC |
| Enemy release | Institutional Release | [] | ['Discharge or release of an individual from any\ntype of correctional or therapeutic residential\nfacility.'] | APAONTO |
| Enemy release | HIPAA release | [] | [] | OPMI |
| Enemy release | signal release | ['signal secretion'] | ['The process in which a signal is secreted or discharged into the extracellular medium from a cellular source.'] | PLANP |
| Enemy release | signal release | ['signal secretion'] | ['The process in which a signal is secreted or discharged into the extracellular medium from a cellular source.'] | UPHENO |
| Enemy release | signal release | [] | [] | CHIRO |
| Enemy release | neurotransmitter release | [] | [] | OCHV |
| Exotic range | range | [] | [] | SWEET |
| Exotic range | Animals, Exotic | ['Exotic Pet', 'Animal, Exotic', 'Exotic Pets', 'Exotic Animal', | ['Animals native to a foreign country or of foreign origin or character, that are not native to the United States.'] | MESH |
| Exotic range | chloroxylenol / hydrocortisone / pramoxine Otic Solution [Exotic-HC] | [] | [] | RXNORM |
| Exotic range | Exotic Other | ['exotic other'] | ["A conceptualization of those with identities other than one's own that they are inherently fascinating in their assumed primitiveness or undevelopedne | GSSO |
| Exotic range | range | ['レンジ', 'ドメイン', '範囲', 'range', '域'] | [] | IOBC |
| Exotic range | exotic molecular entity | ['exotic molecular entities'] | ['A molecular entity in which one or more sub-atomic particles have been replaced by other particles of the same charge.'] | BERO |
| Exotic range | range | [] | [] | ENVTHES |
| Exotic range | Exotic Pets | ['pet, exotic', 'animals, exotic', 'exotic animal', 'animal, exc | [] | MDM |
| Exotic range | Animals, Exotic | [] | [] | OMIT |
| Exotic range | range | ['geographical range', 'geographic range'] | ['An area where a species is found in the wild.'] | INBIO |
| Exotic range | exotic molecular entity | [] | [] | BIOMODELS |
| Exotic range | Exotic shorthaired cat | [] | [] | SNMI |
| Exotic range | exotic molecular entity | [] | ['A molecular entity in which one or more sub-atomic particles have been replaced by other particles of the same charge.'] | VDOT |
| Exotic range | exotic molecular entity | ['exotic molecular entities'] | ['A molecular entity in which one or more sub-atomic particles have been replaced by other particles of the same charge.'] | CHEBI |
| Exotic range | Avian paramyxovirus 1.exotic RNA | ['Avian paramyxovirus 1.exotic ribonucleic acid'] | [] | LOINC |
| Exotic range | Range | ['Range'] | ['The difference between the lowest and highest numerical values; the limits or scale of variation.'] | NCIT |
| Exotic range | Range | [] | ['The difference between the lowest and highest numerical values; the limits or scale of variation.'] | suicideo |
| Exotic range | range | [] | ['the range is a measure of variation which describes the difference between the lowest score and the highest score in a set of numbers (a data set)'] | PSDO |

