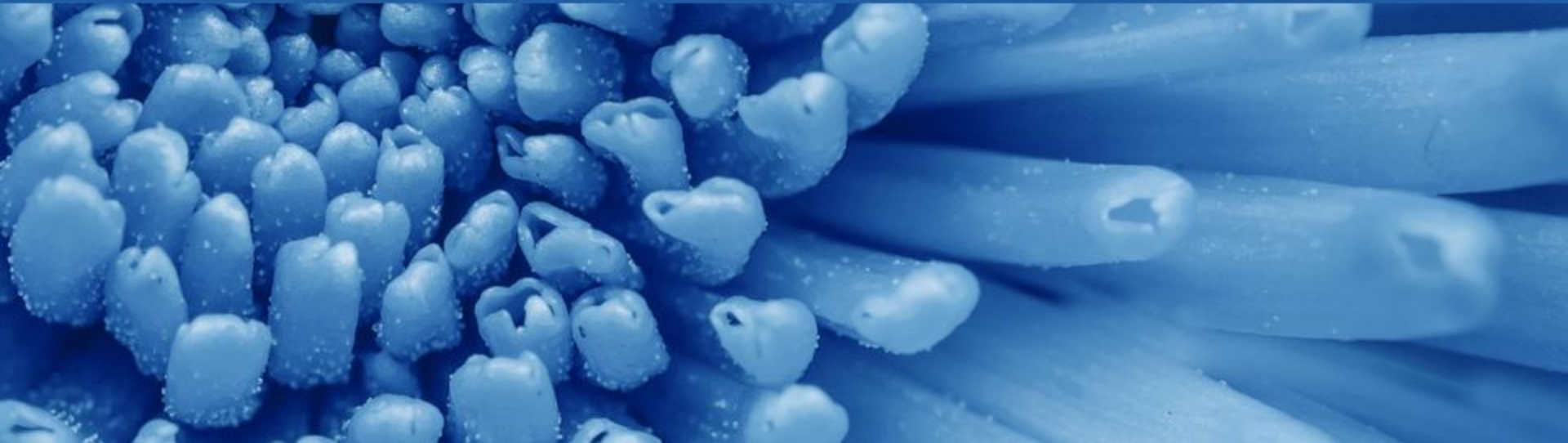
Ontologies for Material Sciences

Prof. Dr. Harald Sack
BMBF Technologiegespräch: "MaterialDigital"
Werkstoffwoche Dresden
Dresden, 18. September 2019

“It is a capital mistake to theorize before one has data.”

Arthur Conan Doyle, A Scandal in Bohemia (1892)



33.6

33.6 m

33.6m

Baleanoptera musculus — is a — Animal

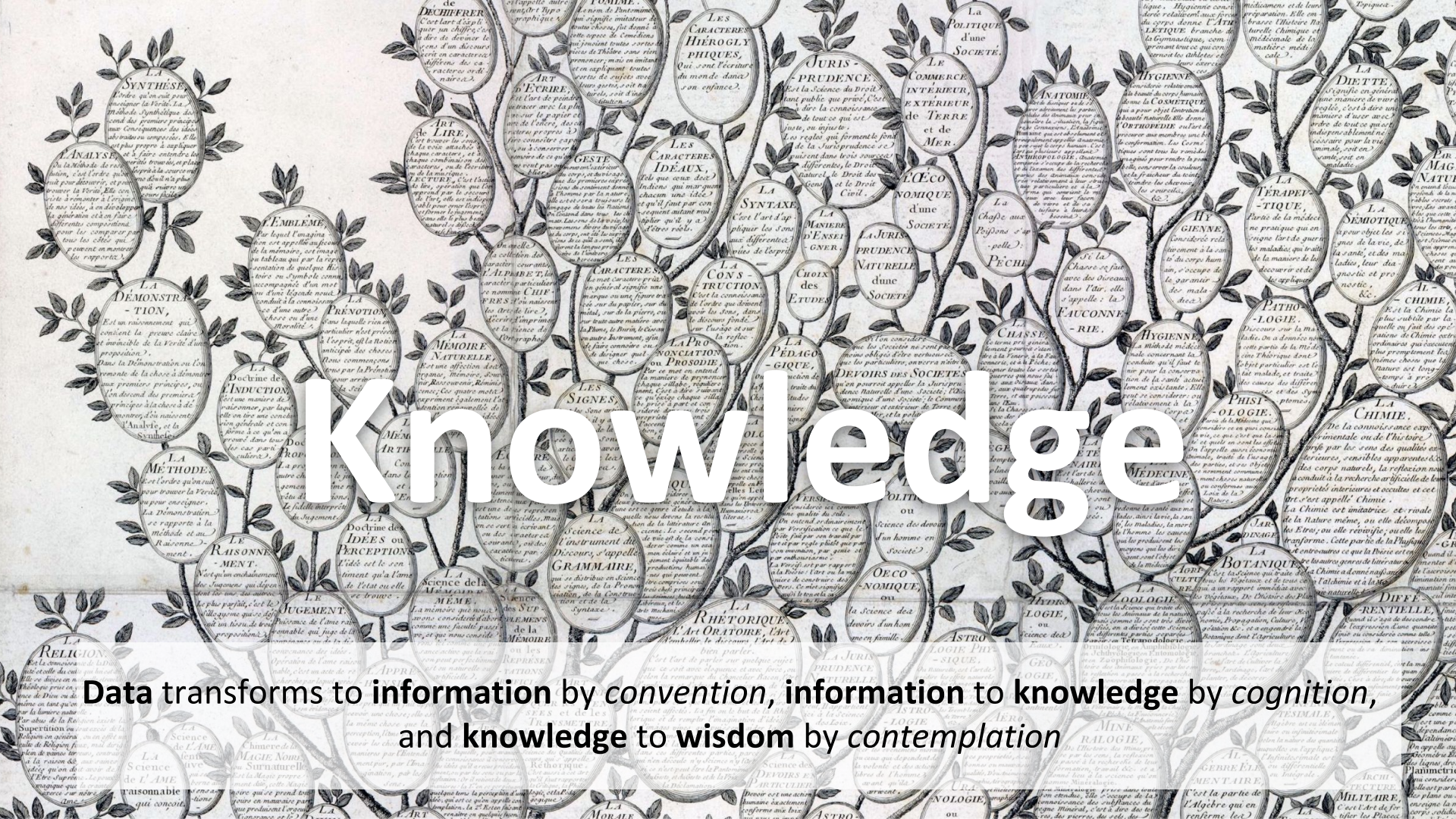
Baleanoptera musculus — maxLength → 33.6 m

$$\texttt{BaleanopteraMusculus} \sqsubseteq \texttt{Animal} \sqcap \forall \texttt{maxLength}.\leq 33.6$$

