

# BioTop - A Top-Domain Ontology for the Life Sciences

Stefan Schulz<sup>1</sup>, Holger Stenzhorn<sup>1</sup>, Martin Boeker<sup>1</sup>, Elena Beisswanger<sup>2</sup>, Udo Hahn<sup>2</sup>, László van den Hoek<sup>3</sup>, Erik van Mulligen<sup>3</sup>

<sup>1</sup> Institute for Medical Biometry and Medical Informatics, University Medical Center Freiburg, Freiburg, Germany

<sup>2</sup> Jena University Language and Information Engineering (JULIE) Lab, Jena, Germany

<sup>3</sup> Department of Medical Informatics, Erasmus University Medical Center, Rotterdam, The Netherlands



## Background

The need for semantic standardization in the Life Sciences has been addressed by a dynamic evolution of domain ontologies (e.g. OBO, SNOMED CT), which tend to adhere to foundational principles of ontology design.

Current biomedical ontologies are characterized by

- large fragmentation and overlap
- missing cross-ontology links
- lack of clear and unambiguous formal definitions of basic terms
- purpose-specific architecture and design decisions

## BioTop - Rationale

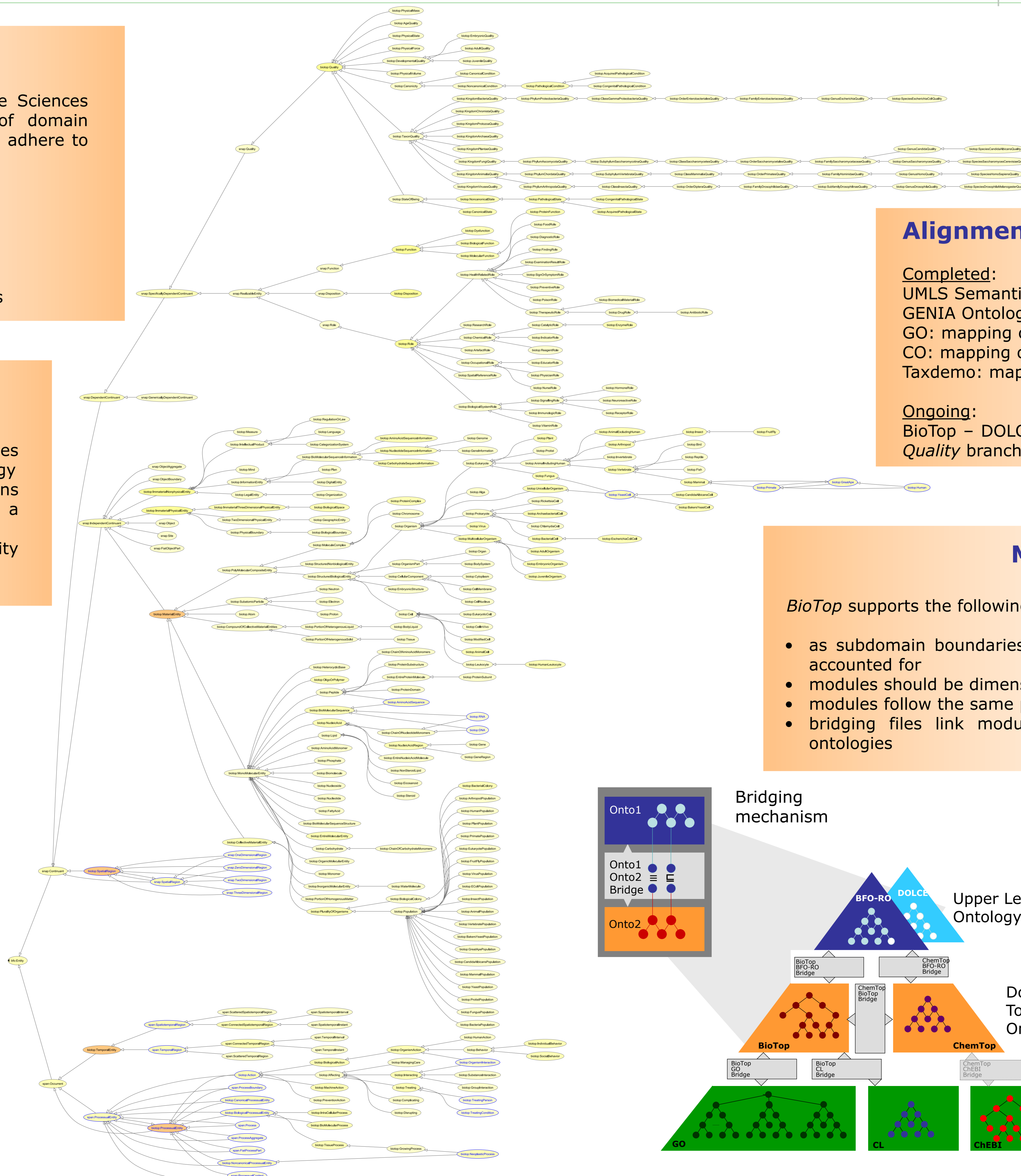
- To consolidate and integrate domain ontologies bridging the gap to a common upper level ontology
- To enforce unambiguous logic-based descriptions of basic entities of biology and medicine using a standardized description language
- To maintain neutrality with regard to granularity and observer-biased views

