Introduction to IEO

What is needed, and why

Straw man first spec (necessary)

- An information entity is a generically dependent continuant that originates with a person, or a machine that was designed to have a function to produce/communicate information.
  - Tie to the originator, versus the receiver
  - Allow for machines
  - Don't allow for information in everything

Goal

- Data integration at Web scale
- A hierarchy with an easily answerable decision at each branch point. So that people who wish to use the ontology can do so effectively and so that representation choices are consistently made the same way by different people.
- Account for decorated content – text and formatting. Text and subtext. Multiplexing.

Three aspects of representing scientific knowledge

- *Record level*: Represent database records. Inconsistent if two sources disagree about contents of a field.
- *Statement level*: Represent what researchers say. Inconsistent if two people disagree about what a paper said
- *Domain level*: OBO Foundry approach. Represent your best understanding of consensus. Inconsistent if facts contradict.
- We need all three (but make clear which is which)

What's in scope (entities)?

- Measurements
- Licenses
- Writings and other narrative objects
- Data sets
- Files and formats
- Software
- Plans, Goals
- Histories

What's in scope(relations)?

- Epistemic relations – support/contradiction
- Encoding relations
- Visualization relations – between data and plot
  - ("A mass spectrum is an intensity vs m/z (mass-to -charge> ratio) plot representing a chemical analysis")
- Aboutness

