N.Minadakis, M.Doerr, C.Bekiari

Lifewatch Marine Thematic Technical Workshop 3-5 June 2014

Possible Scenario of Use

1. Data and Service Providers declare themselves as providers.

2. Providers declare the kind of data or services they wish to provide.

3. They define and provide the metadata.

4. They provide their data.

5. All the harvested metadata is mapped to a unified and normalized XML schema.

6. The system traces and retrieves terms from the metadata and co-references it with external authorities.

7. All metadata is stored in the metadata repository.

8. End user queries the data/metadata/services by using predefined queries or query formulation.
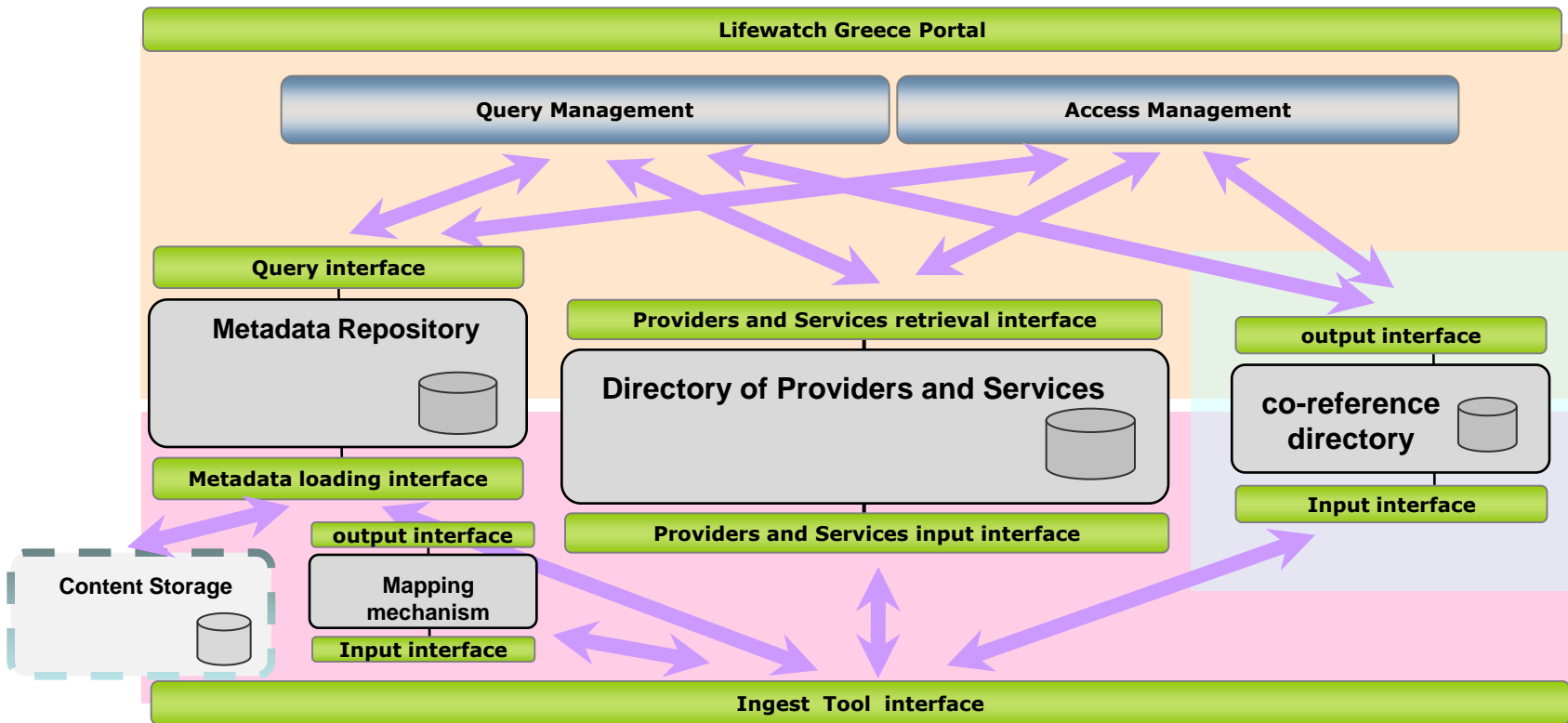
Components of the Infrastructure

1. Metadata Repository:
   Management of Metadata based on Biodiversity Ontology (based on MarineTLO, inspired by ISO 21121)
   Identifier Generation Service.

2. Directory Service (~ Lifewatch Catalogue Service):
   Keeps information on available data /services/tools provided by the infrastructure.

3. Mapping Mechanism:
   Maps the metadata from the data files described in various formats to the unique format accepted by the Metadata Repository.

