

MarineTLO & DarwinCore



N.Minadakis, C.Allocca, Y.Marketakis,
G.Skouradakis, C.Bekiari, M.Doerr, Y.Tzitzikas
14/11/2013



Outline

- What is Darwin Core (DwC)?
- What is MarineTLO?
- Mappings between DwC and MarineTLO
 - Classes
 - Properties
- Examples of modeling using MarineTLO and DwC
 - Scientific Name Assignment
 - Rights
 - Predators and Competitors
- Conclusion

What is Darwin Core?

- Darwin Core is a ***glossary of terms*** intended to facilitate the sharing of information about biological diversity.
 - It was originally conceived to facilitate the discovery, retrieval, and integration of information about modern biological specimens, their spatiotemporal occurrence, and their supporting evidence housed in collections (physical or digital).
 - The Darwin Core is based on the standards developed by the *Dublin Core Metadata Initiative* and should be viewed as an extension of the Dublin Core for biodiversity information.

What is MarineTLO?

- MarineTLO is a ***top-level ontology*** for the marine domain (also applicable to the terrestrial domain)
 - To tackle the need for having integrated sets of facts about marine species, and thus to assist research about species and biodiversity.
 - It provides a unified and coherent core model for schema mapping which enables formulating and answering queries which cannot be answered by any individual source.

MarineTLO – DwC Comparison

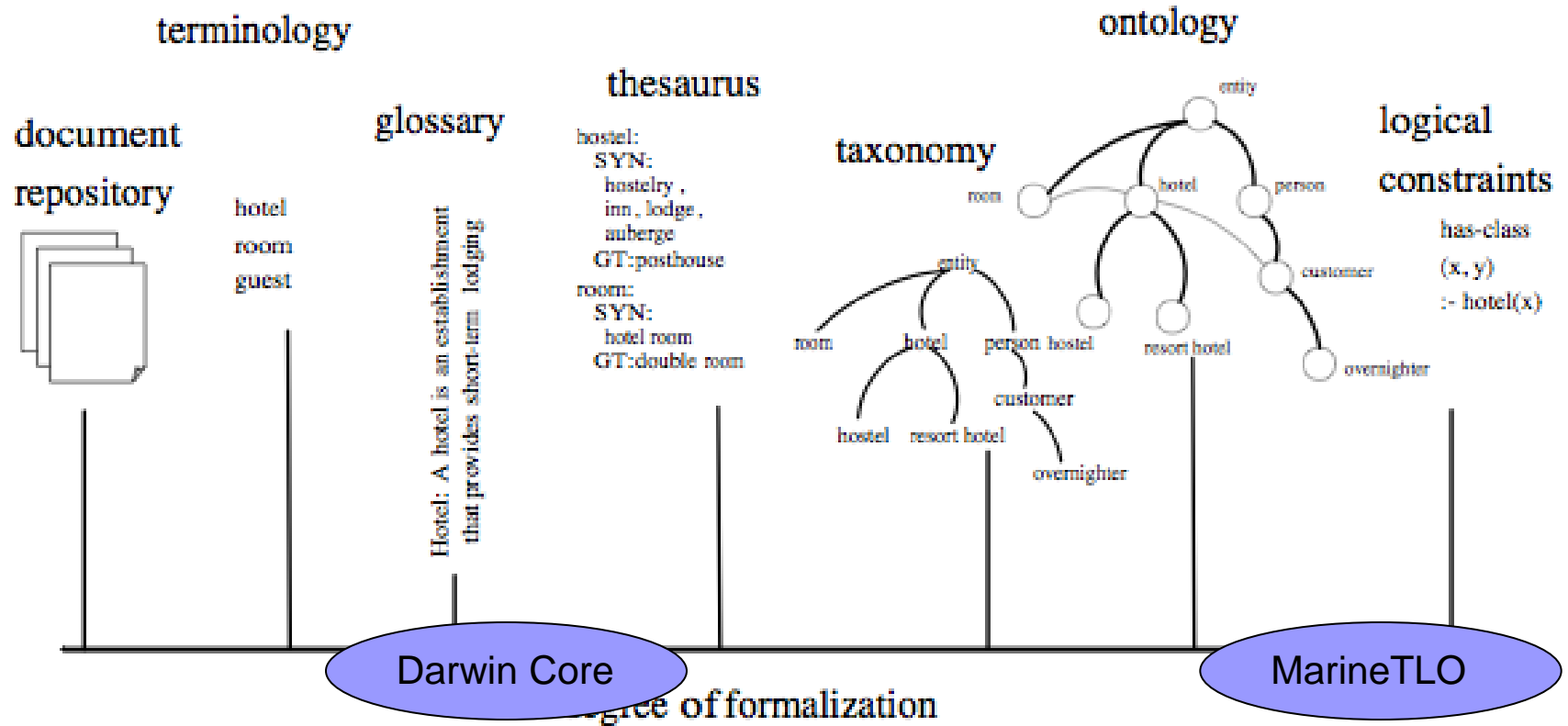


Figure 1. The different degrees of formalization: from unstructured textual content to ontology and logical rules.