KIT Karlsruher Institut für Technologie
FIZ Karlsruhe
Leibniz-Institut für Informationsinfrastruktur

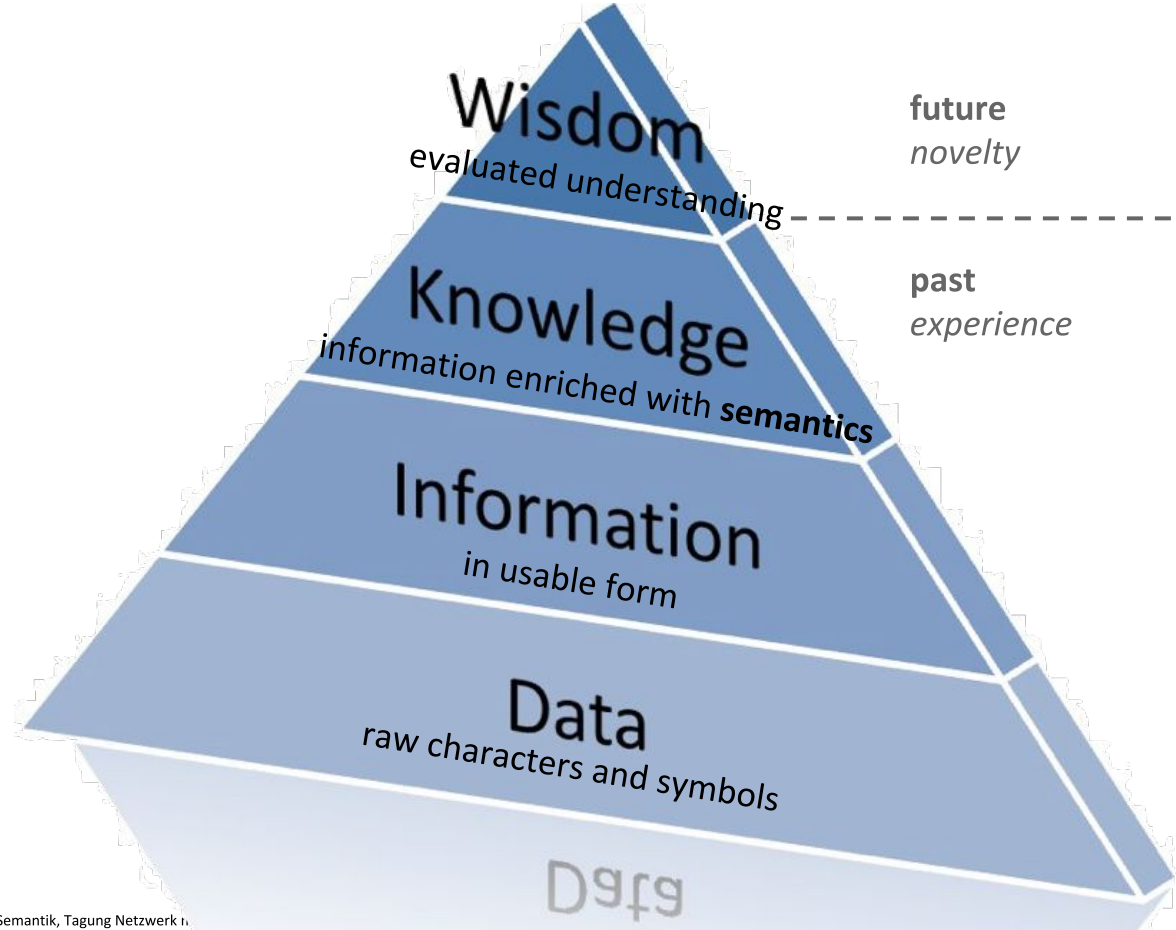

Data


Information



Data transforms to **information** by *convention*, **information** to **knowledge** by *cognition*, and **knowledge** to **wisdom** by *contemplation*

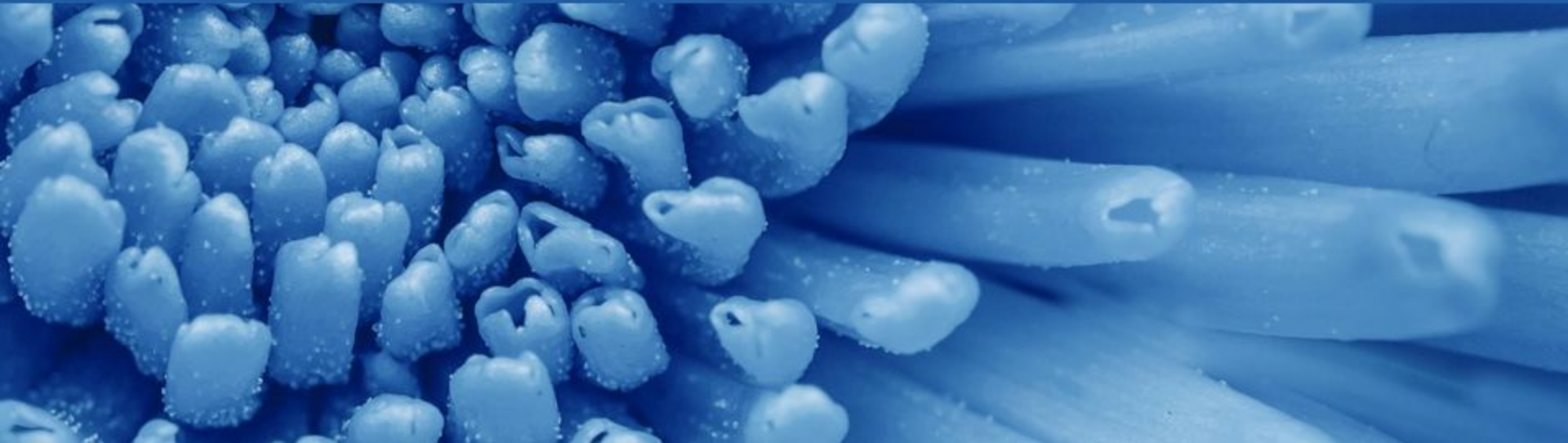
From Data to Knowledge



DIKW Pyramid, *Ackoff 1989 [1]*

“It does not do to leave a live dragon out of your calculations, if you live near him.”

