

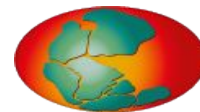
First International Workshop on Reproducible Open Science  
Hannover, Germany, September 9, 2016

# From Data to Machine Readable Information Aggregated in Research Objects

Markus Stocker

PANGAEA / MARUM  
University of Bremen  
Germany

<http://orcid.org/0000-0001-5492-3212>  
@envinf



**marum**  
Center for Marine  
Environmental Sciences



 **Universität Bremen**



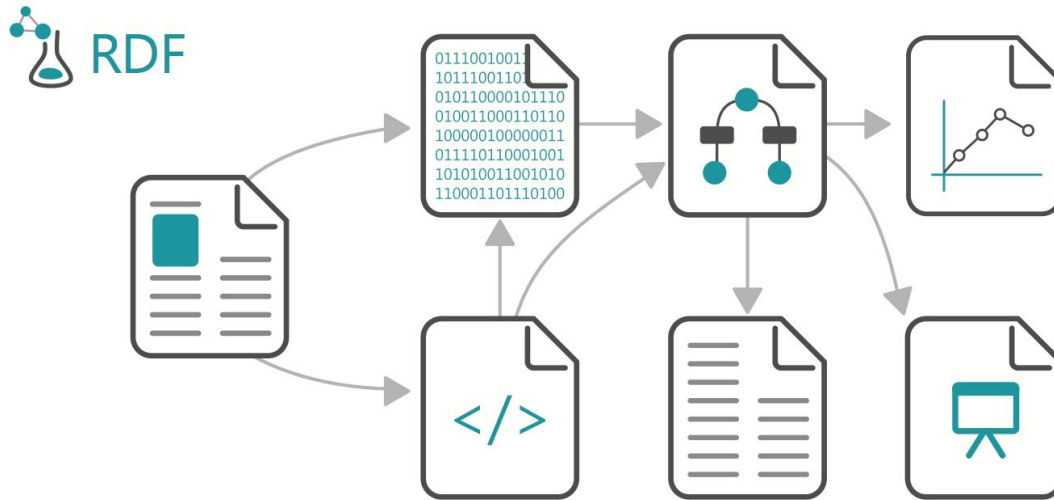
**PLUS**  
ENVRI

# Introduction

- Data interpretation is key in scientific investigations
- Process with data as input and information as output
- Data are uninterpreted symbols, e.g. sensor observation values
- Information are interpreted data, for their meaning in a real-world context
- Record information resulting in data interpretation

# Research Object

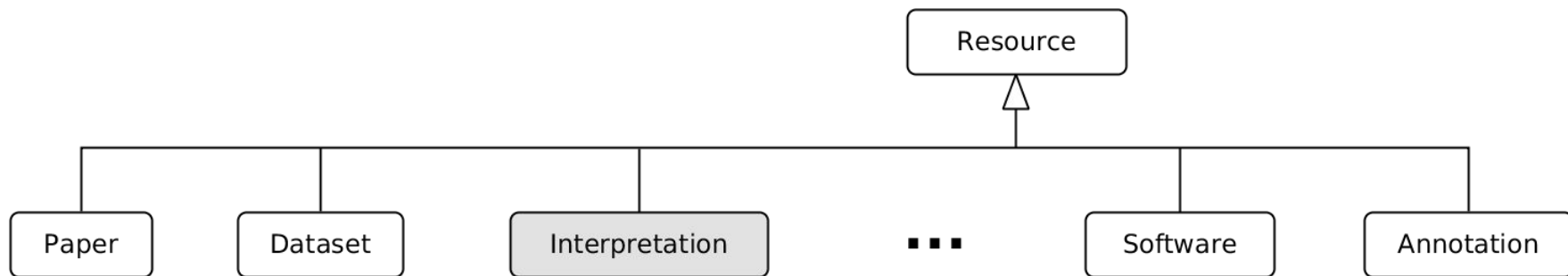
- “Semantically rich aggregations of resources that bring together data, methods and people in scientific investigations” (Bechhofer et al., 2013)