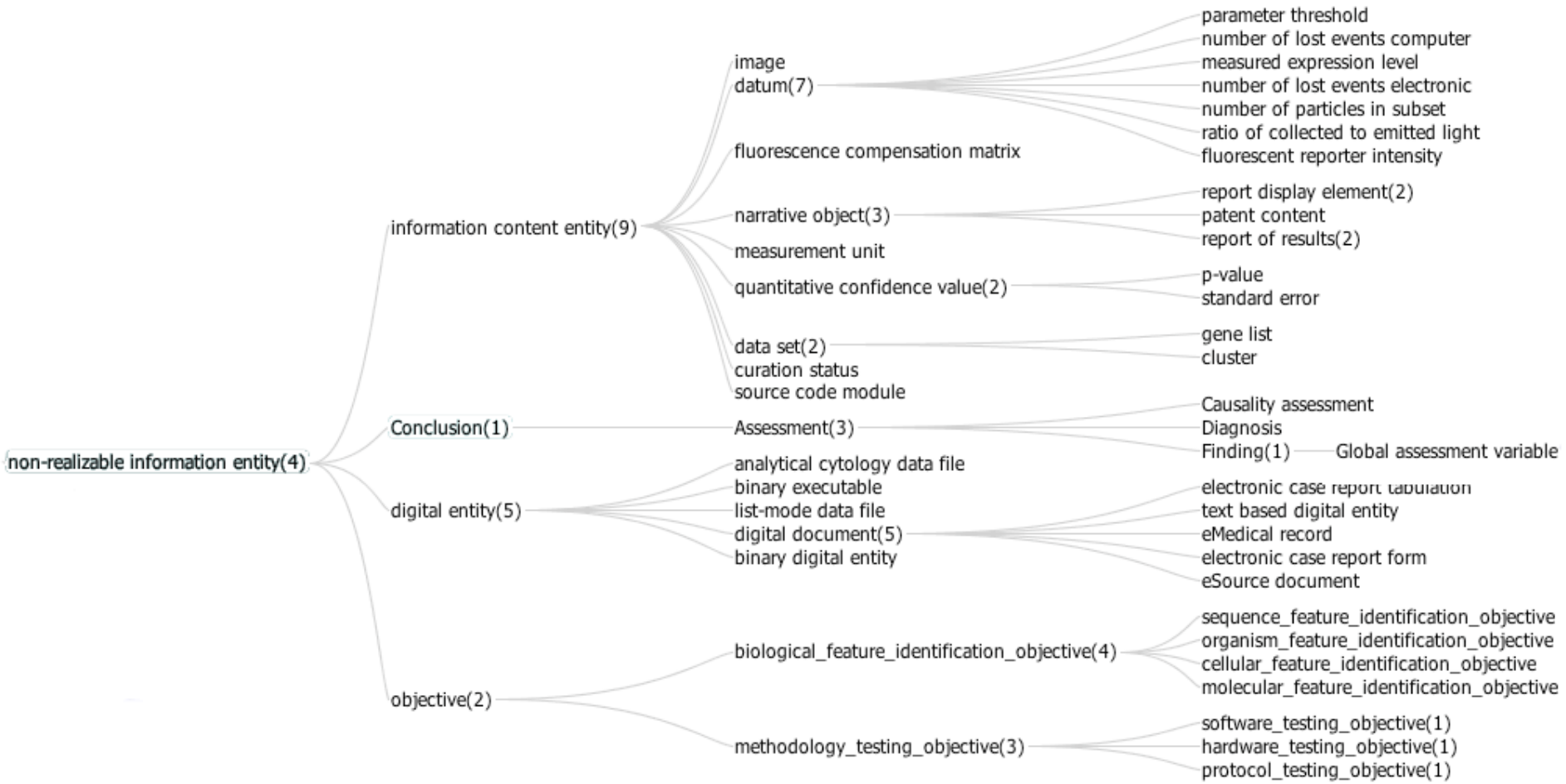
Information and realism

- Realism - a practice of representation that traces entities one represents back to particulars in the physical world.
- Realism is a pragmatic approach for gaining consensus in scientific ontologies.
- Challenge: To what do we trace “information” – what bears it? What is it about?
-

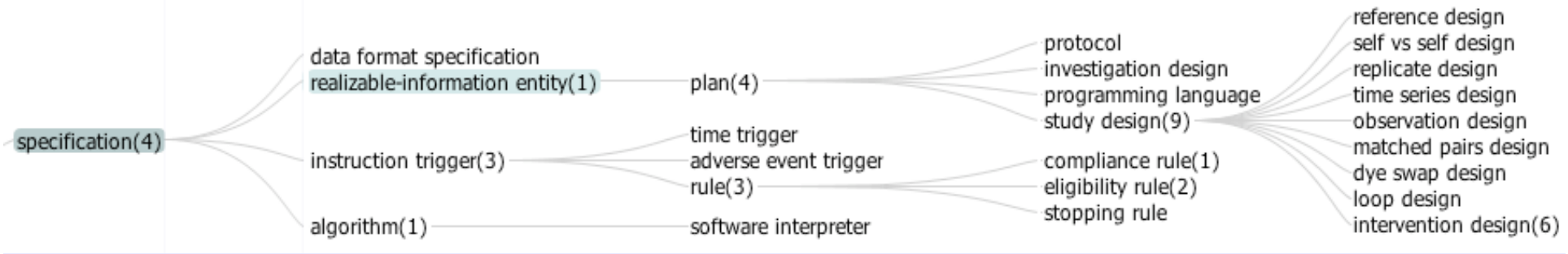
Persistent problems

- Confusion between database records and what they are about.
- Confusion between sequence information and biological polymers
- Worries about including "propositions" in the ontology – self reflection
- Plain lack of understanding of where law goes in this framework
- What entities are individuals versus classes
- "Information Resources" on the Semantic Web
- Change over time
- "Qualities" of information – e.g. compressibility
- Constant reinvention of bibliographic ontologies
- How reliable is this information. Do I believe this (Tom)