J.R.R. Tolkien, The Hobbit or There and Back again (1937)



Traditional vs Semantic Technology

Baleanoptera musculus

maxLength:	33.6
avgLength:	26.0
maxWeight:	173
minPopulation:	10.000
maxPopulation:	20.000
genus:	Balaenoptera
family:	Balaenopteridae
class:	Mammalia
phylum:	Chordata

Unit of measurement required

maxLength:	**the maximum length given in metres.**
avgLength:	**the average length given in metres.**
maxWeight:	**maximum weight given in tons**
genus:	**A genus is a taxonomic rank used in the biological classification of living and fossil organisms, as well as viruses**
…	…

Natural Language

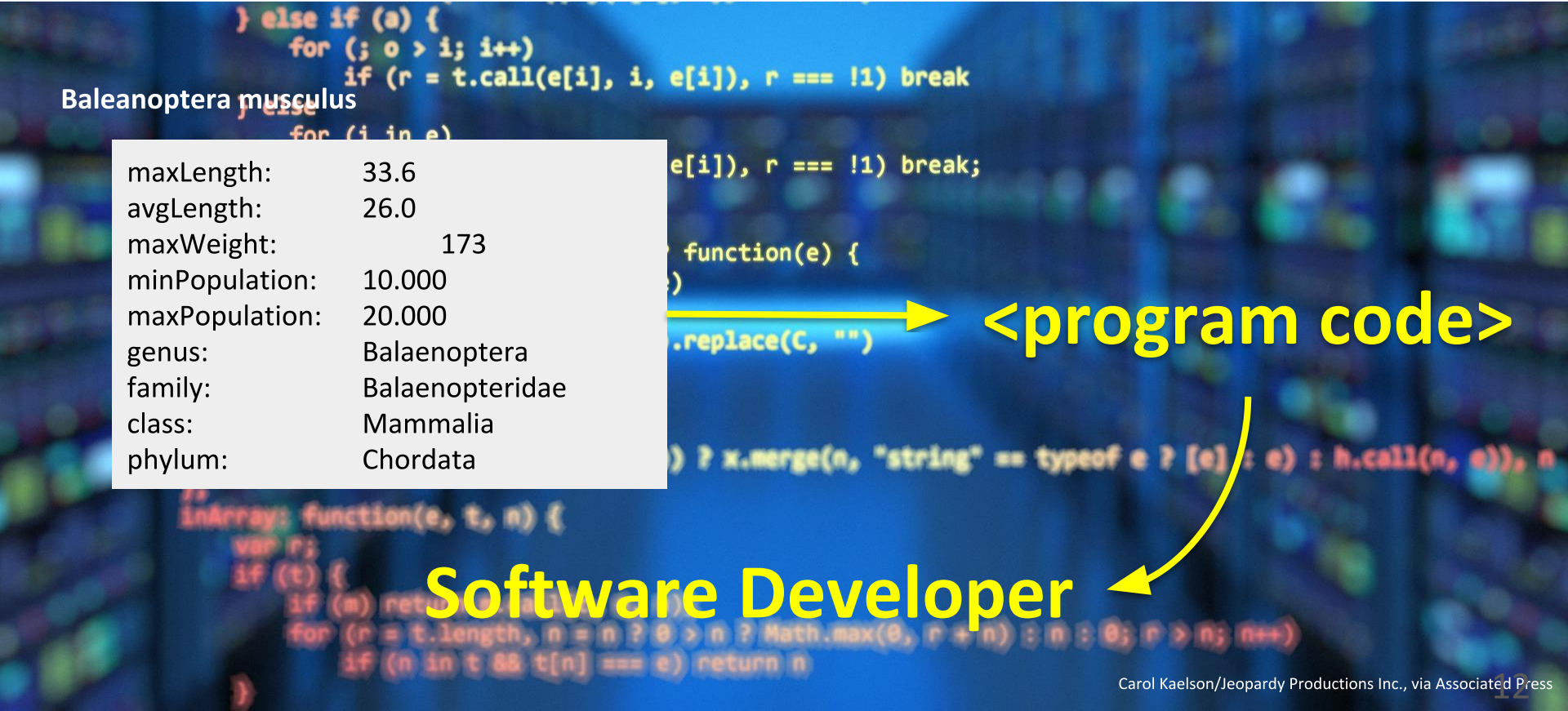
Traditional vs Semantic Technology

Baleanoptera musculus

maxLength:	33.6
avgLength:	26.0
maxWeight:	173
minPopulation:	10.000
maxPopulation:	20.000
genus:	Balaenoptera
family:	Balaenopteridae
class:	Mammalia
phylum:	Chordata

<program code>

Software Developer



Carol Kaelson/Jeopardy Productions Inc., via Associated Press

Reading vs Understanding

This sentence no verb.

This sentence rides a bicycle.

Syntax

Semantics

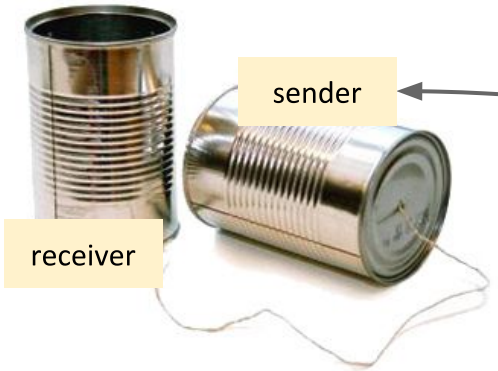
Determines rules according to which correct (**well formed**) sentences are constructed.

Determines how the **meaning** of sentences can be constructed from the meaning of smaller language units (words).

Reading (parsing) checks only syntactic rules to find out whether the text is well formed.