4. Co-reference Directory:
   Co-reference resolution on instances/terms/services etc.

5. Ingest Tool:
   Ingest the metadata from various selected databases to the central metadata repository.

6. Access Manager:
   Authorization/authentication control.

7. Query Manager:
   Query Formulation & Execution

8. Content Storage:
   Data storage for datasets/images/maps/sequences etc.

Lifewatch Greece Infrastructure

Lifewatch Greece Portal

Query Management

Access Management

Query interface

Metadata Repository

Metadata loading interface

Content Storage

output interface

Mapping mechanism

Input interface

Providers and Services retrieval interface

Directory of Providers and Services

Providers and Services input interface

output interface

co-reference directory

Input interface

Ingest Tool interface

Ingest tool

sources

Declaration of Services and Data

Meta data

Local DB interface

Local Data Base

Data Flow Example

```
<field index="2" term="dwc:country"/>
<field index="3" term="dwc:accessRights"/>
<field index="4" term="dwc:ownerInstitutionCode"/>
<field index="5" term="dwc:rightsHolder"/>
<field index="7" term="dwc:identifiedBy"/>
<field index="8" term="dwc:institutionID"/>
<field index="9" term="dwc:datasetID"/>
<field index="10" term="dwc:collectionCode"/>
<field index="11" term="dwc:eventDate"/>
<field index="12" term="dwc:taxonRank"/>
<field index="13" term="dwc:genus"/>
<field index="14" term="dwc:catalogueNumber"/>
```
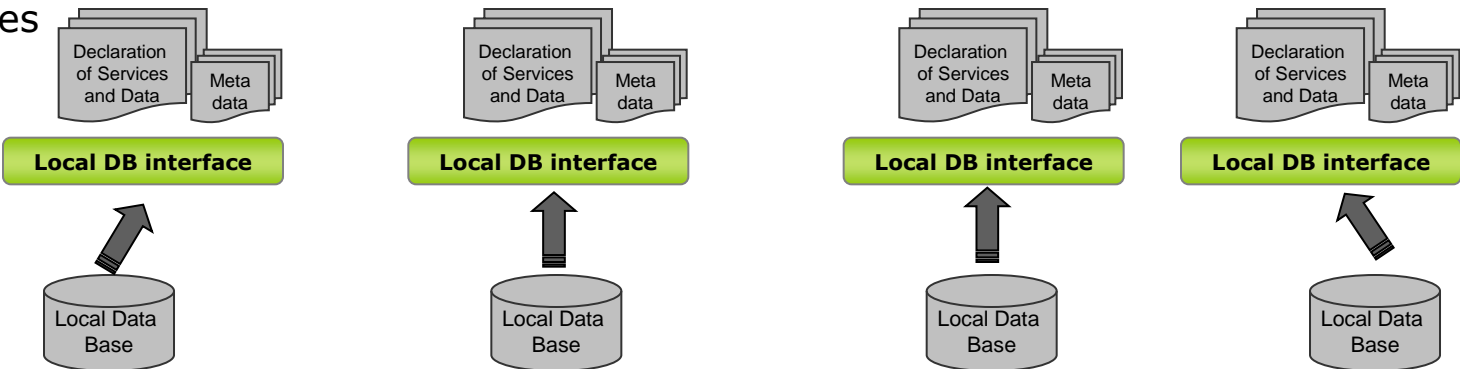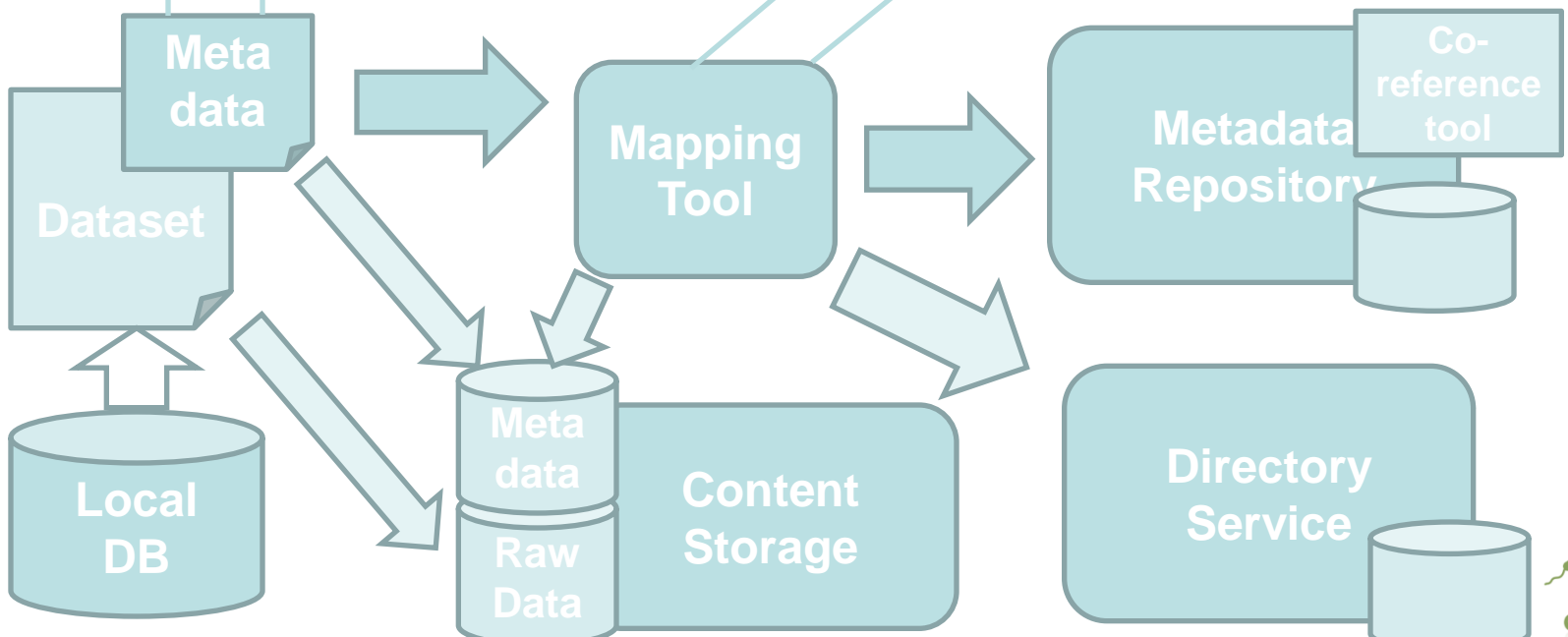
BC19 Actor