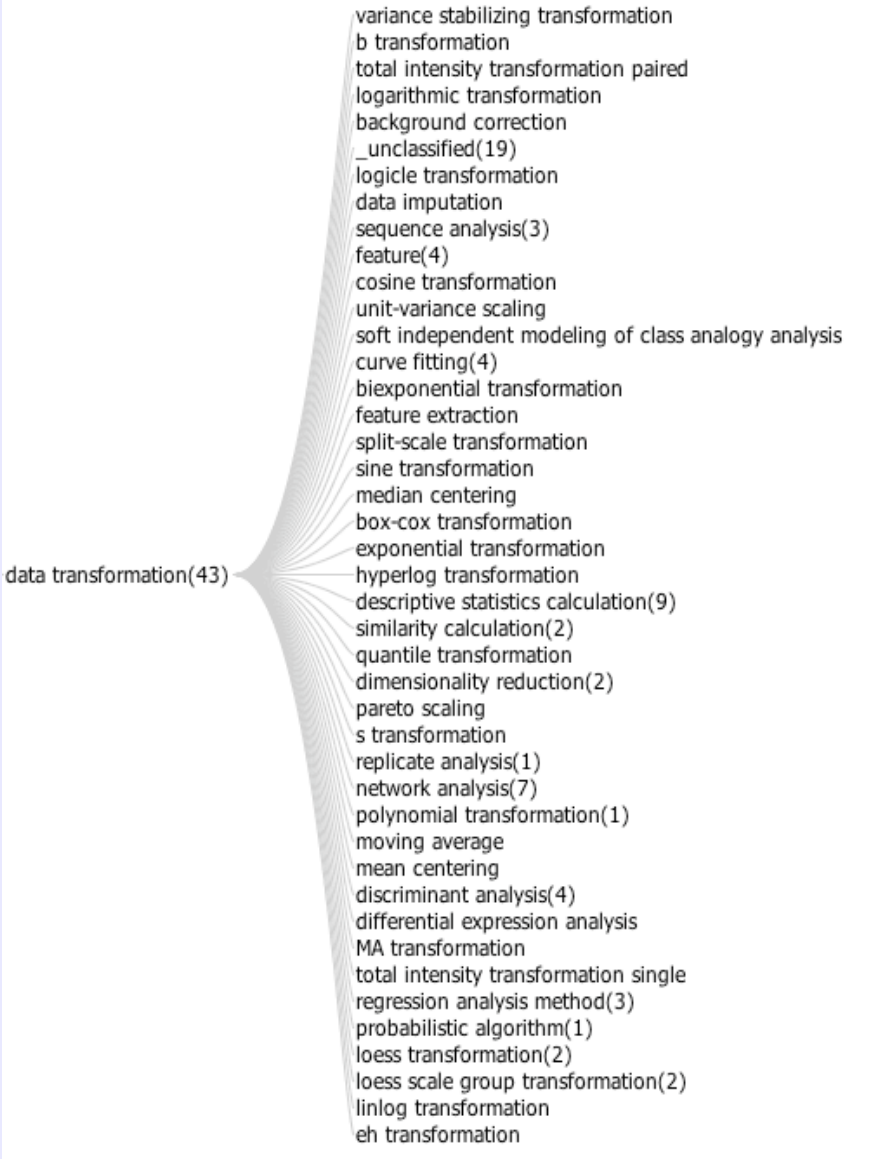
OBI


non-realizable information entity(4)
information content entity(9)
image
datum(7)
parameter threshold
number of lost events computer
measured expression level
number of lost events electronic
number of particles in subset
ratio of collected to emitted light
fluorescent reporter intensity
fluorescence compensation matrix
narrative object(3)
report display element(2)
patent content
report of results(2)
measurement unit
quantitative confidence value(2)
p-value
standard error
data set(2)
gene list
cluster
curation status
source code module
Conclusion(1)
Assessment(3)
Causality assessment
Diagnosis
Finding(1)
Global assessment variable
digital entity(5)
analytical cytology data file
binary executable
list-mode data file
digital document(5)
electronic case report tabulation
text based digital entity
eMedical record
electronic case report form
eSource document
binary digital entity
objective(2)
biological_feature_identification_objective(4)
sequence_feature_identification_objective
organism_feature_identification_objective
cellular_feature_identification_objective
molecular_feature_identification_objective
methodology_testing_objective(3)
software_testing_objective(1)
hardware_testing_objective(1)
protocol_testing_objective(1)

OBI


specification(4)
data format specification
realizable-information entity(1)
instruction trigger(3)
algorithm(1)
plan(4)
time trigger
adverse event trigger
rule(3)
software interpreter
protocol
investigation design
programming language
study design(9)
compliance rule(1)
eligibility rule(2)
stopping rule
reference design
self vs self design
replicate design
time series design
observation design
matched pairs design
dye swap design
loop design
intervention design(6)

OBI


data transformation(43)
variance stabilizing transformation
b transformation
total intensity transformation paired
logarithmic transformation
background correction
_unclassified(19)
logicle transformation
data imputation
sequence analysis(3)
feature(4)
cosine transformation
unit-variance scaling
soft independent modeling of class analogy analysis
curve fitting(4)
biexponential transformation
feature extraction
split-scale transformation
sine transformation
median centering
box-cox transformation
exponential transformation
hyperlog transformation
descriptive statistics calculation(9)
similarity calculation(2)
quantile transformation
dimensionality reduction(2)
pareto scaling
s transformation
replicate analysis(1)
network analysis(7)
polynomial transformation(1)
moving average
mean centering
discriminant analysis(4)
differential expression analysis
MA transformation
total intensity transformation single
regression analysis method(3)
probabilistic algorithm(1)
loess transformation(2)
loess scale group transformation(2)
linlog transformation
eh transformation

Alternatives/Inspirations

- LOA Information Object
- Software resource ontology
- FRBR
  - Work, Expression, Manifestation, Item
- ACM Computing Classification System