Understanding = correct interpretation

Understanding = Successful Communication



sender

receiver

Experience

Concept

symbolizes

refers to

Context

Experience

Symbol

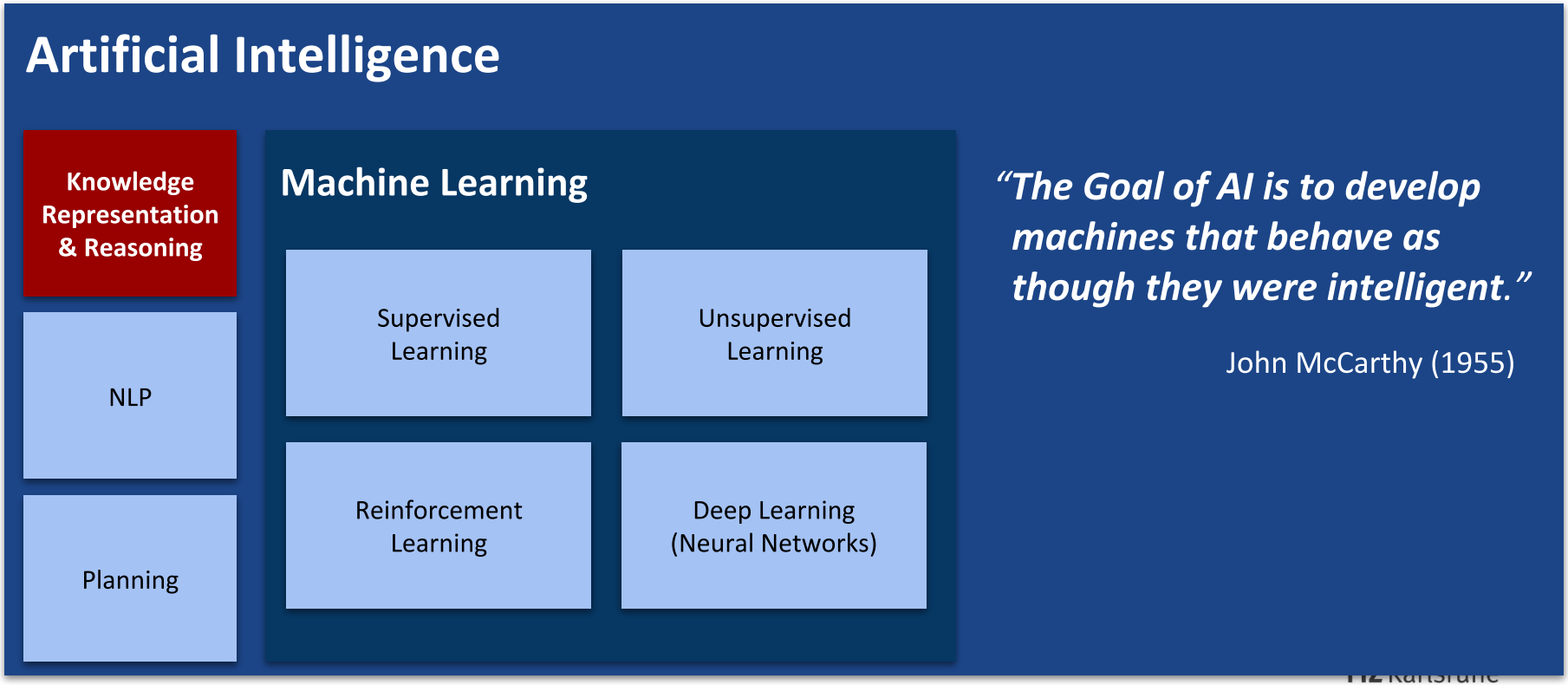
Object

stands for

"Jaguar"

Pragmatics

Symbolic Knowledge Representation



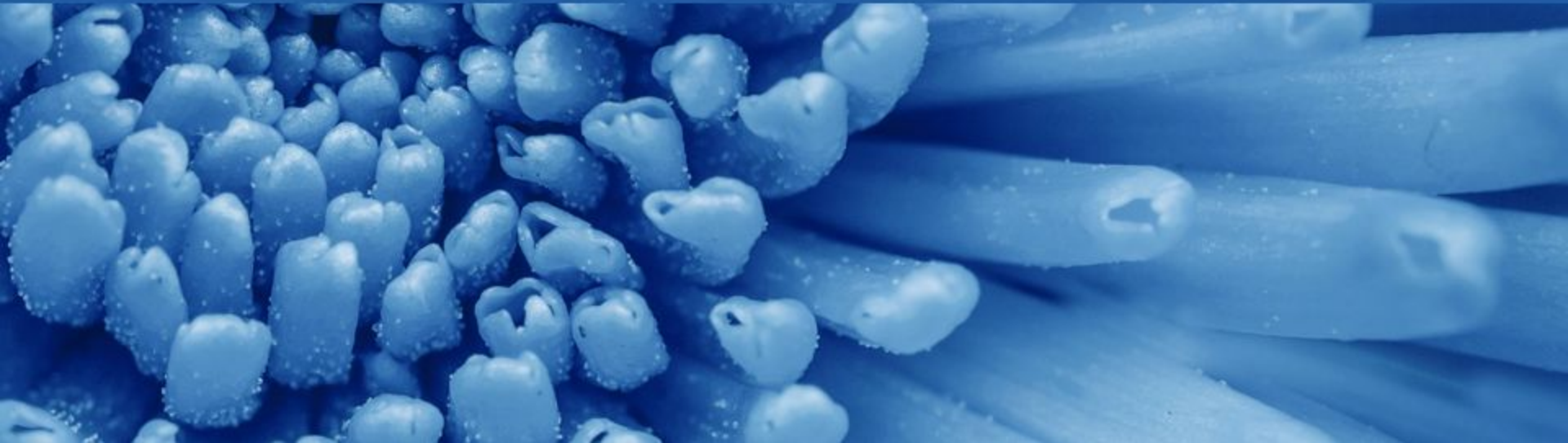
Prof. Dr. Harald Sack: Ein Skateboard für den Papst - Warum es maschinelles Lernen so schwer hat ohne Semantik, Tagung Netzwerk machinelle Verfahren in der Erschließung, DNB Frankfurt, 11. Oktober 2019

15


PARENTAL
ADVISORY
EXPLICIT SEMANTICS

“90% of most Magic
merely consists of knowing one extra fact.”