Figure taken from: "From Glossaries to Ontologies: Extracting Semantic Structure from Textual Definitions", Roberto NAVIGLI and Paola VELARDI

Example of DwC2MarineTLO mappings

DwC Term	Mapping Operator	MarineTLO Class/Property
dwc: Event	<i>Sub Class Of</i>	core: BC4 Temporal Phenomenon
dwc: Identification (the assignment of a scientific name)	<i>CbTi</i>	core: BC44 Attribute Assignment
dwc: MeasurementOrFact	<i>CbTi</i>	core: BC54 Measurement
dwc: measurementType	<i>CbTi</i>	core: BC54 Measurement core: BC5 Dimension
dwc: measurementValue	<i>CbTi</i>	String Number
dwc: measurementUnit	<i>Equivalent To</i>	core: BC55 Measurement Unit
dcterms:type	<i>Sub Property Of</i>	core: LX3 has type
dwc: measurementDeterminedDate	<i>Sub Property Of</i>	core: LC5 has time span
Dwc:measurementDeterminedBy	<i>Sub Property Of</i>	core:LC13 is carried out by

CbTi: *Can be Transformed Into*

An extended mapping can be found at:

http://www.ics.forth.gr/isl/MarineTLO/files/DwC_MarineTLO_mapping_firstDraft.pdf

Scientific Names 1/2

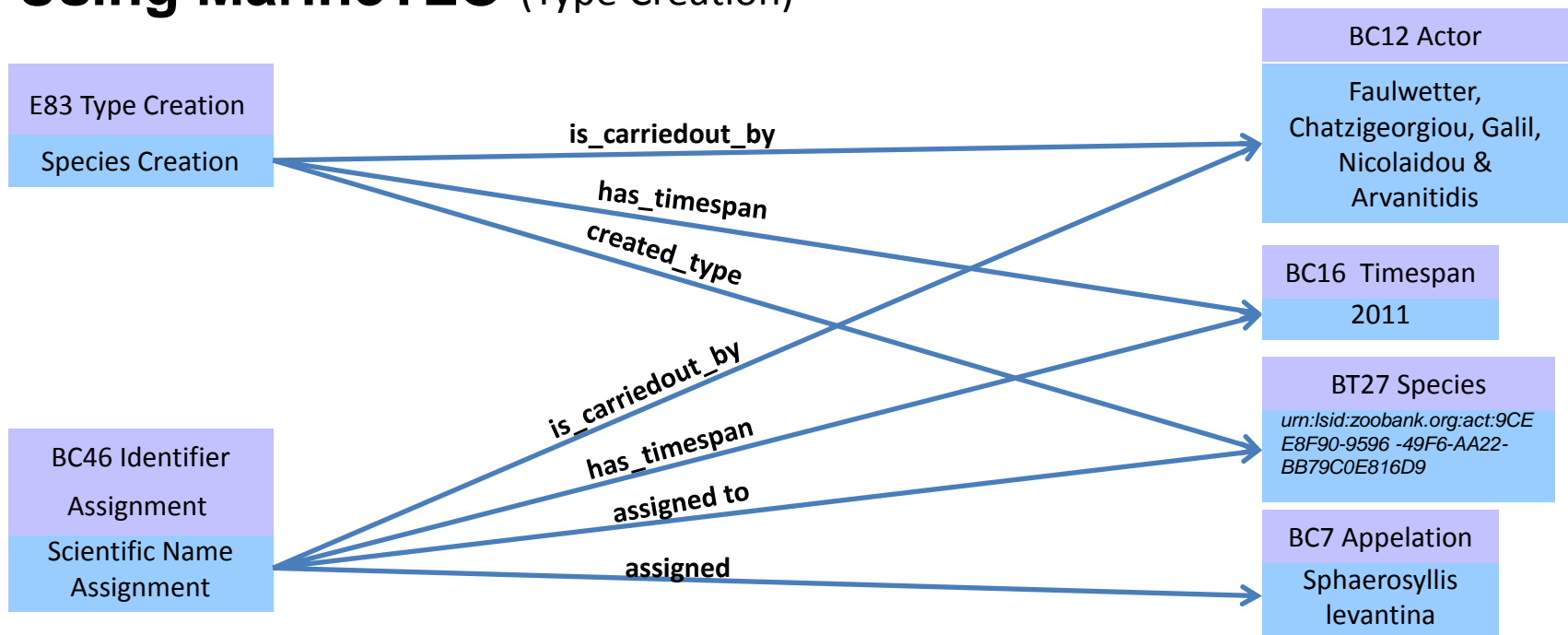
Using DwC

scientificNameID	urn:lsid:zoobank.org:act:9CEE8F90-9596-49F6-AA22-BB79C0E816D9
scientificName	Sphaerosyllis levantina
scientificNameAuthorship	Faulwetter, Chatzigeorgiou, Galil, Nicolaidou & Arvanitidis, 2011