BC16 Timespan — has timespan

BC3 Place — took place at

BC53 Specimen — assigned to

BT27 Species — assigned

is about

BC21 Dataset — is composed of

BC42 Collection

is subject to — E30 Rights

rights held by

BC13 Institution

E17 Type Assignment — is carried out by

Meta data

Dataset

Local DB

Mapping Tool

Meta data

Raw Data

Content Storage

Metadata Repository

Co-reference tool

Directory Service

LifeWatch Greece

DwC and Semantic Models

- DwC can be translated without loss into Semantic Model.

- Semantic Model can be reduced to DwC.

- Complete compatibility with DwC services.

- Easy addition of features for increased precision (data tolerances, modalities etc.).

DwC drawbacks

▪ Often leaves the interpretation of the relationships between the whole record and one of its fields to the intuition of the human reader (it cannot be used to draw logical conclusion without human intervention) - No reasoning can be applied.

▪ Complex events causally related cannot be described.

▪ Fields like prey of and predator of are missing. Extensions by adding fields without **property subsumption** cause a proliferation of queries to collect related information.

▪ A well-defined ontology - a formal definition of **relationships** between concepts in this domain and between its disciplines can solve this problem without increasing complexity.

▪ Shorter queries to retrieve the same kind of information than with DwC

MarineTLO

- Inspired by ISO 21121, OBOE, INSPIRE, DwC and others. It follows the event pattern
    - Things, people and ideas meeting in space and time
    - Strong mechanism for heterogeneous complementary information

- defines the underlying semantics of database schemata used in marine domain in terms of a formal ontology.

- does not define any of the terminology typically appearing as data in the respective data structures.

- foresees the characteristic relationships for its use.

- enables semantic interoperability.

- facilitates the knowledge sharing, information exchange integration, between heterogeneous sources found distributed across various infrastructures by providing a common understanding of the concepts and entities and relationships holding between them.