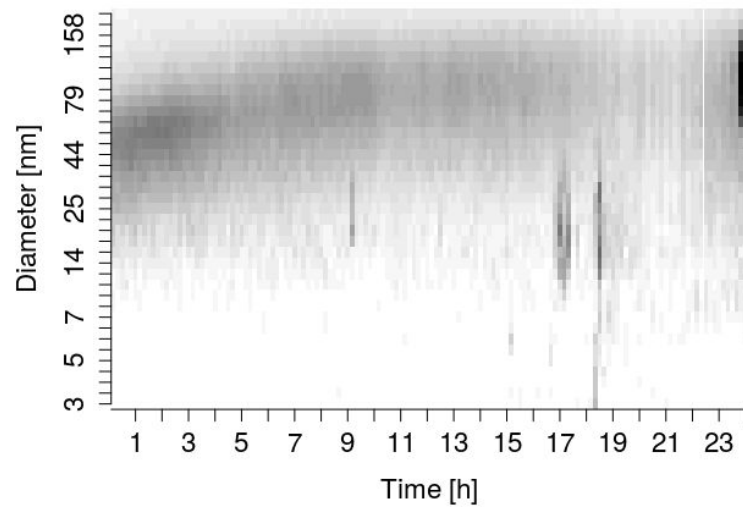
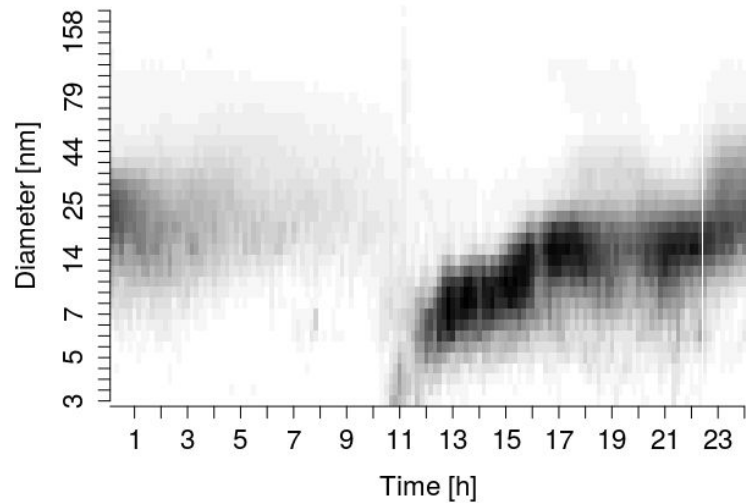
# Proposal

- Extend the Research Object model
- Additional Resource type called *Interpretation*
- Instances represent information resulting from data interpretation
- Instances are machine readable