# Step 4: Structure the results in csv files

| | A | B | C | D | E | F | G | H | I |
|---|---|---|---|---|---|---|---|---|---|
| 1 | value | comments | classes | parents | children | relationship with | prefLabel | definition | id |
| 2 | Enemy release | | | | | | Enemy Release Hypothesis | | http://purl.jp/bio/4/id/20130607940385504 |
| 3 | Enemy release | | enemy | | specialist, generalist | | natural enemy | | http://purl.jp/bio/4/id/20090602252307711 |
| 4 | Exotic range | | | | | | exotic | | http://sweetontology.net/matrPlant/Exotic |
| 5 | Exotic range | | | | | | Animals, Exotic | ['Animals native to a foreign cc | http://purl.bioontology.org/ontology/MESH/D000068881 |
| 6 | Exotic range | | | | | | Exotic-HC | | http://purl.bioontology.org/ontology/RXNORM/352639 |
| 7 | Range | | range | | exotic, native | | range | ['The difference between the lc | http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C38013 |
| 8 | Invasion Success | | invasion success | | | | | | |
| 9 | Invasion success | | | | | | success | ['favorable outcome of an activ | http://purl.bioontology.org/ontology/CSP/2483-5678 |
| 10 | Invasion success | | | | | | Success | ['An event that accomplishes it | http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C25362 |
| 11 | Ecosystem | | ecosystem | | undistributed, distributed | | Ecosystem | ['A functional system which inc | http://purl.bioontology.org/ontology/MESH/D017753 |
| 12 | Ecosystem | | | | | | ecosystem | ['complex of organisms formin | http://purl.bioontology.org/ontology/CSP/4000-0170 |
| 13 | Ecosystem | | | | | | Ecosystem | | http://www.semanticweb.org/mca/ontologies/2018/8/untitled-onto |
| 14 | Biodiversity | | | | | | Biodiversity | | http://purl.jp/bio/4/id/20090607788434703 |
| 15 | Biodiversity | | biodiversity | | | | Biodiversity | ['The variety of all native living | http://purl.bioontology.org/ontology/MESH/D044822 |
| 16 | Biodiversity | | | | | | Biodiversity | ['The degree of variation of life | http://edamontology.org/topic_3050 |
| 17 | Species | | species | | native, non-native, closely related, lack of non-native | | species richness | ['A group of organisms that diff | http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C45293 |
| 18 | Non-native species | | | | | | native species | | http://data.loterre.fr/ark:/67375/BLH-NZ9X4S2Z-W |
| 19 | Non-native species | | | | | | native tree species | | http://vocabs.lter-europe.net/EnvThes/20738 |
| 20 | Genetic changes | | Genetic change | change | | | | | |
| 21 | Genetic changes | | | | | | Genetic | ['Having to do with information | http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#C17938 |
| 22 | Genetic changes | | | | | | genetic | ['disease that is a direct result | http://radlex.org/RID/RID5659 |
| 23 | Genetic changes | | | | | | genetic | | http://sbmi.uth.tmc.edu/ontology/ochv#28472 |
| 24 | Growth | | | | | | Growth | | http://purl.bioontology.org/ontology/LNC/LA31297-7 |
| 25 | Growth | | Growth | | | | Growth | ['Gradual increase in the numt | http://purl.bioontology.org/ontology/MESH/D006128 |
| 26 | Growth | | | | | | growth | ['The increase in size or mass | http://purl.obolibrary.org/obo/GO_0040007 |
| 27 | Reproduction | | | | | | reproduction (biology) | | http://purl.jp/bio/4/id/20090606773660616 |
| 28 | Reproduction | | Reproduction | | | affects (class) pop | Reproduction | ['The total process by which or | http://purl.bioontology.org/ontology/MESH/D012098 |
| 29 | Reproduction | | | | | | reproduction | ['total process by which organi | http://purl.bioontology.org/ontology/CSP/2586-6130 |
| 30 | Defense mechanisms | | | | | | Defense Mechanisms | ['Unconscious process used by | http://purl.bioontology.org/ontology/MESH/D003674 |
| 31 | Defense mechanisms | | | | | | Defense Mechanisms | ['Any intrapsychic strategies th | http://ontology.apa.org/apaonto/termsonlyOUT%20(5).owl#Defen |
| 32 | Defense mechanisms | | | | | | defense mechanisms | | http://sbmi.uth.tmc.edu/ontology/ochv#C0011142 |
| 33 | Specialist enemies | | | | | | Specialist | | http://purl.bioontology.org/ontology/LNC/LA28116-4 |
| 34 | Specialist enemies | | | | | | specialist | | http://sbmi.uth.tmc.edu/ontology/ochv#C0087009 |
| 35 | Specialist enemies | | | | | | specialist | | http://purl.jp/bio/4/id/20090607150802661 |