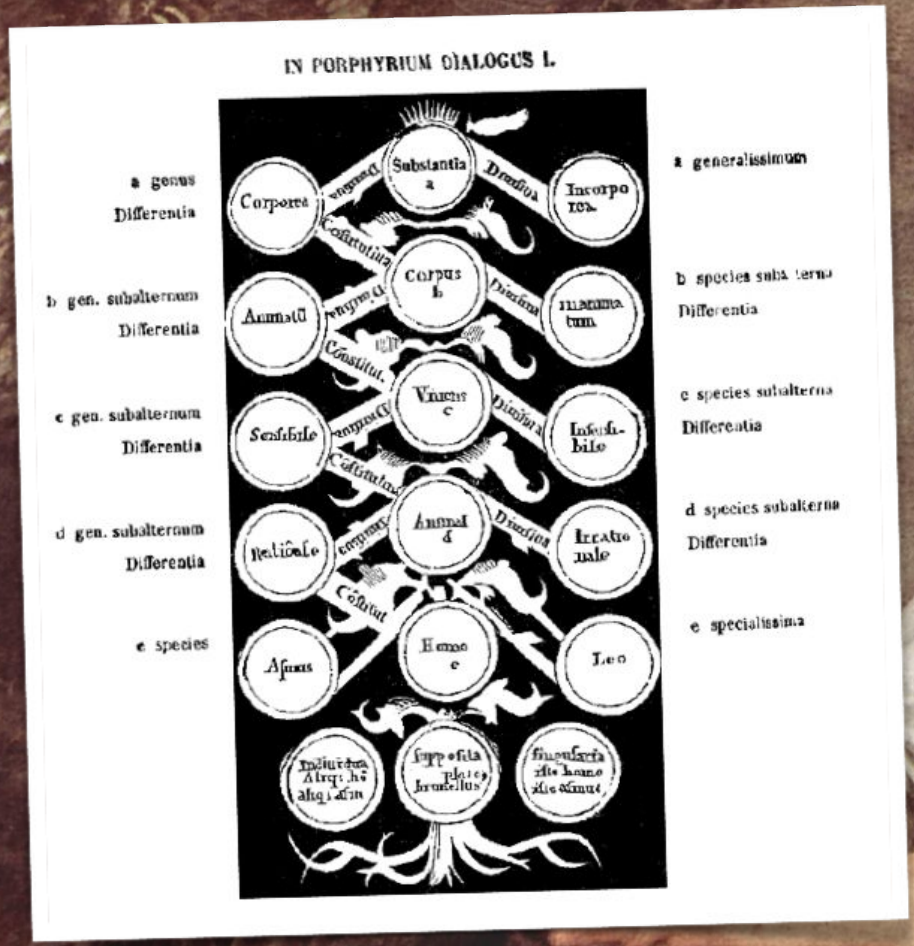
Terry Pratchett, Night Watch (2002)



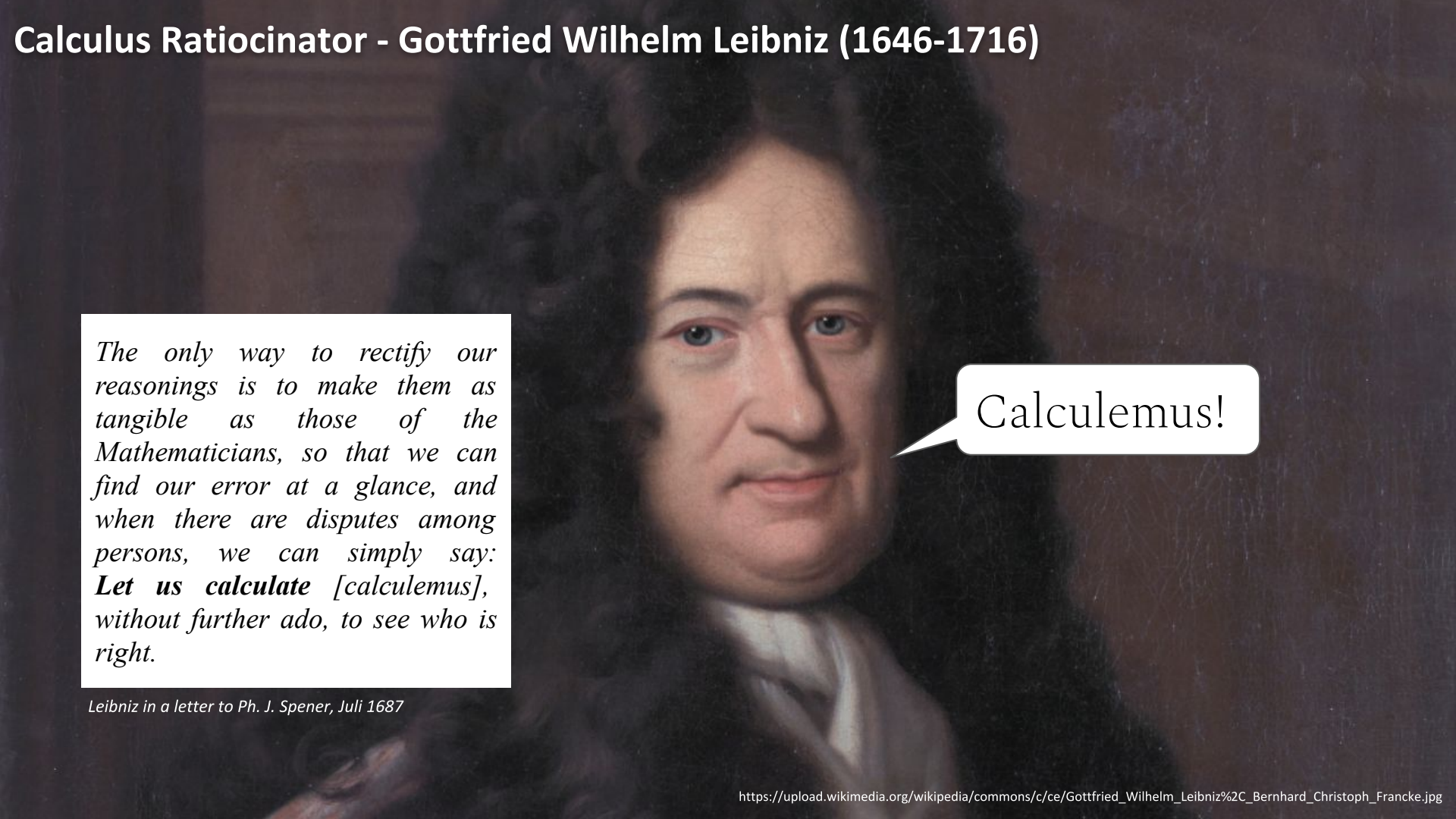
Knowledge Representation

<u>Ontology</u> is the philosophical study of the nature of being, existence, or reality, as well as the basic categories of being and their relations...