## BioTop - Characteristics

- 257 classes
- 42 relation types
- 193 logical restriction on classes.
- 57 sufficient criteria for full class definitions
- OWL-DL as representation language
- BFO as upper ontology
- OBO RO relations

Example of a formal class definition (*biotop:Tissue*):

<ul style="list-style-type: none"> <li>• <code>biotop:CompoundOfCollectiveMaterialEntity</code></li> <li>• <code>biotop:PortionOfHeterogeneousLiquid</code></li> <li>• <code>biotop:BodyLiquid</code></li> <li>• <code>biotop:PortionOfHeterogeneousSolid</code></li> <li>• <code>biotop:Tissue</code></li> <li>• <code>biotop:MonomolecularEntity</code></li> <li>• <code>biotop:AmphiphilicDimer</code></li> <li>• <code>biotop:BiologicalSequence</code></li> <li>• <code>biotop:BiologicalSequenceStructure</code></li> <li>• <code>biotop:BiologicalMolecule</code></li> <li>• <code>biotop:Carbohydrate</code></li> </ul>	<ul style="list-style-type: none"> <li>• <code>biotop:PortionOfHeterogeneousSolid</code></li> <li>• <code>biotop:derivesFrom some biotop:MulticellularOrganism or (biotop:properPartOf some biotop:MulticellularOrganism)</code></li> <li>• <code>biotop:hasComponentPart some (biotop:hasOranularPart some biotop:ExtracellularEntity)</code></li> <li>• <code>biotop:hasComponentPart some (biotop:hasOranularPart some biotop:Cell)</code></li> <li>• <code>biotop:hasComponentPart only biotop:CollectiveMaterialEntity</code></li> <li>• <code>biotop:hasComponentPart some biotop:CollectiveMaterialEntity</code></li> <li>• <code>biotop:hasReference some biotop:PhysicalMass</code></li> <li>• <code>biotop:hasReference some biotop:PhysicalVolume</code></li> </ul>
---	--



## Availability

<http://purl.org/biotop/>:

- Sources
- Publications
- Discussion List

## Alignment and mapping

Completed:

- UMLS Semantic Network Mapping
- GENIA Ontology mapping
- GO: mapping of the three GO ontologies
- CO: mapping of the cell ontology
- Taxdemo: mapping of sample biological taxonomy