# Step 5: Terms analysis

| | A | B | C | D | E | F | G | H | I | J | K | L | M |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | value | comments | classes | parents | children | relationship with | prefLabel | definition | id | ontology_url | synonyms | children (imported when INBIO) | parents |
| 2 | Enemy release | split to release and enemy | enemy | | specialist, generalist | | natural enemy | | | https://data.bioor | ['天敵', 'natural enemy'] | generalist, parasite, predator, specialist | role |
| 3 | Enemy release | | release | | | | | ['To make available; set free.'] | | | | data release; in our case enermy release | action |
| 4 | Range | | range | | exotic, native | | range | ['The difference between the lowe | | https://data.bioor | Range | (types of ranges) | conceptua |
| 5 | Invasion success | | invasion | | | | | ['The process of transportation, establishment and | | | ['biological invasion', 'species invasion'] | | process |
| 6 | Ecosystem | | ecosystem | | undistributed, distributed | | Ecosystem | ['A functional system which includ | | https://data.bioor | ['Ecological System', 'Ecosystems', 'Systems, Ecological', 'System, Ec | | environme |
| 7 | Biodiversity | | biodiversity | | | | Biodiversity | ['The variety of all native living org | | https://data.bioor | ['Biological Diversity', 'Diversity, Biological'] | | ecosystem |
| 8 | Species | | species | | native, non-native, closely related, lack of non-native | | species richness | ['A group of organisms that differ f | | https://data.bioontology.org/ontologies/NCIT | | | Taxon/Tax |
| 9 | Genetic changes | | genetic change | change | | | | ['Having to do with information that is passed from parents to offspring through genes in sperm and egg cells.'] | | | | | Molecula |
| 10 | Growth | | growth | | | | Growth | ['Gradual increase in the number, | | https://data.bioontology.org/ontologies/MESH | | | action / ph |
| 11 | Reproduction | | reproduction | | | affects (class) population size (not in hypotheses) | Reproduction | ['The total process by which orgar | | https://data.bioor | ['Period, Reproductive', 'Periods, Reproductive', 'Reproductive Index, | | life cycle / |
| 12 | Phenotypic plasticity | | phenotypic plasticity | | | | Phenotypic plasticity | | | https://data.bioor | ['Phenotypic plasticity', '表現型可塑性'] | | life history |
| 13 | Richness of closely related species | split into species richness and closely related species | species richness | biodiversity | | | species richness | ["p[Henderson's] ecological divers | | https://data.bioor | ['alpha diversity', 'richnes | - / | diversity in |
| 14 | Richness of closely related species | | closely related species - see above - | | | | | | | | | | |
| 15 | Islands | | islands | | | | Islands | ['Tracts of land completely surrour | | https://data.bioontology.org/ontologies/MESH | | | geographic |
| 16 | Continental areas | no api output for the context | area | | | islands (sibling) | | | | | | | areas |
| 17 | Ecological impacts | was added to the list | ecological impact | impact | | | Environmental impacts | | | | | | Ecology |
| 18 | Limiting similarity | it is a property: is similar to | | | | | Similarity | ['The quality of being alike; having | | https://data.bioor | ['Like', 'Similar', 'Similarity'] | | |
| 19 | Propagule pressure | | propagule pressure | | | | | ['A composite measure consisting of the number o | | | ['introduction effort'] | | conceptua |
| 20 | Introduction events | nothing similar found in the context of inbio | event | | introduction | | | | | | | | prospectiv |
| 21 | Disturbance | | disturbance | | | | Disturbance | ['A departure or divergence from t | | https://data.bioor | ['Disturbance'] | | conceptua |
| 22 | Biotic resistance | resistance | | | | | | ['[NCIT] A Resistance Process consists of activities of biologic molecules or complexes involved in processes that maintain | | | | | diversity |
| 23 | Energy allocation | split to energy and allocat | energy | | | | | ['The capacity of a physical system to do work.'] | | | | | resource |
| 24 | Energy allocation | hypo name | allocation | | | | | ['To set apart or re-distribute for a special purpose'] | | | | | process |
| 25 | Defense mechanisms | | defense | | | | | ['A trait or behaviour that is useful for repelling an enemy.', 'Protection from harm.'] | | | | | compound |
| 26 | Darwin's naturalization | no outputs; hypo name => area | | | | | | Darwin's naturalization hypothesis predicts that invaders less related to native flora are more likely to be successful than those that | | | | | |
| 27 | Invasional meltdown | no outputs; the name of 11th hypo => new class "survival" | | | | | | The invasional meltdown hypothesis (IMH) posits that positive interactions among invaders initiate positive population-level feedba | | | | | |