Issues

- Select all the scientific names that were assigned in the 19th century
- The authorship given as string prevent us to further explore/navigate on it (highly required in LOD)

Using MarineTLO (Type Creation)



Rights 1/2

■ Using DwC

- Using a real occurrence dataset:

datasetID	d50024ac-5268-477e-8559-933779553b34
datasetName	EasternMedSyllids
rights	http://creativecommons.org/licenses/by/3.0/
rightsHolder	Israel Oceanographic & Limnological Research Ltd

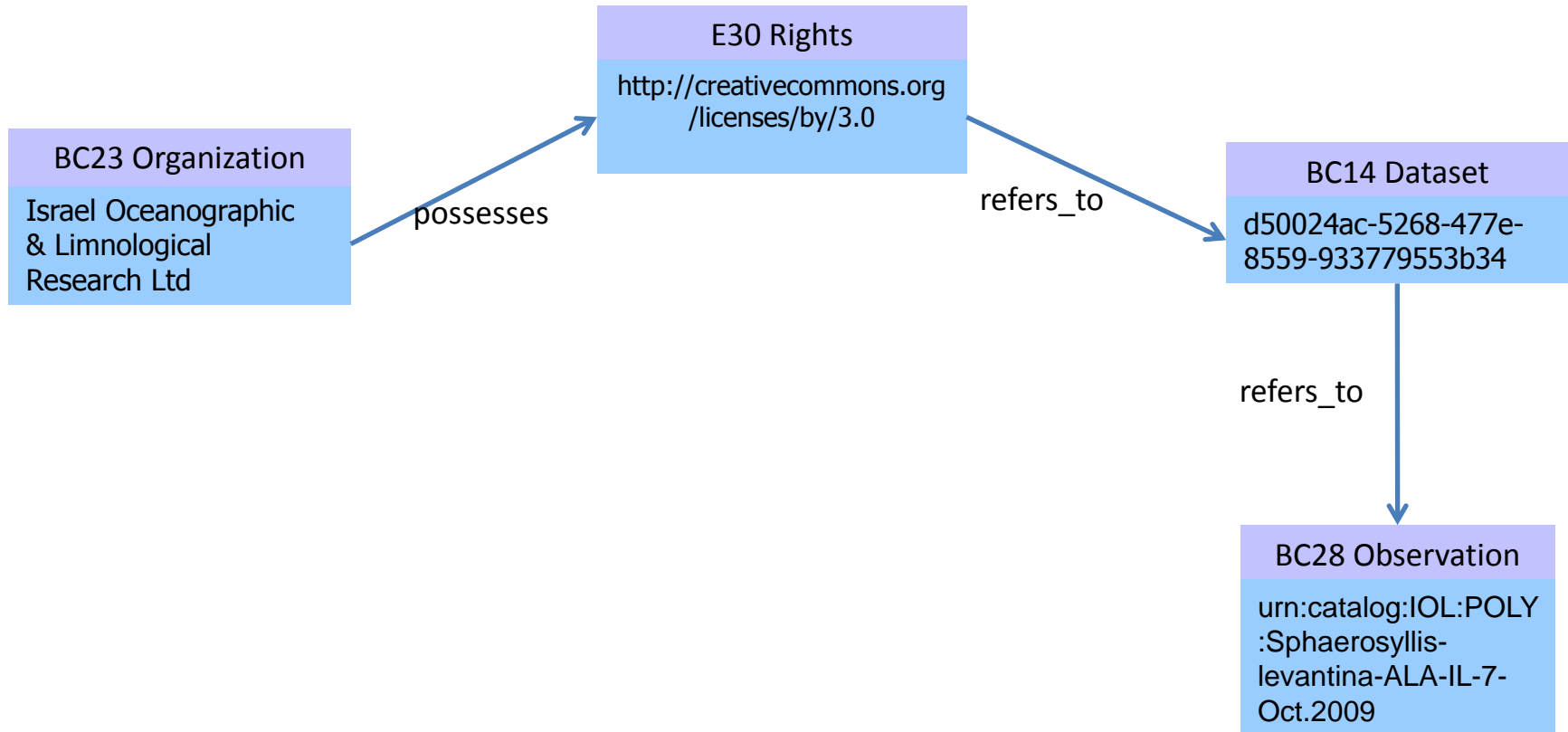
- DwC scope note:

Information about rights held in and over the resource. Typically, rights information includes a statement about various property rights associated with the resource, including intellectual property rights.