The Universal Categories - Aristotle (384–322 BCE)



Calculus Ratiocinator - Gottfried Wilhelm Leibniz (1646-1716)

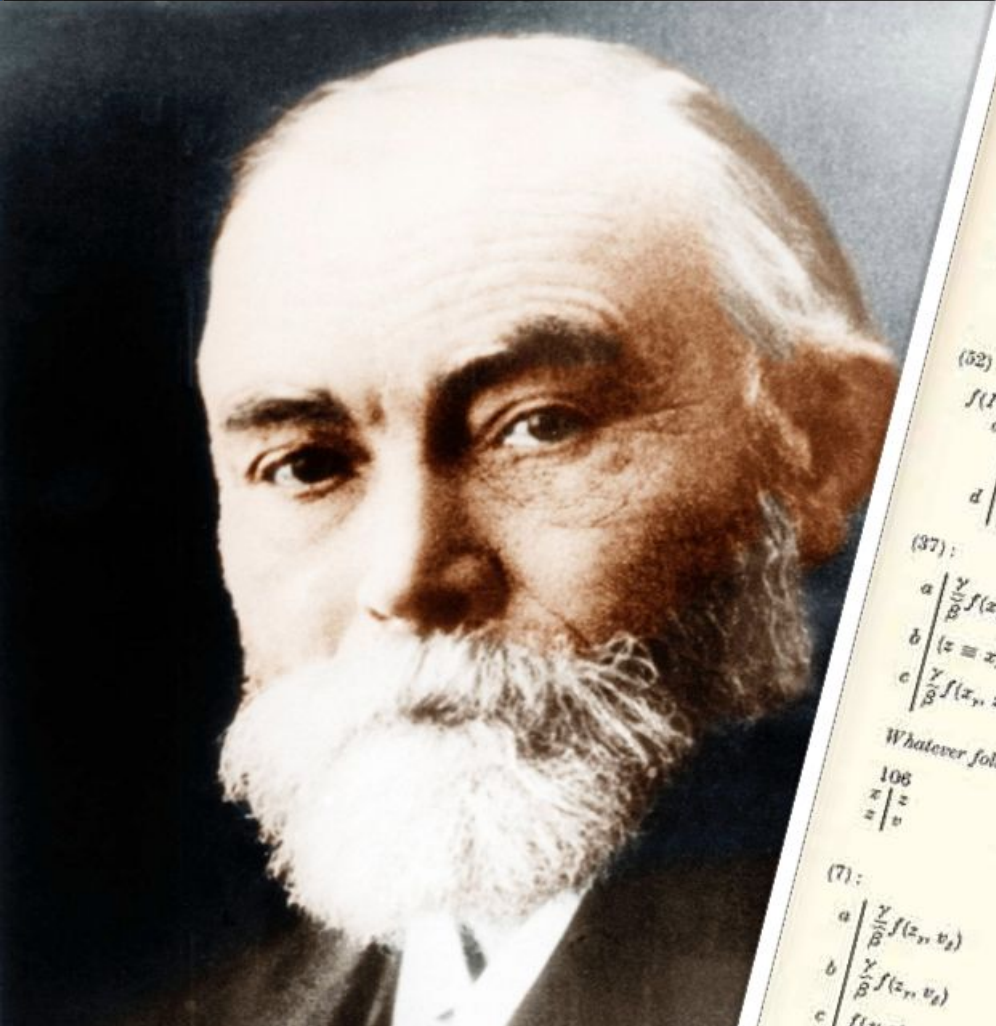


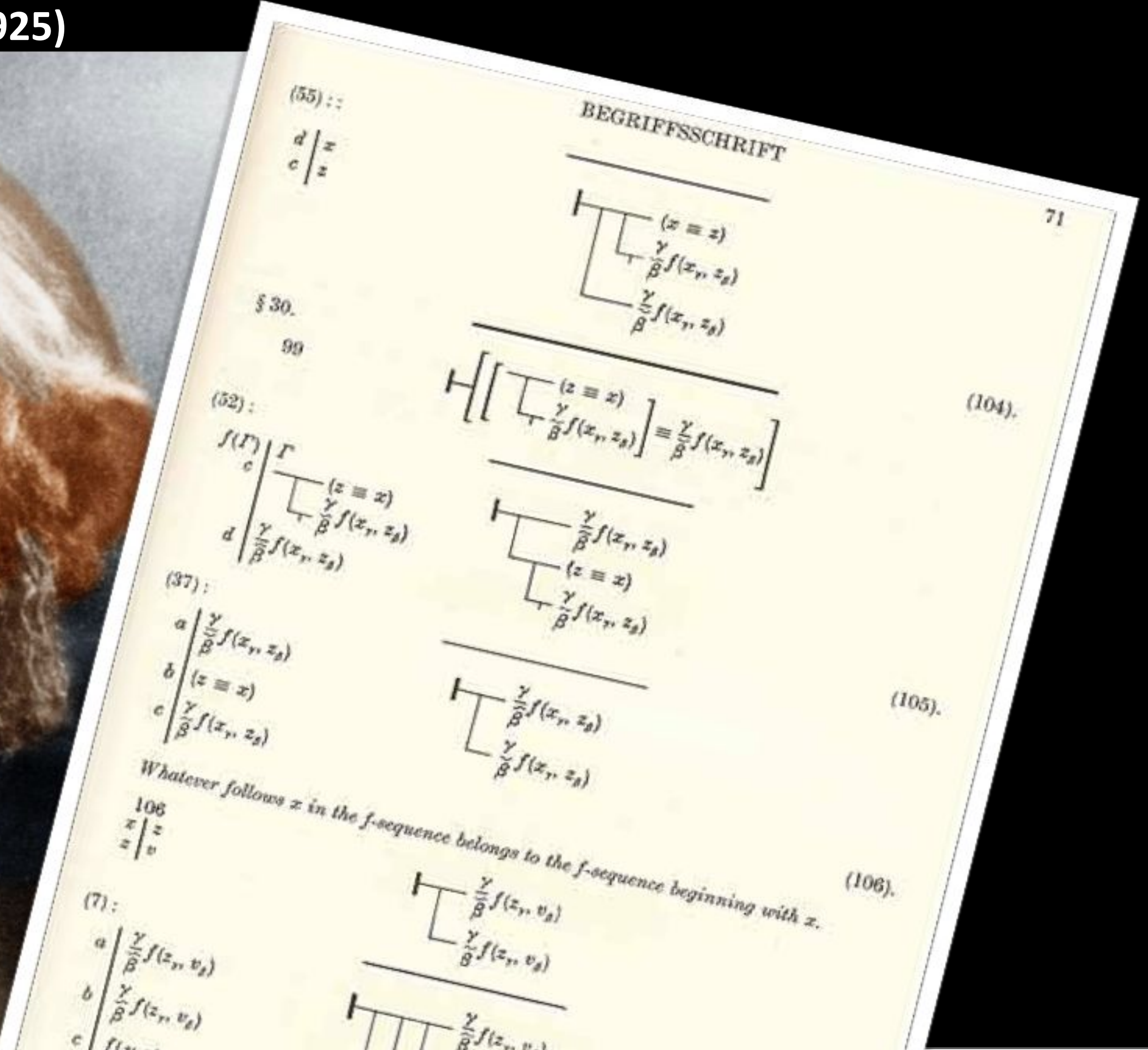
The only way to rectify our reasonings is to make them as tangible as those of the Mathematicians, so that we can find our error at a glance, and when there are disputes among persons, we can simply say: ***Let us calculate*** *[calculemus], without further ado, to see who is right.*