# Step 6: Define classes and their hierarchies

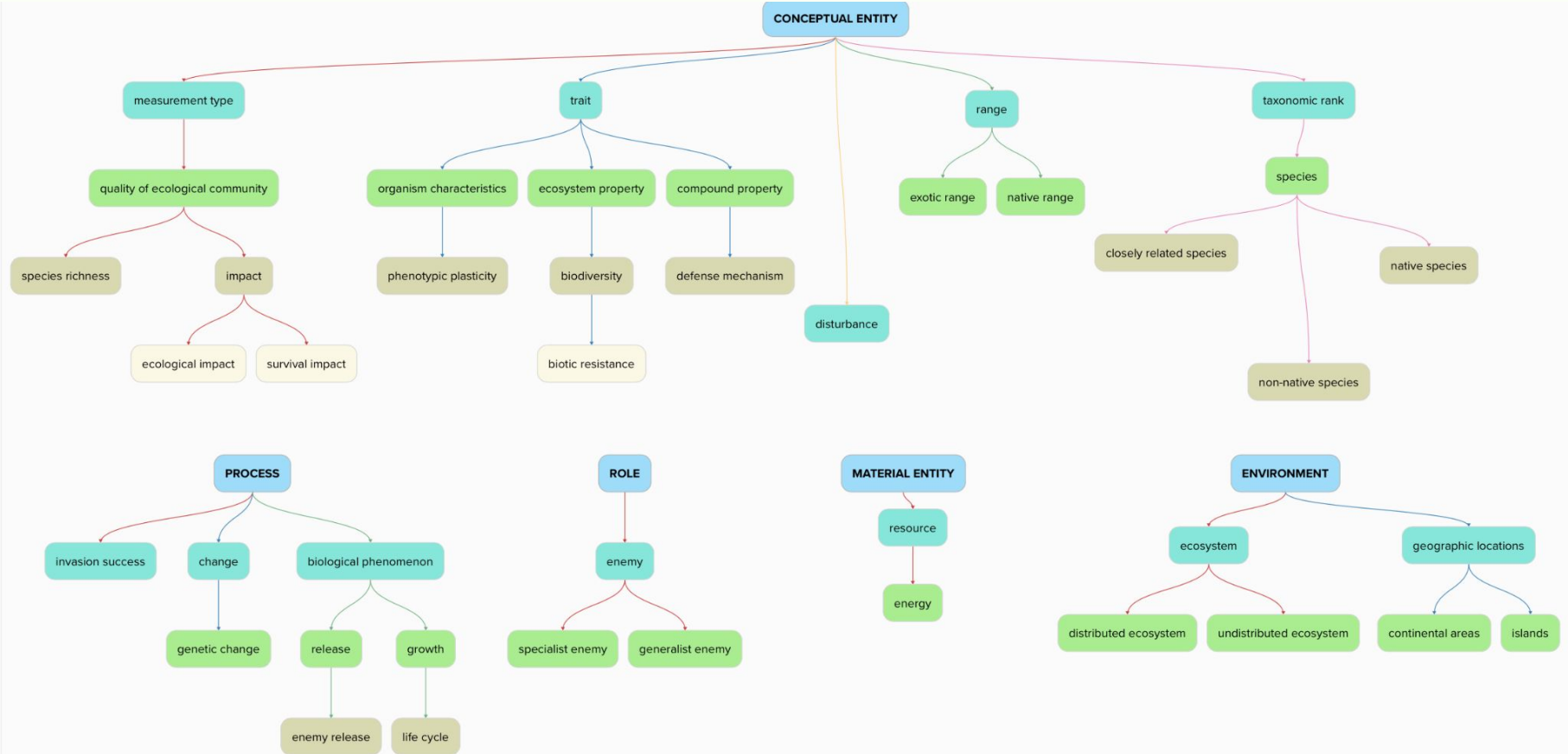
**NLP output for “Is_a” relationships:**

- Invasion success is caused by the absence of enemies in the exotic range.
- Non-native species allocate more energy in cheap defenses against generalist enemies after being released from natural specialist enemies.
- Invasive species are more phenotypically plastic than non-invasive or native ones.
- The invasion success of non-native species is higher in areas that are poor in closely related species.
- An ecosystem with high biodiversity is more resistant against non-native species.
- Non-native species allocate more energy in growth and/or reproduction after being released from natural enemies.
- Non-native species are more likely to become established and have major ecological impacts on islands.
- The invasion success of non-native species is high if they strongly differ from native species, and it is low if they are similar to native species.
- A high propagule pressure is a cause of invasion success.
- The invasion success of non-native species is higher in highly disturbed ecosystems.
- The presence of non-native species in an ecosystem facilitates invasion by additional species, increasing their likelihood of survival or ecological impact.

