

RECOMMENDATION REPORTS AND METRICS FOR VARIOUS ONTOLOGIES AS NEEDED IN THE PAPER:

Original GO.OWL:

```
{
  "name": "go_converted.owl",
  "timestamp": "2025-08-18T08:47:00.291835Z",
  "metrics": {
    "ANOnto": 5.954898607378451,
    "AROnto": 0.0,
    "CBOnto": 1.3602420705678933,
    "CROnto": 0.0,
    "DITOnto": 13.0,
    "INROnto": 1.4119716589298803,
    "LCOMOnto": 3.6167058655833513,
    "NACOnto": 1.1674150096215523,
    "NOCOnto": 1.561355162911331,
    "NOMOnto": 0.1818714879061813,
    "POnto": 1.230100171023699,
    "PROnto": 0.8858912257035992,
    "RFCOnto": 5811.363916355974,
    "RROnto": 0.11410877429640076,
    "TMOnto": 1.230100171023699,
    "WMCOnto": 1.5938431468360617,
    "modularityScore": 2.954085217403955,
    "reusabilityScore": 5825.938517898373,
    "analysabilityScore": 5831.116578926868,
    "changeabilityScore": 5832.677934089779,
    "modificationStabilityScore": 5819.496062601873,
    "testabilityScore": 5831.116578926868
  },
  "subcharacteristics": {
    "formalRelationsSupport": 0.11410877429640076,
    "redundancy": 5.954898607378451,
    "tangledness": -1.230100171023699,
    "cohesion": -3.6167058655833513,
    "controlledVocabulary": 5.954898607378451,
    "schemaValueReconciliation": 0.11410877429640076,
    "consistentSearchQuery": 7.480979040604732,
    "knowledgeAcquisition": 5.88713589376867,
  }
}
```

```

    "clustering": 5.954898607378451,
    "similarity": 0.11410877429640076,
    "indexingLinking": 1.526080433226281,
    "resultRepresentation": 0.0,
    "guidance": 1.4119716589298803,
    "decisionTrees": 0.18187148790618135,
    "knowledgeReuse": 3.568292912818798,
    "inference": 0.11410877429640076,
    "replaceability": 16.337069797653573,
    "adaptability": 5829.061228224195,
    "learnability": 5819.677934089778,
    "modularity": 2.954085217403955,
    "reusability": 5829.061228224195,
    "analyzability": 5831.116578926867,
    "changeability": 5832.677934089778,
    "modificationStability": 5817.902219455036
  }
}

```

GO.OWL (10% perturbations):

```

{
  "name": "go_10p_converted.owl",
  "timestamp": "2025-08-18T08:42:02.582688Z",
  "metrics": {
    "ANOnto": 1.9308298373068657,
    "AROnto": 0.0,
    "CBOnto": 1.1704233750745379,
    "CROnto": 0.0,
    "DITOnto": 12.0,
    "INROnto": 1.0471707766294893,
    "LCOMOnto": 3.9488520408163263,
    "NACOnto": 0.9838010204081633,
    "NOCOnto": 1.2205128205128206,
    "NOMOnto": 0.0429755448685153,
    "POnto": 1.0041952317609741,
    "PROnto": 0.9605781865965834,
    "RFCOnto": 512.6153846153846,
    "RROnto": 0.03942181340341656,
    "TMOnto": 1.0041952317609741,
    "WMCOnto": 1.0901463214980047,