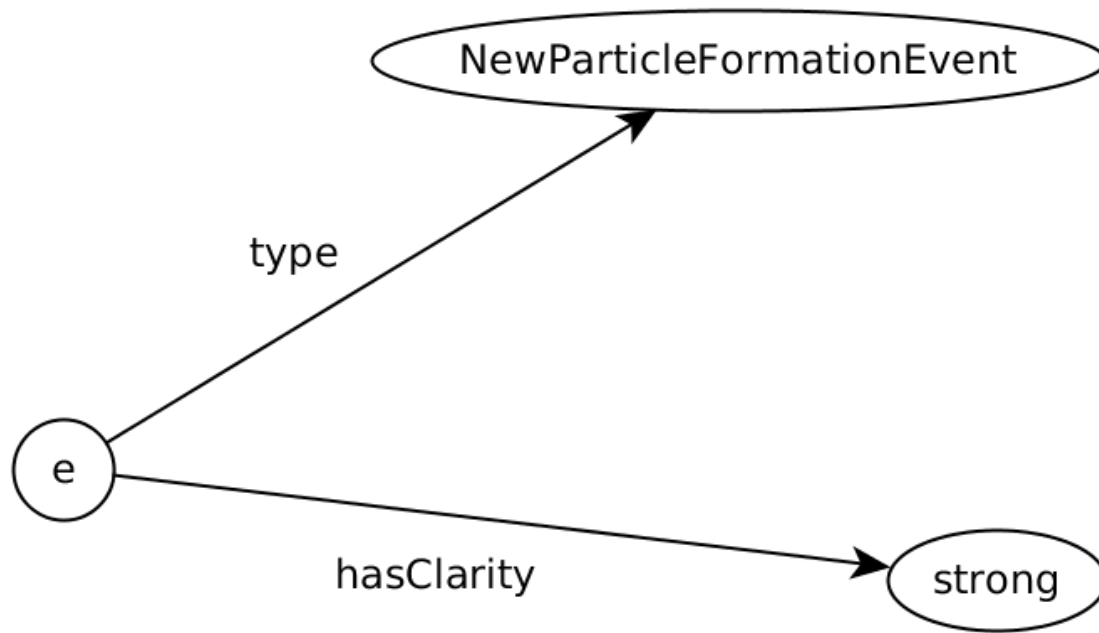
# Proposed Extension



# Application



# Application

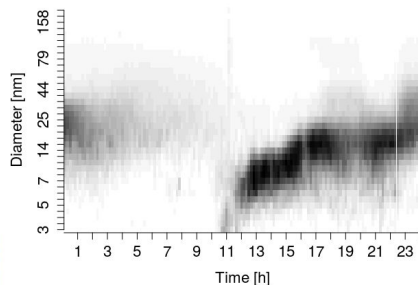


# Application

```

@prefix qb: <http://purl.org/linked-data/cube#>.
@prefix ex: <http://example.org#>.
@prefix dimension: <http://purl.org/linked-data/sdmx/2009/dimension#>.

ex:c819b0f-24b5-449a-8d66-54cbb17afe7
  a
    qb:Observation ;
  ex:d10b2 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d10b3 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d12a2 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d12a3 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d12e2 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d12e3 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d14e2 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d14e3 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d15e2 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d15e3 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d17b2 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d17e3 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d20e2 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d20b3 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d20a2 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d25e2 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d25e3 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d31e2 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d35e1 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d35e2 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d39e1 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d39e2 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d447e1 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d447e2 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d50e1 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d50e2 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d52e1 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d52e2 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d63e1 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d63e2 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d70e2 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d74e1 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d794e1 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d891e1 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d91e1 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d91e2 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d93e1 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d93e2 <http://www.w3.org/2001/XMLSchema#double>;
  ex:d1 <http://www.w3.org/2001/XMLSchema#dateTime>.
dimension:timePeriod
  <2011-03-26T00:00:00+00:00> <http://www.w3.org/2001/XMLSchema#dateTime>.
  
```



```

@prefix ex: <http://example.org#>.
@prefix dimension: <http://purl.org/linked-data/sdmx/2009/dimension#>.
@prefix qb: <http://purl.org/linked-data/cube#>.

ex:5df2e06e-5288-4cf9-a04e-6e8c204646b1
  a
    qb:Observation ;
  ex:p01 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p02 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p03 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p04 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p05 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p06 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p07 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p08 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p09 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p10 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p11 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p12 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p13 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p14 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p15 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p16 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p17 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p18 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p19 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p20 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p21 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p22 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p23 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p24 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p25 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p26 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p27 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p28 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p29 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p30 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p31 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p32 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p33 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p34 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p35 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p36 <http://www.w3.org/2001/XMLSchema#double>;
  ex:p37 <http://www.w3.org/2001/XMLSchema#double>;
  qb:dataSet
    ex:d2 ;
  dimension:timePeriod
    <2011-03-26+00:00> <http://www.w3.org/2001/XMLSchema#date>.
  
```

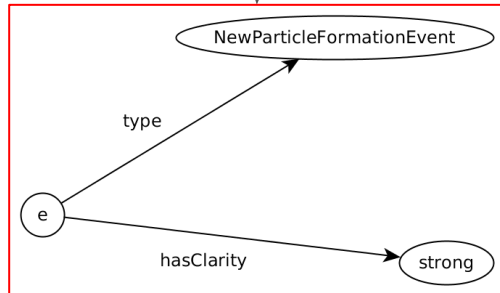
```

# -----
# Author: Markus Stocker
# ORCID: 0000-0001-5492-3212
# Date: 2016-08-11
# License: CC-BY
# -----
library(Rcur1)
# https://github.com/egonw/rrdf
# Requires rJava, see:
# http://stackoverflow.com/a/2593282
library(rrdf)
library(uuid)

# ---- CONFIGURATION ----
Sys.setlocale("LC_TIME", "en_US.UTF-8")

host <- "http://avaa.tdata.fi"
table <- "HYV_DNPS"
from <- "2011-03-26 00:00:00.000"
to <- "2011-03-27 00:00:00.000"
quality <- "ANY"
averaging <- "NONE"
type <- "NONE"
variables <- c("d31e1","d35e1","d39e1","d44e1","d50e1","d56e1",
"d63e1","d70e1","d794e1","d891e1","d10oe2","d11ze2",
"d12e2","d14e2","d15e2","d17e2","d20oe2","d224e2",

```





# Application

```
<>      a          ro:ResearchObject ;
         ore:aggregates ex:d1, ex:f1, ex:d2, ex:s1, ex:e ;
         dct:created   "2016-08-11"^^xsd:dateTime ;
         dct:creator   [ a foaf:Person; foaf:name "Markus Stocker" ] .

ex:d1    a          wf4ever:Dataset, qb:DataSet ;
         swrc:doi    <https://doi.org/10.6084/m9.figshare.3565635> .

ex:f1    a          wf4ever:Image ;
         swrc:doi    <https://doi.org/10.6084/m9.figshare.3567273> .

ex:d2    a          wf4ever:Dataset, qb:DataSet ;
         swrc:doi    <https://doi.org/10.6084/m9.figshare.3571146> .

ex:s1    a          wf4ever:Software ;
         swrc:doi    <https://doi.org/10.6084/m9.figshare.3571212> .

ex:e     a          ex:NewParticleFormationEvent, ro:Interpretation ;
         ex:hasClarity ex:strong .
```

# Discussion

- Relevance of ontology because interpretation follows a conceptualization
- Share semantics between humans and machines
- Utilize interpretations to build models, e.g. machine learning classifiers
- Applicable also to Distributed Scholarly Compound Object (DiSCO)
- Link other PID types, e.g. ORCID iD

```
<> a ro:ResearchObject ;  
    dct:creator [ a foaf:Person; dbo:orcidId "0000-0001-5492-3212" ] .
```

# Conclusion

- Data interpretations are artefacts in scientific investigations
- Record interpretations in artefact aggregations, e.g. Research Object
- Record for humans *and* machines, not just images and natural language text