# Step 6: Define classes and their hierarchies

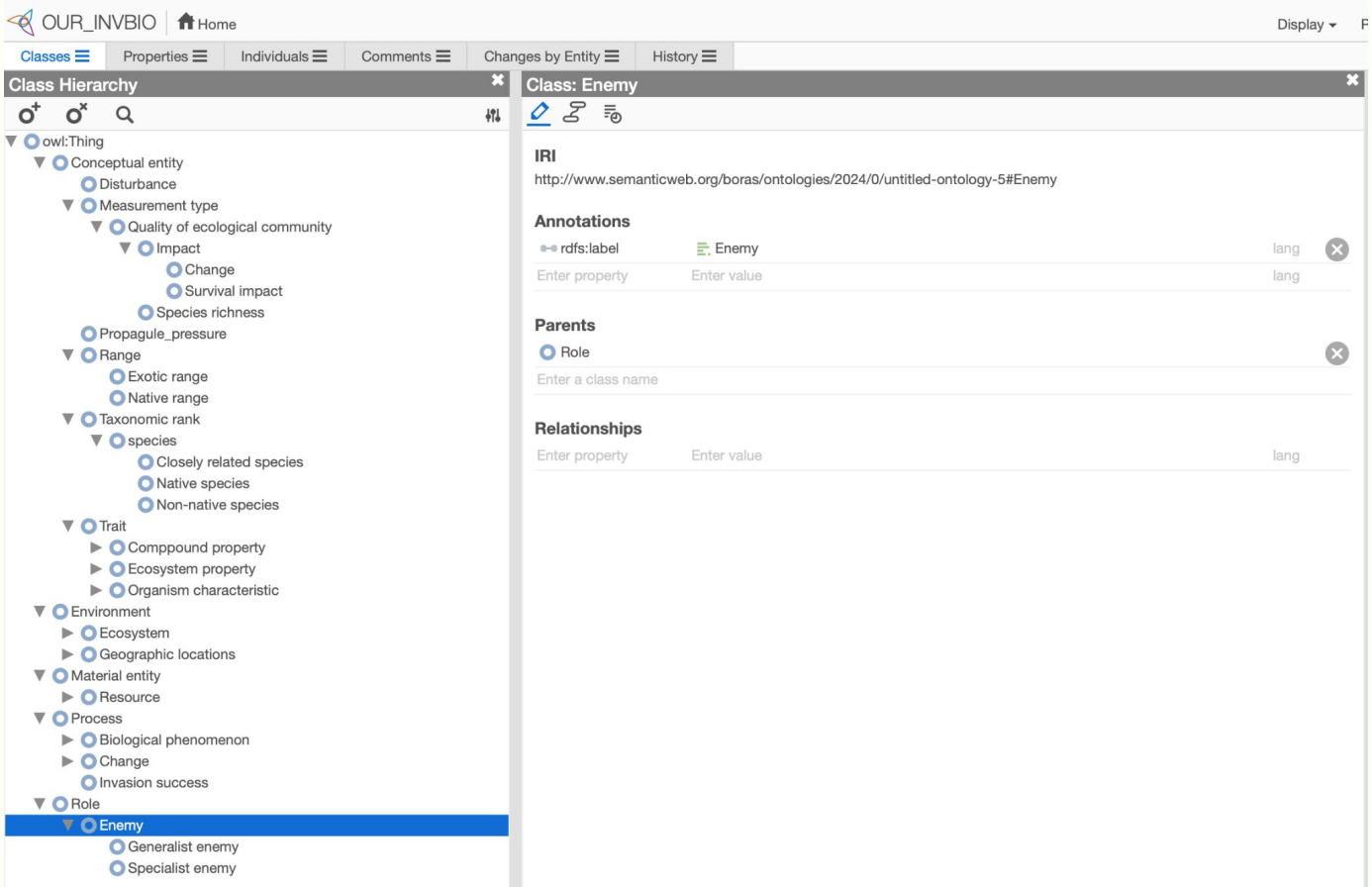
| | A | B | C | D | E | F | G | H | I | J | K | L | M |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | value | comments | classes | parents | children | relationship with | prefLabel | definition | id | ontology_url | synonyms | children (imported when INBIO) | parents |
| 2 | Enemy release | split to release and enemy | enemy | | specialist, generalist | | natural enemy | | | https://data.bioor | ['天敵', 'natural enemy'] | generalist, parasite, predator, specialist | role |
| 3 | Enemy release | | release | | | | | ['To make available; set free.'] | | | | data release; in our case enermy release | action |
| 4 | Range | | range | | exotic, native | | range | ['The difference between the lowe | | https://data.bioor | Range | (types of ranges) | conceptua |
| 5 | Invasion success | | invasion | | | | | ['The process of transportation, establishment and | | | ['biological invasion', 'species invasion'] | | process |
| 6 | Ecosystem | | ecosystem | | undistributed, distributed | | Ecosystem | ['A functional system which includ | | https://data.bioor | ['Ecological System', 'Ecosystems', 'Systems, Ecological', 'System, Ec | | environme |
| 7 | Biodiversity | | biodiversity | | | | Biodiversity | ['The variety of all native living org | | https://data.bioor | ['Biological Diversity', 'Diversity, Biological'] | | ecosystem |
| 8 | Species | | species | | native, non-native, closely related, lack of non-native | | species richness | ['A group of organisms that differ f | | https://data.bioontology.org/ontologies/NCIT | | | Taxon/Tax |
| 9 | Genetic changes | | genetic change | change | | | | ['Having to do with information that is passed from parents to offspring through genes in sperm and egg cells.'] | | | | | Molecula |
| 10 | Growth | | growth | | | | Growth | ['Gradual increase in the number, | | https://data.bioontology.org/ontologies/MESH | | | action / ph |
| 11 | Reproduction | | reproduction | | | affects (class) population size (not in hypotheses) | Reproduction | ['The total process by which orgar | | https://data.bioor | ['Period, Reproductive', 'Periods, Reproductive', 'Reproductive Index, | | life cycle / |
| 12 | Phenotypic plasticity | | phenotypic plasticity | | | | Phenotypic plasticity | | | https://data.bioor | ['Phenotypic plasticity', '表現型可塑性'] | | life history |