```

```

    "modularityScore": 2.2605696965725426,
    "reusabilityScore": 525.6984170363129,
    "analyzabilityScore": 530.867781897642,
    "changeabilityScore": 532.0882947181549,
    "modificationStabilityScore": 520.0453191732863,
    "testabilityScore": 530.867781897642
  },
  "subcharacteristics": {
    "formalRelationsSupport": 0.03942181340341656,
    "redundancy": 1.9308298373068657,
    "tangledness": -1.0041952317609741,
    "cohesion": -3.9488520408163263,
    "controlledVocabulary": 1.9308298373068657,
    "schemaValueReconciliation": 0.03942181340341656,
    "consistentSearchQuery": 3.0174224273397714,
    "knowledgeAcquisition": 1.927276105841767,
    "clustering": 1.9308298373068657,
    "similarity": 0.03942181340341656,
    "indexingLinking": 1.0865925900329059,
    "resultRepresentation": 0.0,
    "guidance": 1.0471707766294893,
    "decisionTrees": 0.04297554486851518,
    "knowledgeReuse": -1.013826971748487,
    "inference": 0.03942181340341656,
    "replaceability": 14.353634686879339,
    "adaptability": 528.1394426773385,
    "learnability": 520.0882947181548,
    "modularity": 2.2605696965725426,
    "reusability": 528.1394426773385,
    "analyzability": 530.867781897642,
    "changeability": 532.0882947181548,
    "modificationStability": 518.9551728517883
  }
}

```

GO.OWL (Selected seed terms for modularization):

```

{
  "AR": [
    {
      "term": "GO_0016891",

```

```

    "iri": "http://www.geneontology.org/owl/#GO_0016891"
  }
],
"CR": [
  {
    "term": "GO_0016891",
    "iri": "http://www.geneontology.org/owl/#GO_0016891"
  }
],
"DIT": [
  {
    "term": "GO_0015599",
    "iri": "http://www.geneontology.org/owl/#GO_0015599"
  }
],
"RFC": [
  {
    "term": "GO_0030704",
    "iri": "http://www.geneontology.org/owl/#GO_0030704"
  }
],
"RR": [
  {
    "term": "GO_0016891",
    "iri": "http://www.geneontology.org/owl/#GO_0016891"
  }
]
]
}

```

GO.OWL (with 10% perturbations) full basic recommendation report:

Basic Ontology Recommendations

Overview

This ontology has good adaptability and learning potential but struggles with basic organization. Key issues include shallow relationships, missing data instances, and an overly complex hierarchy that needs simplification.

Key Strengths

1. ****High Adaptability****: Scores show excellent potential for modification and reuse
2. ****Easy to Learn****: Clear structure helps new users understand the system
3. ****Good Search Support****: The current setup helps users find information effectively

Areas for Improvement

1. Weak property-class connections (AROnto = 0)
2. No real-world examples (CROnto = 0)
3. Overly deep hierarchy (DITOnto = 12)
4. Too many properties per class (RFOnto = 512)

Simple Recommendations

1. ****Add Basic Descriptions to Classes****

- **Why**: Your classes (like AR, CR, DIT) lack connecting properties
- **How**: Add 2-3 simple properties (like "has_function" or

"related_to") to each class

2. ****Create Example Instances****

- **Why**: Classes have no real-world examples to show how they work
- **How**: Add 3-5 concrete examples per class (e.g., "AR_Example1" for

AR class)

3. ****Simplify the Family Tree****

- **Why**: The hierarchy is 12 levels deep - too complex for most users
- **How**: Combine similar classes (like AR/CR/RR that all use

GO_0016891) and flatten the structure

Next Steps

Start by merging duplicate-seeming classes (AR/CR/RR) and add 2-3 simple properties to each remaining class. Review the metrics monthly to track improvement.

GO.OWL(10% perturbations) modular basic recommendation report:

Basic Ontology Recommendations

Overview

This ontology has good adaptability and ease of learning but struggles with core relationships and structure. Key gaps include missing instances, limited property definitions, and weak connections between concepts.