Leibniz in a letter to Ph. J. Spener, Juli 1687

https://upload.wikimedia.org/wikipedia/commons/c/ce/Gottfried_Wilhelm_Leibniz%2C_Bernhard_Christoph_Francke.jpg

Begriffsschrift - Gottlob Frege (1848-1925)





(55) ::

BEGRIFFSSCHRIFT

71

§ 30.

99

(52) :

(104).

(37) :

(105).

Whatever follows x in the f-sequence belongs to the f-sequence beginning with x.

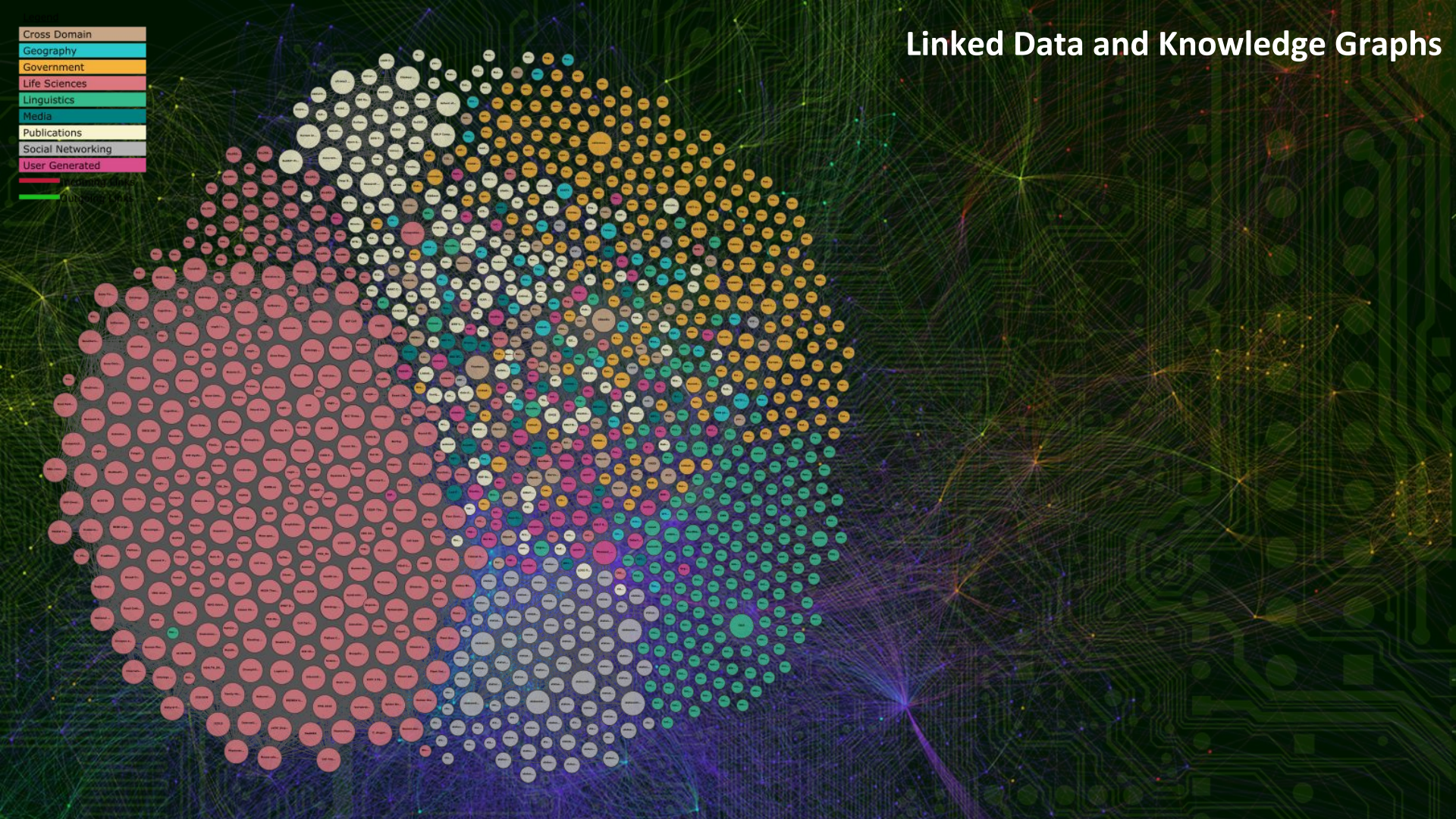
106

(106).