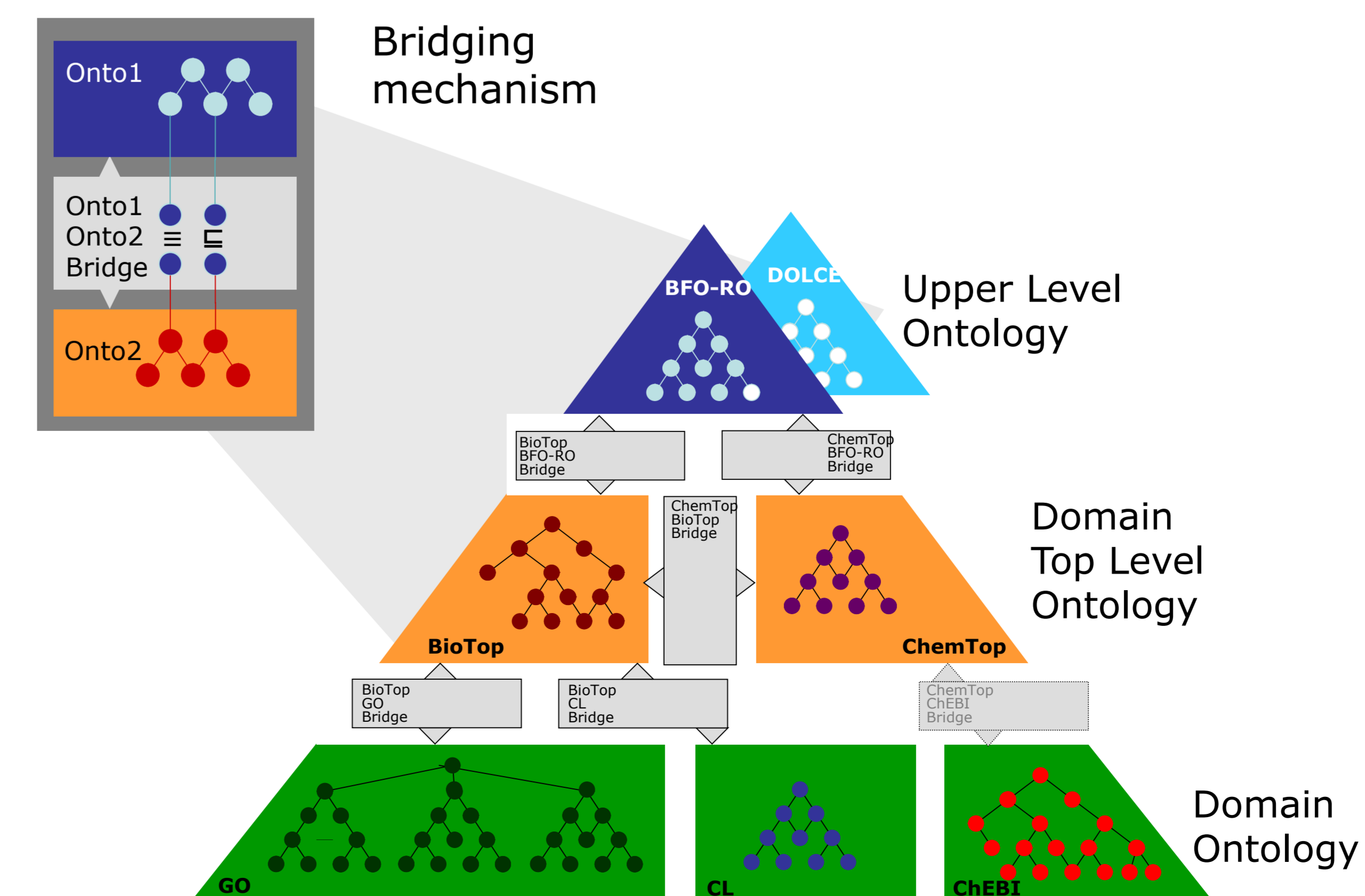
Ongoing:

BioTop – DOLCE mapping (requires redesign of the BioTop *Quality* branch)

## Modularization

BioTop supports the following modularization principles:

- as subdomain boundaries are fuzzy a limited degree of overlap must be accounted for
- modules should be dimensioned in a way that they classify rapidly
- modules follow the same principles of upper-level arrangement
- bridging files link modules between themselves and with upper-level ontologies



## ChemTop

According to the modularization principles the more detailed biochemistry descriptions were separated into an optional add-on called ChemTop.

Integration of ChemTop and ChEBI planned.