| 13 | Richness of closely related species | split into species richness and closely related species | species richness | biodiversity | | | species richness | ["p[Henderson's] ecological divers | | https://data.bioor | ['alpha diversity', 'richnes | - / | diversity in |
| 14 | Richness of closely related species | | closely related species - see above - | | | | | | | | | | |
| 15 | Islands | | islands | | | | Islands | ['Tracts of land completely surrour | | https://data.bioontology.org/ontologies/MESH | | | geographic |
| 16 | Continental areas | no api output for the context | area | | | islands (sibling) | | | | | | | areas |
| 17 | Ecological impacts | was added to the list | ecological impact | impact | | | Environmental impacts | | | | | | Ecology |
| 18 | Limiting similarity | it is a property: is similar to | | | | | Similarity | ['The quality of being alike; having | | https://data.bioor | ['Like', 'Similar', 'Similarity'] | | |
| 19 | Propagule pressure | | propagule pressure | | | | | ['A composite measure consisting of the number o | | | ['introduction effort'] | | conceptua |
| 20 | Introduction events | nothing similar found in the context of inbio | event | | introduction | | | | | | | | prospectiv |
| 21 | Disturbance | | disturbance | | | | Disturbance | ['A departure or divergence from t | | https://data.bioor | ['Disturbance'] | | conceptua |
| 22 | Biotic resistance | resistance | | | | | | ['[NCIT] A Resistance Process consists of activities of biologic molecules or complexes involved in processes that maintain | | | | | diversity |
| 23 | Energy allocation | split to energy and allocat | energy | | | | | ['The capacity of a physical system to do work.'] | | | | | resource |
| 24 | Energy allocation | hypo name | allocation | | | | | ['To set apart or re-distribute for a special purpose'] | | | | | process |
| 25 | Defense mechanisms | | defense | | | | | ['A trait or behaviour that is useful for repelling an enemy.', 'Protection from harm.'] | | | | | compound |
| 26 | Darwin's naturalization | no outputs; hypo name => area | | | | | | Darwin's naturalization hypothesis predicts that invaders less related to native flora are more likely to be successful than those that | | | | | |
| 27 | Invasional meltdown | no outputs; the name of 11th hypo => new class "survival" | | | | | | The invasional meltdown hypothesis (IMH) posits that positive interactions among invaders initiate positive population-level feedba | | | | | |