Key Strengths

1. ****Easy to adapt****: Scores high in adaptability (4.5/5) for use in different contexts
2. ****Reusable components****: Parts can be reused in other ontologies (4.5/5)
3. ****Clear diagnostics****: Easy to analyze for issues (5.5/5)

Areas for Improvement

1. ****Missing relationships****: Only basic "is a type of" relationships exist (formalRelationsSupport: 0.0)
2. ****Empty classes****: No real-world examples/instances (CROnto: 0.0)
3. ****Duplicate information****: Some terms like GO_0016891 appear multiple times

Simple Recommendations

1. ****Add relationship types****
 - Why: Currently only uses basic "type of" connections
 - How: Create new relationships like "part_of" or "participates_in" between "vitelline membrane formation" and related processes
2. ****Add real-world examples****
 - Why: Classes like "vitelline membrane formation sensu insecta" have no concrete instances
 - How: Add specific insect examples (e.g., "Drosophila vitelline formation") with dates/researchers
3. ****Merge duplicate terms****
 - Why: GO_0016891 appears in AR, CR, and RR - could cause confusion
 - How: Either combine these into one parent concept or clarify their unique roles

Next Steps

Start by reviewing the repeated GO_0016891 entries, then add 2-3 simple relationship types between your seed terms. Consider consulting an entomologist to identify key real-world examples to populate the classes.

UNIV-BENCH-OWL2DL.owl scores:

```
{
  "name": "UNIV-BENCH-OWL2DL_converted.owl",
  "timestamp": "2025-06-02T08:41:11.658949Z",
  "metrics": {
    "ANOnto": 0.0227272727272728,
    "AROnto": 0.52272727272727,
    "CBOnto": 0.991869918699187,
    "CROnto": 0.553030303030303,
    "DITOnto": 6.0,
    "INROnto": 0.92424242424242,
    "LCOMOnto": 2.8272727272727,
    "NACOnto": 0.98181818181818,
    "NOCOnto": 0.991869918699187,
    "NOMOnto": 0.0,
    "POnto": 0.92424242424242,
    "PROnto": 1.0,
    "RFCOnto": 0.0,
    "RROnto": 0.0,
    "TMOnto": 0.92424242424242,
    "WMCOnto": 0.92424242424242,
    "modularityScore": 1.916112342941611,
    "reusabilityScore": 6.9242424242424,
    "analysabilityScore": 10.743385070214337,
    "changeabilityScore": 11.735254988913525,
    "modificationStabilityScore": 5.735254988913525,
    "testabilityScore": 10.743385070214337
  },
  "subcharacteristics": {
    "formalRelationsSupport": 0.0,
    "redundancy": 0.0227272727272728,
    "tangledness": -0.92424242424242,
    "cohesion": -2.8272727272727,
    "controlledVocabulary": 0.0227272727272728,
    "schemaValueReconciliation": 0.52272727272727,
    "consistentSearchQuery": 1.46969696969697,
  }
}
```

```

    "knowledgeAcquisition": 0.022727272727272728,
    "clustering": 0.022727272727272728,
    "similarity": 0.5227272727272727,
    "indexingLinking": 1.4469696969696968,
    "resultRepresentation": 1.0757575757575757,
    "guidance": 1.4469696969696968,
    "decisionTrees": 0.5227272727272726,
    "knowledgeReuse": -1.3575757575757574,
    "inference": 0.5530303030303030,
    "replaceability": 7.916112342941611,
    "adaptability": 8.907982261640798,
    "learnability": 5.735254988913525,
    "modularity": 1.9161123429416111,
    "reusability": 8.907982261640798,
    "analyzability": 10.74338507021434,
    "changeability": 11.735254988913526,
    "modificationStability": 4.811012564671101
  }
}

```

UNIV-BENCH_OWL2DL_advanced_recommendations:

Advanced Ontology Analysis and Recommendations

1. Comprehensive Assessment

The ontology demonstrates strong structural flexibility (changeabilityScore=11.74) but suffers from critical semantic deficiencies. Key weaknesses include:

- ****Severe annotation poverty**** (ANOnto=0.0227): 98% below best practice thresholds
- ****Relational impoverishment**** (RROnto=0.0): Complete absence of non-taxonomic relationships
- ****Hierarchical entanglement**** (TMOnto=-0.9242): Excessive multi-inheritance patterns
- ****Cohesion deficit**** (LCOnto=-2.8273): Weak conceptual clustering

Strengths lie in:

- Deep hierarchical organization (DITOnto=6.0)
- High modifiability (changeability=11.74)
- Good query support (consistentSearchQuery=1.47)

2. Technical Analysis by Key Quality Dimensions

Structural Dimension

Metrics Assessment:

- WMCOnto=1.0 (suboptimal schema density)
- TMOnto2=3.2 (excessive multi-parent classes)

Identified Issues:

1. Overuse of disjoint unions creating rigid taxonomies
2. Property starvation (0 object properties beyond taxonomy)
3. Redundant equivalences (e.g., "basket ball" \equiv "basketball")

Technical Recommendations:

1. Convert 40% of disjoint unions to property restrictions using OWL ObjectPropertyDomain/Range
2. Implement property chains for role hierarchies (e.g., ``isHeadOf o inverse(hasHead) \sqsubseteq holdsPosition``)
3. Normalize lexical variants using skos:altLabel instead of owl:sameAs

Functional Dimension

Metrics Assessment:

- KnowledgeReuse=-1.36 (low export potential)
- Inference=0.55 (limited reasoning capacity)

Identified Issues:

1. Missing property axioms prevent logical inference
2. No cross-hierarchy relationships between domains
3. Absence of SWRL rules or property chains

Technical Recommendations:

1. Introduce 5-7 core object properties connecting seed terms:
``owl
ObjectProperty: teachesCourse
Domain: Faculty
Range: Course
``
2. Implement inverse properties for organizational structure:
``owl
InverseObjectProperties(:hasHead :headOf)
``

```

3. Add transitivity to academic roles:
    ```owl
 TransitiveObjectProperty(:supervises)
    ```

### Semantic Dimension
**Metrics Assessment**:
- ControlledVocabulary=0.02 (terminological inconsistency)
- Similarity=0.52 (weak comparability)

**Identified Issues**:
1. Mixed naming conventions (e.g., "petroleuml engineering")
2. Undifferentiated class labels ("college" ≡ "school")
3. Missing semantic typing of relationships

**Technical Recommendations**:
1. Apply OWL annotation patterns:
    ```owl
 Class: Program
 rdfs:label "Academic Program"@en
 skos:definition "Structured set of courses..."@en
    ```
2. Implement lexical reconciliation:
    ```owl
 Class: PetroleumEngineering
 skos:prefLabel "Petroleum Engineering"@en
 skos:altLabel "Petroleuml Engineering"@en
    ```
3. Add property hierarchies:
    ```owl
 SubObjectPropertyOf(:teachesCourse :participatesInAcademicActivity)
    ```

## 3. Critical Metrics Improvement Plan

### ANOnto (Current score: 0.0227)
**Technical Explanation**: Measures annotation density per class. Current score indicates <0.03 annotations/class versus ideal >0.8.

```

****Impact****: Directly affects 6 subcharacteristics including KnowledgeReuse (+0.4 per 0.1 ANOnto gain) and Clustering (+0.25).

****Seed Terms****: Program (<http://benchmark/OWL2Bench#Program>)

****Improvement Approach****:

1. Annotate Program class with 5 standard properties:

```
```owl
Class: Program
 rdfs:comment "Organized academic curriculum..."@en
 dcterms:created "2023-07-20"^^xsd:date
 skos:example "Undergraduate CS Program"@en
```
```

2. Implement annotation propagation:

```
```owl
AnnotationProperty: academicYear
SubAnnotationPropertyOf(academicYear dcterms:date)
```
```

3. Expected improvement: ANOnto=0.42 (+0.4) with 15 annotations across 36 classes

RROnto (Current score: 0.0)

****Technical Explanation****: Ratio of relationships to total axioms. Current 0 value indicates pure taxonomy.

****Impact****: Critical for FormalRelationsSupport (current 0) and Similarity (+0.2 per 0.1 RROnto gain).

****Seed Terms****: Program (<http://benchmark/OWL2Bench#Program>)

****Improvement Approach****:

1. Create 3 core relationships for Program:

```
```owl
ObjectProperty: hasCurriculum
Domain: Program
Range: Course
SubPropertyOf: hasComponent
```
```