Occurrence Records in DwC

```xml
    <field index="1"
term="http://rs.tdwg.org/dwc/terms/locality"/>
    <field index="2" term="http://purl.org/dc/terms/modified"/>
    <field index="3"
term="http://rs.tdwg.org/dwc/terms/identificationReferences"/>
    <field index="4"
term="http://rs.tdwg.org/dwc/terms/country"/>
    <field index="5" term="http://purl.org/dc/terms/language"/>
    <field index="6"
term="http://rs.tdwg.org/d
    <field index="7"
term="http://purl.org/dc/t
    <field index="8"
term="http://rs.tdwg.org/d
    <field index="9" term=
    <field index="10"
term="http://purl.org/dc/t
```

```xml
        <dataset>
            <title xml:lang="eng">Eastern Mediterranean Syllidae
</title>
            <creator>
            <individualName>
                <givenName>Sarah</givenName>
                <surName>Faulwetter</surName>
            </individualName>
            <organizationName>Hellenic Centre for Marine Research
</organizationName>
                <positionName>Dat
                <address>
                    <deliveryPoint>
Gournes</deliveryPoin
                    <city>Heraklion
                    <administrative
                    <postalCode>710
                    <country>GR</co
                </address>
```
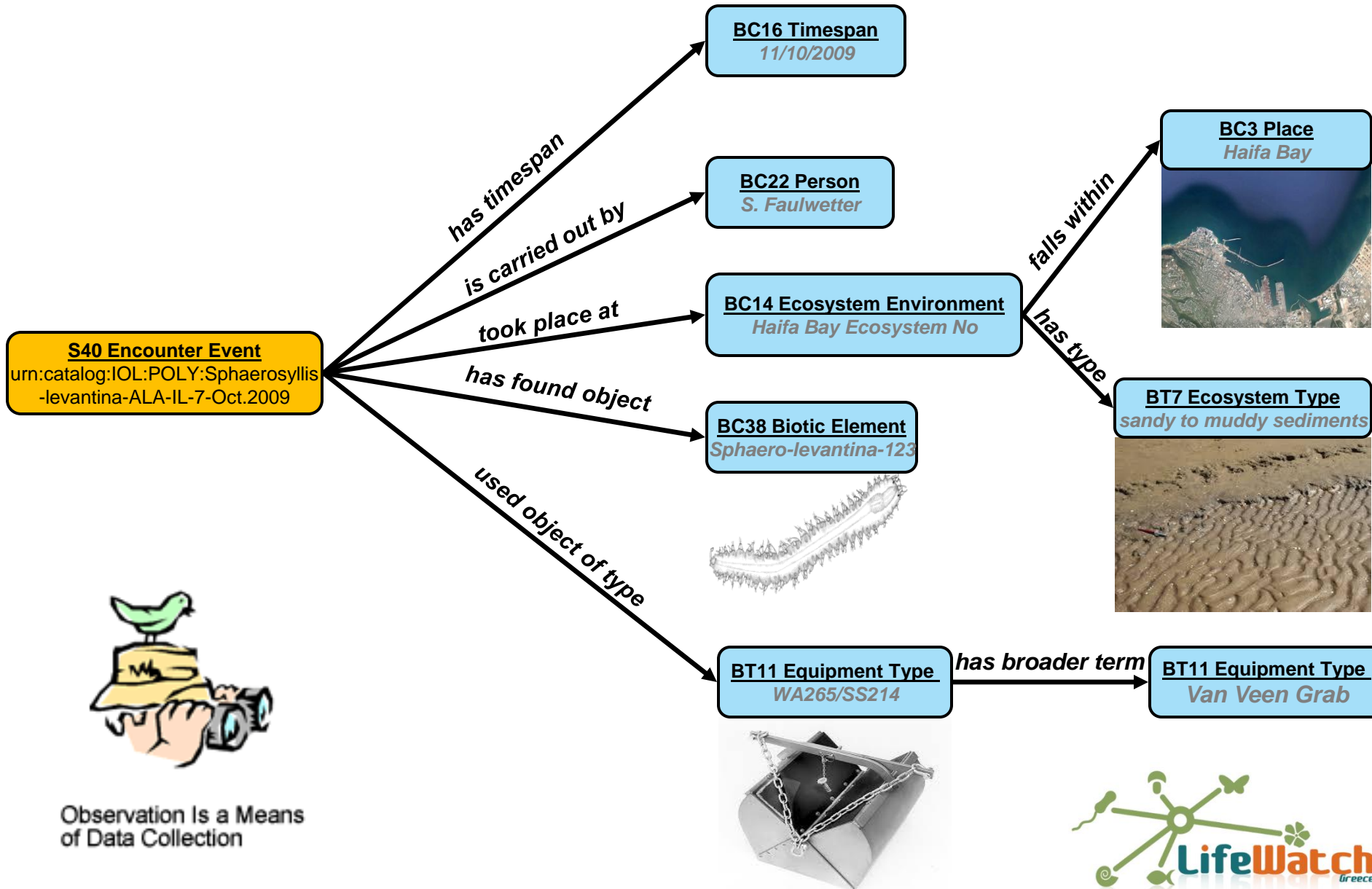
EasternMedSyllids_94 Alykes   2011-09-26 San Martῖn 2003,
2004, Nygren 2004   Greece   en   Animalia
      http://www.opendatacommons.org/licenses/by/1.0/ HCMR
      Palpata   HCMR ICZN 3   Syllidae   Whole animal,
preserved in ethanol Mediterranean Sea   Giorgos
Chatzigeorgiou, Sarah Faulwetter, Christos ArvanitidisPolychaeta
      HCMR HCMR Phyllodocida   EasternMedSyllids
      Odontosyllis fulgurans (Audouin & Milne Edwards, 1834)