- **Issue:** Does the rights refer to the occurrence records or to the dataset?

Rights 2/2

■ Using MarineTLO



Biological Interaction Between Species 1/5

- Problem Statement

- How to express associations such as

- The species X is predator of the species Y
 - The species Y is prey of the species X

- *e.g. Thunnus Albacares is predator of Katsuwonus Pelamis*

Biological Interaction Between Species 2/5

■ Using DwC (DwC: Resource Relationship patter)

resourceID	http://www.marinespecies.org/aphia.php?p=taxdetails&id=127027
relatedresourceID	http://www.marinespecies.org/aphia.php?p=taxdetails&id=127018
relationshipofResource	predator of

Query Formulation:

```
Select ?preys
Where
{
  ?x dwc:relationshipofResource "predator of" .
  ?x dwc:resourceID <http://www.marinespecies.org/aphia.php?p=taxdetails&id=127027> .
  ?x dwc:relatedresourceID ?preys
}
```

■ Using MarineTLO (usually is predator of)

BT33 Marine Animal Type

<http://www.marinespecies.org/aphia.php?p=taxdetails&id=127027>

usually is predator of

BT33 Marine Animal Type

<http://www.marinespecies.org/aphia.php?p=taxdetails&id=127018>

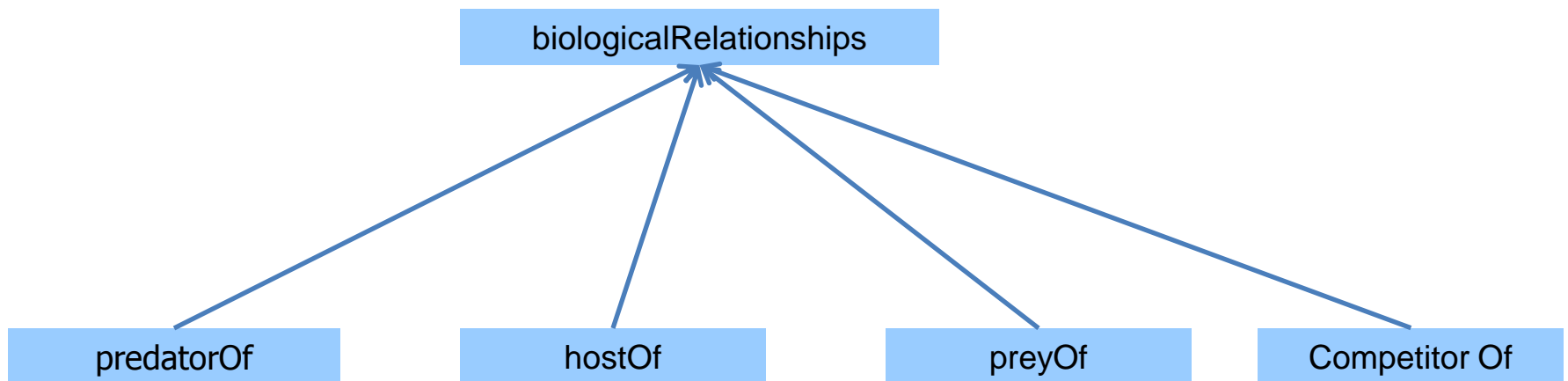
```
Select ?preys
Where
{
  <http://www.marinespecies.org/aphia.php?p=taxdetails&id=127027> tlo:is_predator_of ?preys
}
```

Biological Interaction Between Species 3/5

■ Using DwC (DwC: Resource Relationship patter)

resourceID	http://www.marinespecies.org/aphia.php?p=taxdetails&id=127027
relatedresourceID	http://www.marinespecies.org/aphia.php?p=taxdetails&id=127018
relationshipofResource	predator of

ISSUE: Using the above patter, it becomes even more complex for representing and formulating queries over more complex and structured relationships, such as:



Conclusion

- Comparing MarineTLO (*top-level ontology*) and Darwin Core (*glossary of terms*)
- MarineTLO over performs DwC in representing and query formulation
 - Scientific Name Assignment
 - Rights
 - Predators and Competitors
- As next
 - Extending MarineTLO to cover part of DwC which is currently under investigation.

NOTE: This presentation is a first draft showing the outcomes of the ongoing activity on MarineTLO and DwC comparison.