2. Convert 20% subclass axioms to property restrictions:

```
```owl
```

```
Class: CoEdCollege
EquivalentTo: College and (hasAdmissionPolicy value CoEd)
...
```

3. Expected improvement: RROnto=0.35 with 12 new properties/restrictions

## ## 4. Implementation Roadmap

### ### Immediate Technical Improvements (Week 1-2)

1. Annotate all seed terms with minimum:
  - 1 rdfs:comment
  - 1 skos:definition
  - 1 dcterms:created
2. Implement 5 core object properties connecting Program to Department, Course, and Faculty
3. Fix lexical inconsistencies using skos:altLabel

### ### Medium-term Refactoring (Month 1-3)

1. Modularize into 3 components:
  - AcademicStructure (organizations/roles)
  - Curriculum (courses/programs)
  - ResearchOutput (publications/articles)
2. Replace 30% of disjoint unions with property-based restrictions
3. Introduce property chains for role inheritance

### ### Long-term Quality Strategy

1. Implement CI/CD pipeline with:
  - OOPS! for structural validation
  - ROBOT for annotation checks
  - SPARQL-based metric monitoring
2. Establish annotation governance policy requiring:
  - Minimum 3 annotations per new class
  - Property typing for all relationships
3. Integrate with upper ontologies:

```
```owl
Import(<http://purl.org/vocab/frbr/core>)
Import(<http://www.w3.org/2004/02/skos/core>)
```
```

This technical plan prioritizes metric-driven improvements while maintaining structural integrity. Focus on property engineering and

annotation rigor will yield 40-60% gains in target metrics within 3 development cycles.

## UNIV-BENCH\_OWL2DL\_basic\_recommendations:

### # Basic Ontology Recommendations

#### ## Overview

This ontology organizes academic concepts like engineering disciplines, staff roles, and publications, but needs better documentation and richer relationships. It shows good adaptability but lacks descriptive details and connections between concepts.

#### ## Key Strengths

1. **\*\*Flexible structure\*\***: Scores high in adaptability (8.9/10) and modifiability (11.7/10), meaning it's easy to update and reuse
2. **\*\*Clear organization\*\***: Handles academic hierarchies well (e.g., professor types, college disciplines)
3. **\*\*Good foundation for expansion\*\***: Scores well in guidance (1.4/10) for domain knowledge specification

#### ## Areas for Improvement

1. **\*\*Missing descriptions\*\***: Classes lack annotations/metadata (ANOnto score 0.02/5)
2. **\*\*Limited relationships\*\***: Only uses basic "is a" relationships (RROnto score 0/5)
3. **\*\*Potential duplicates\*\***: Some terms like "basket ball" vs "basketball" need cleanup
4. **\*\*Shallow connections\*\***: No properties linking concepts like "Program" to departments or courses

#### ## Simple Recommendations

1. **\*\*Add simple class descriptions\*\***
  - *\*Why\**: Helps users understand what each term means (e.g., difference between "assistant professor" and "associate professor")
  - *\*How\**: Add 1-2 sentence explanations using ``rdfs:comment`` for each class
2. **\*\*Connect related concepts with properties\*\***
  - *\*Why\**: Currently missing relationships like "teaches" between faculty and courses, or "publishedIn" for articles

```
- *How*:
 1. Create object properties like "hasDepartment" between Program and
Department
 2. Add "teachesCourse" between Professor and Course

3. **Clean up duplicates and add synonyms**
 - *Why*: Terms like "basket ball" (two words) and "basketball" (one
word) might confuse users
 - *How*:
 1. Use `owl:sameAs` for true duplicates
 2. Add `skos:altLabel` for alternative spellings
```

## **## Next Steps**

Start by adding descriptions to 5 key classes (like "Program" and "faculty"), then create 2-3 new relationships connecting courses to departments. Use free tools like Protégé to implement these changes and validate with reasoning tools.