      http://www.opendatacommons.org/licenses/by/1.0/ 2007-09-18
      Odontosyllis fulgurans (IMBG-NaGISA-CALB-20B_07)20
      024.9875000   NaGISA/HAIFA   35.4158333 20   fulgurans
      doi: 10.4067/S0717-65382003000200004B rocky substrate with
photophilic algae (Sargassum spp., Cystoseira spp).
      Odontosyllis   Odontosyllis_fulgurans-CALB-20B_07
      species
EasternMedSyllids_110 Alykes   2011-09-26 San Martῖn 2003,
2004, Nygren 2004   Greece   en   Animalia
      http://www.opendatacommons.org/licenses/by/1.0/ HCMR
      Palpata   HCMR ICZN 1   Syllidae   Whole animal,
preserved in ethanol Mediterranean Sea   Giorgos
Chatzigeorgiou, Sarah Faulwetter, Christos ArvanitidisPolychaeta
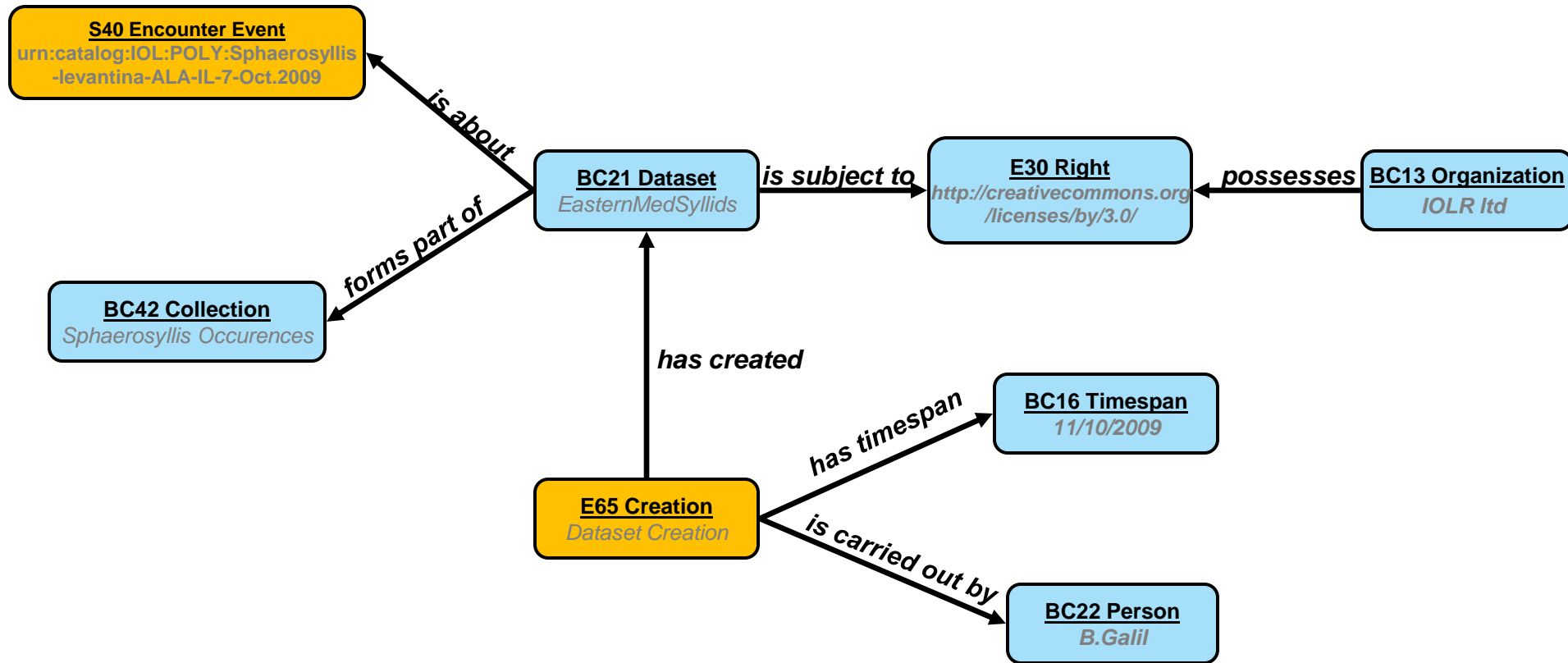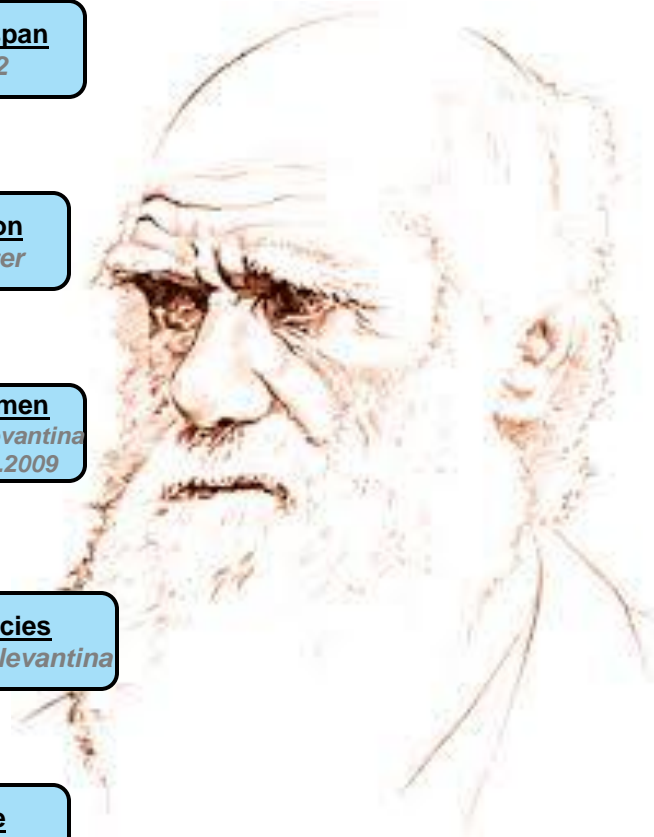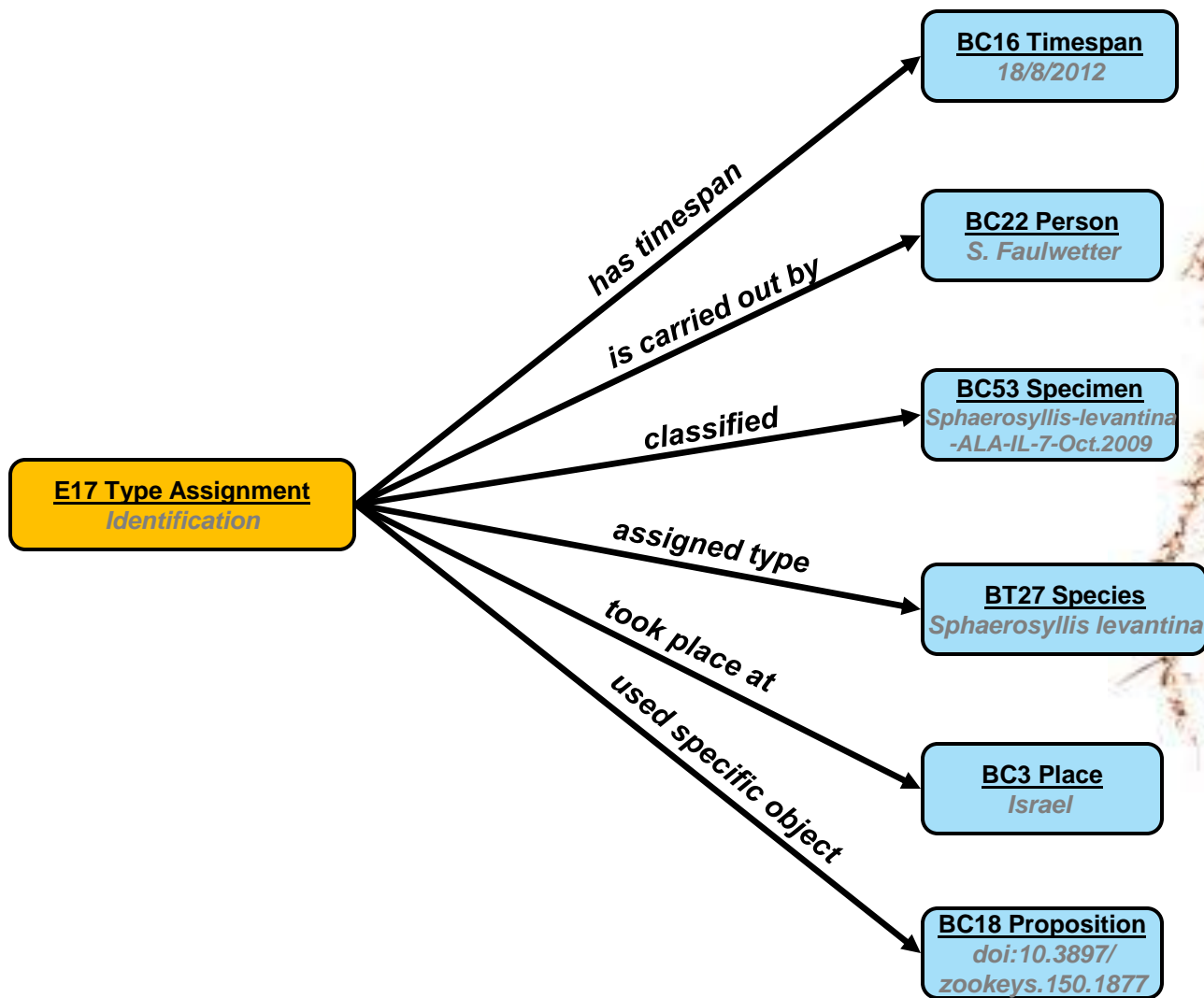
Occurrence