# Step 6: Define classes and their hierarchy



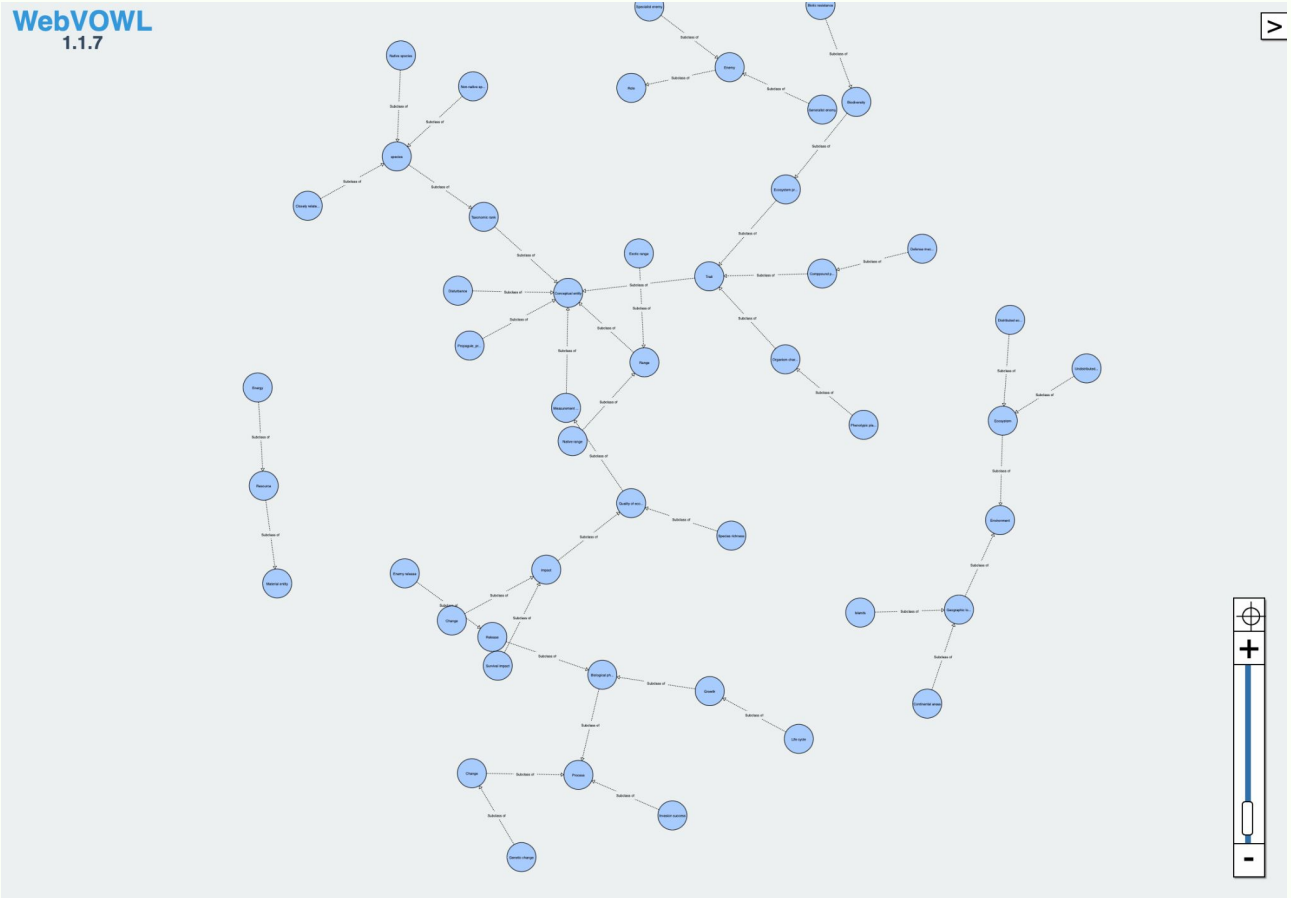
Click

# Step 7: Convert into OWL format

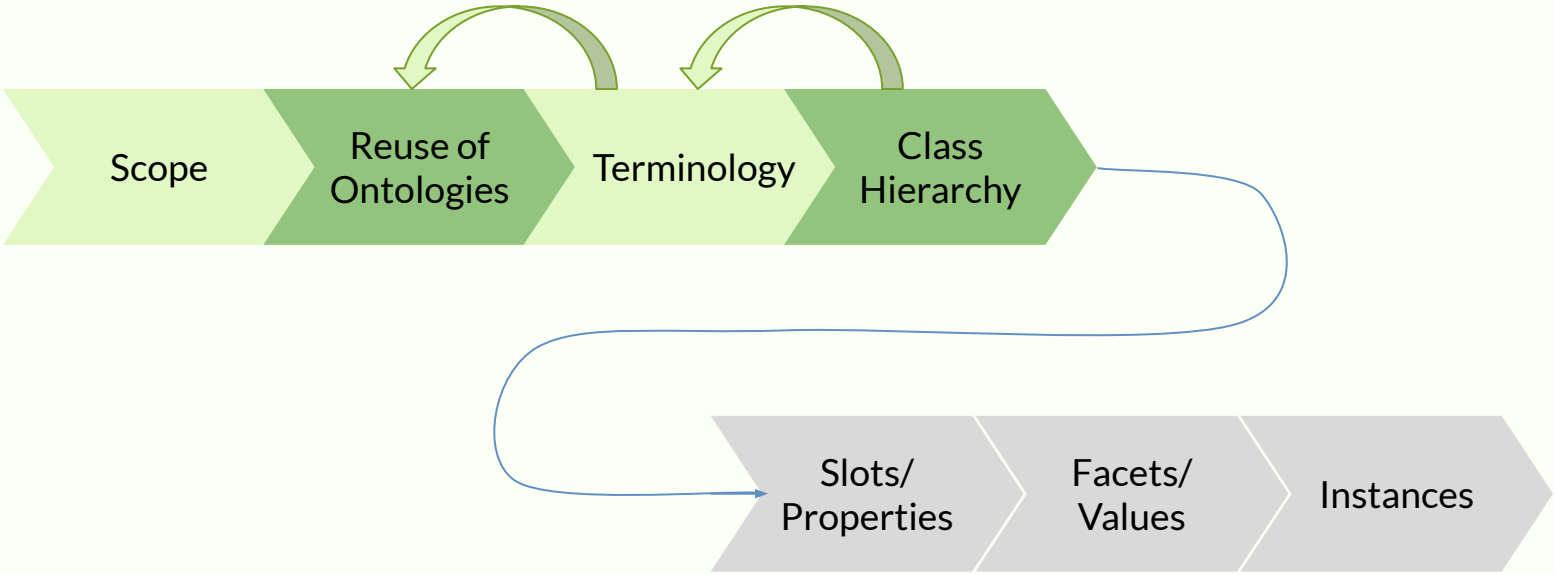


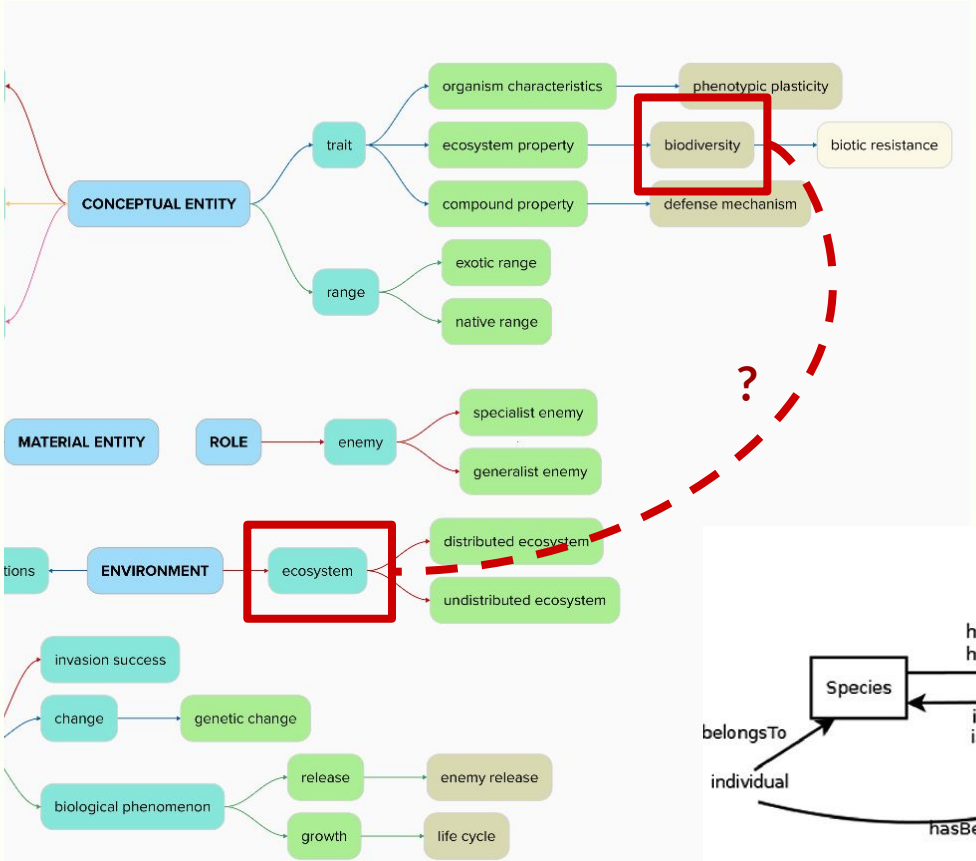
Click

# Visualization



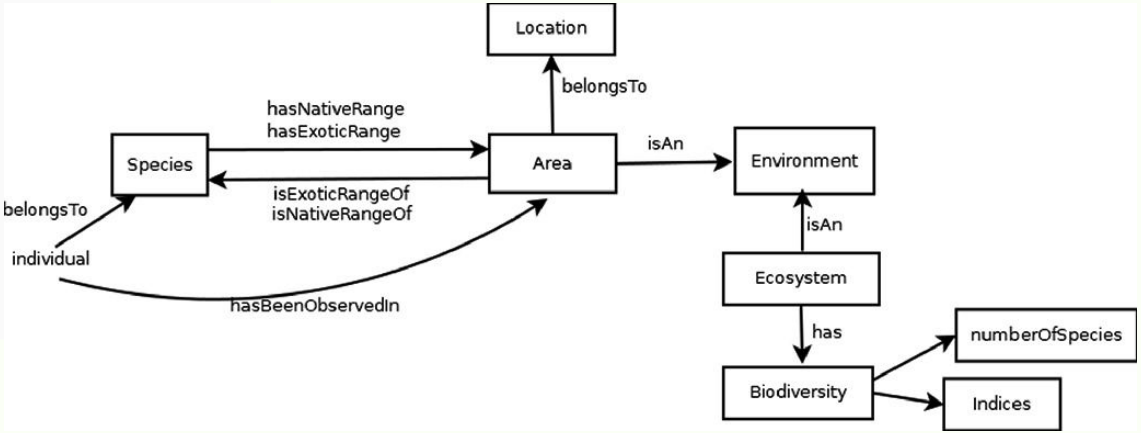
# Ontology Development Pipeline






*Hypothesis II -* ***Biotic resistance hypothesis:***

*"An ecosystem with high biodiversity is more resistant against non-native species than an ecosystem with lower biodiversity"*



# Conclusion

## Hierarchy-of-Hypotheses

**H01 – Enemy release hypothesis**

**H02 – Biotic resistance hypothesis**

**H03 – Evolution of increased competitive ability**

**H04 – Shifting defense hypothesis**

**H05 – Phenotypic plasticity hypothesis**

**H06 – Darwin's naturalization hypothesis**

**H07 – Island susceptibility hypothesis**

**H08 – Limiting similarity hypothesis**

**H09 – Propagule pressure hypothesis**

**H10 – Disturbance hypothesis**

**H11 – Invasional meltdown hypothesis**