S40 Encounter Event
urn:catalog:IOL:POLY:Sphaerosyllis
-levantina-ALA-IL-7-Oct.2009

BC16 Timespan
11/10/2009

BC22 Person
S. Faulwetter

BC14 Ecosystem Environment
Haifa Bay Ecosystem No

BC38 Biotic Element
Sphaero-levantina-123

BC3 Place
Haifa Bay

BT7 Ecosystem Type
sandy to muddy sediments

BT11 Equipment Type
WA265/SS214

BT11 Equipment Type
Van Veen Grab

has timespan

is carried out by

took place at

has found object

used object of type

falls within

has type

has broader term

Observation Is a Means
of Data Collection

LifeWatch Greece

Occurrence (more info)

Identification

E17 Type Assignment
Identification

- has timespan → **BC16 Timespan** *18/8/2012*
- is carried out by → **BC22 Person** *S. Faulwetter*
- classified → **BC53 Specimen** *Sphaerosyllis-levantina -ALA-IL-7-Oct.2009*
- assigned type → **BT27 Species** *Sphaerosyllis levantina*
- took place at → **BC3 Place** *Israel*
- used specific object → **BC18 Proposition** *doi:10.3897/ zookeys.150.1877*

LifeWatch Greece

Measurement Records Example

EventDate	11/10/2009
verbatimEventDate	Autumn 2009
year	2009
month	10
day	11
habitat	sandy to muddy sediments
samplingProtocol	Van-Veen grab (KAHLSICO, model WA265/SS214) 32x35cm,
measurementType	total length
measurementValue	1.2
measurementUnit	mm
measurementDeterminedBy	Bella Galil
measurementDeterminedDate	11/10/2009
identifiedBy	Sarah Faulwetter
dateIdentified	18/8/2012

Measurement

BC16 Timespan
11/10/2009

BC22 Person
S. Faulwetter

BC54 Measurement
Length Measurement

BC53 Specimen
*Sphaerosyllis-levantina
-ALA-IL-7-Oct.2009*

Number
1.2

BC35 Dimension
Length

BC36 Unit
mm

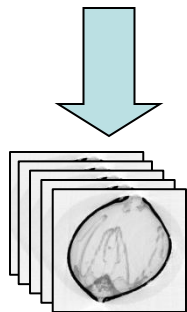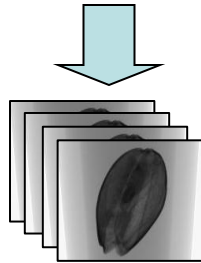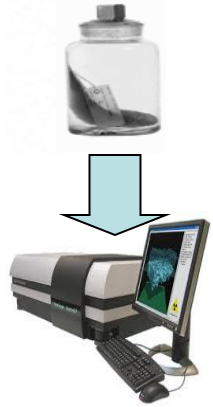has timespan

is carried out by

measured

observed dimension

has value

has unit
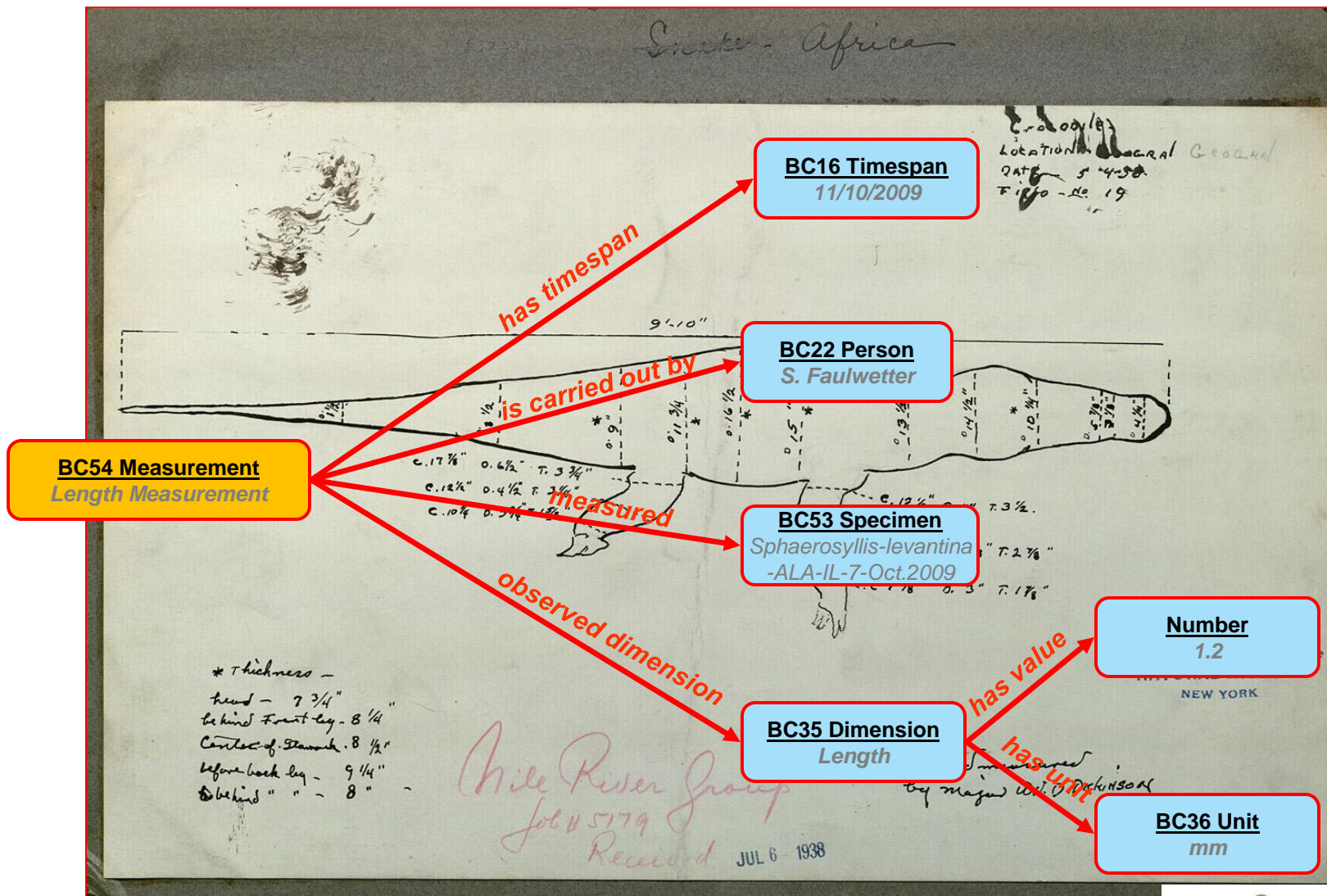
MicroCT Scanning Records

```
[System]
Scanner=Skyscan1172
Instrument S/N=08F01106
Hardware version=F
Software=Version 1. 5 (build 13)
Home directory=C:\Skyscan1172F_11MP_Hamamatsu
Source Type=Hamamatsu 100/250
Camera=Hamamatsu C9300 11Mp camera
Camera Pixel Size (um)=    8.99
CameraXYRatio=1.0000
Incl in lifting (um/mm)=0.0000
[Acquisition]
Data directory=C:\Results\Lumbrineridae
\lumbrineridae_maria_none_eth-air
Filename Prefix=lumbrineridae_
Number of Files= 1442
Source Voltage (kV)= 59
Source Current (uA)= 167
Number of Rows= 2672
Number of Columns= 4000
Image crop origin X=    0
Image crop origin Y=0
```

```
=0.2500

(Gaussian)
ction=10

FOV=OFF
ROI=ON
46
=1251
61
2677
h=4000
ive to Nyquisit frequency=100

(1)=0: Hamming (Ramp in case of optical
scanner); 1: Hann; 2: Ramp; 3: Almost Ramp;
Filter type meaning(2)=11: Cosine; 12: Shepp-Logan; [100,200]:
Generalized Hamming, alpha=(iFilter-100)/100
Undersampling factor=2
Threshold for defect pixel mask (%)=0
Beam Hardening Correction (%)=20
CS Static Rotation (deg)=0.0
Minimum for CS to Image Conversion=0.0000
Maximum for CS to Image Conversion=0.3052
```
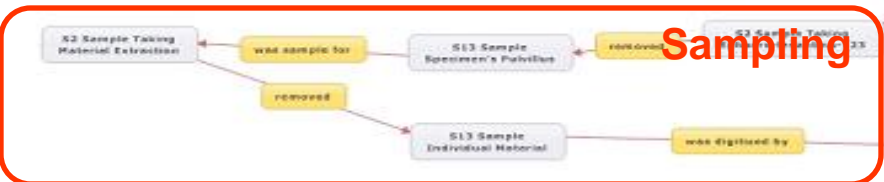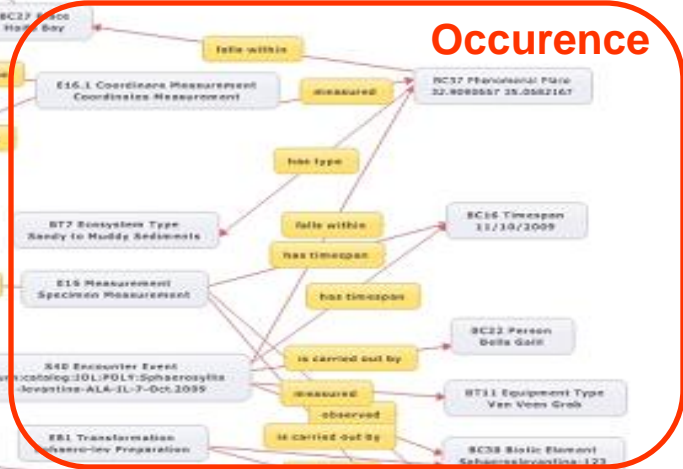
LifeWatch
Greece

Micro CT Scanning

The Graph

Thank You!!!