(7) :

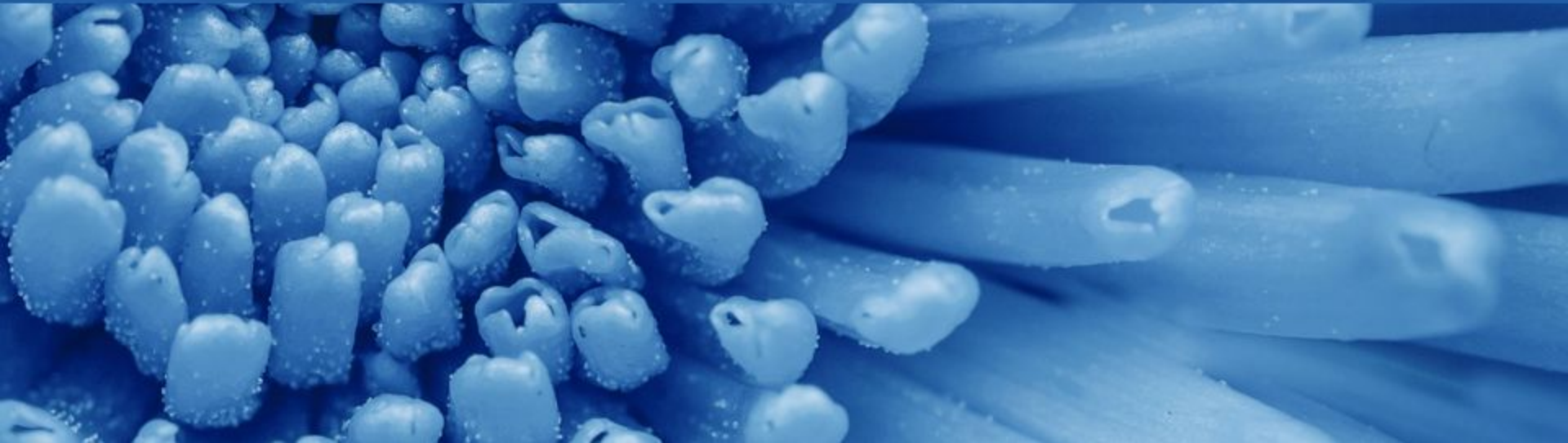
Frame Logic for Knowledge Representation - Marvin Minsky (1974)

https://www.flickr.com/photos/x-ray_delta_one/


Legend
Cross Domain
Geography
Government
Life Sciences
Linguistics
Media
Publications
Social Networking
User Generated
Linked Data and Knowledge Graphs

“In theory there is no difference between theory and practice. In practice there is.”

Yogi Berra



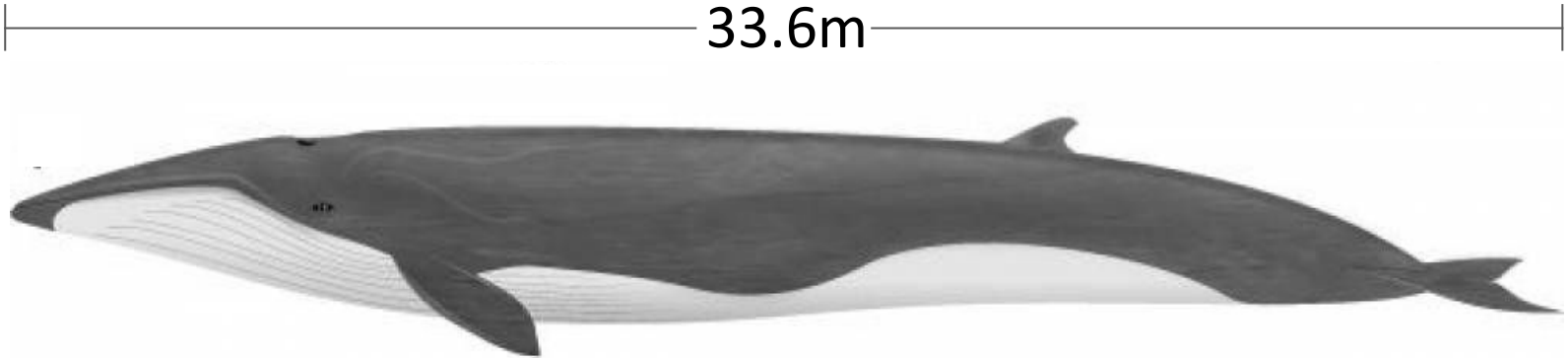
Ontologies in Computer Science

An ontology is an
explicit, formal specification of a shared conceptualization.

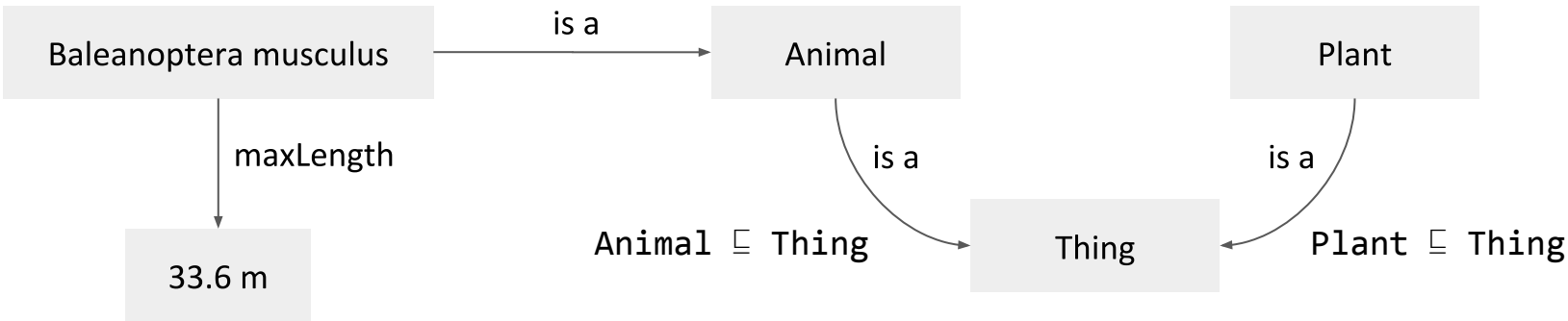
according to Thomas R. Gruber: A Translation Approach to Portable Ontology Specifications. Knowledge Acquisition, 5(2):199-220, 1993.

conceptualization:	abstract model (domain, identified relevant concepts, relations)
explicit:	meaning of all concepts must be defined
formal:	machine understandable
shared:	consensus about ontology

Miniature Example Ontology



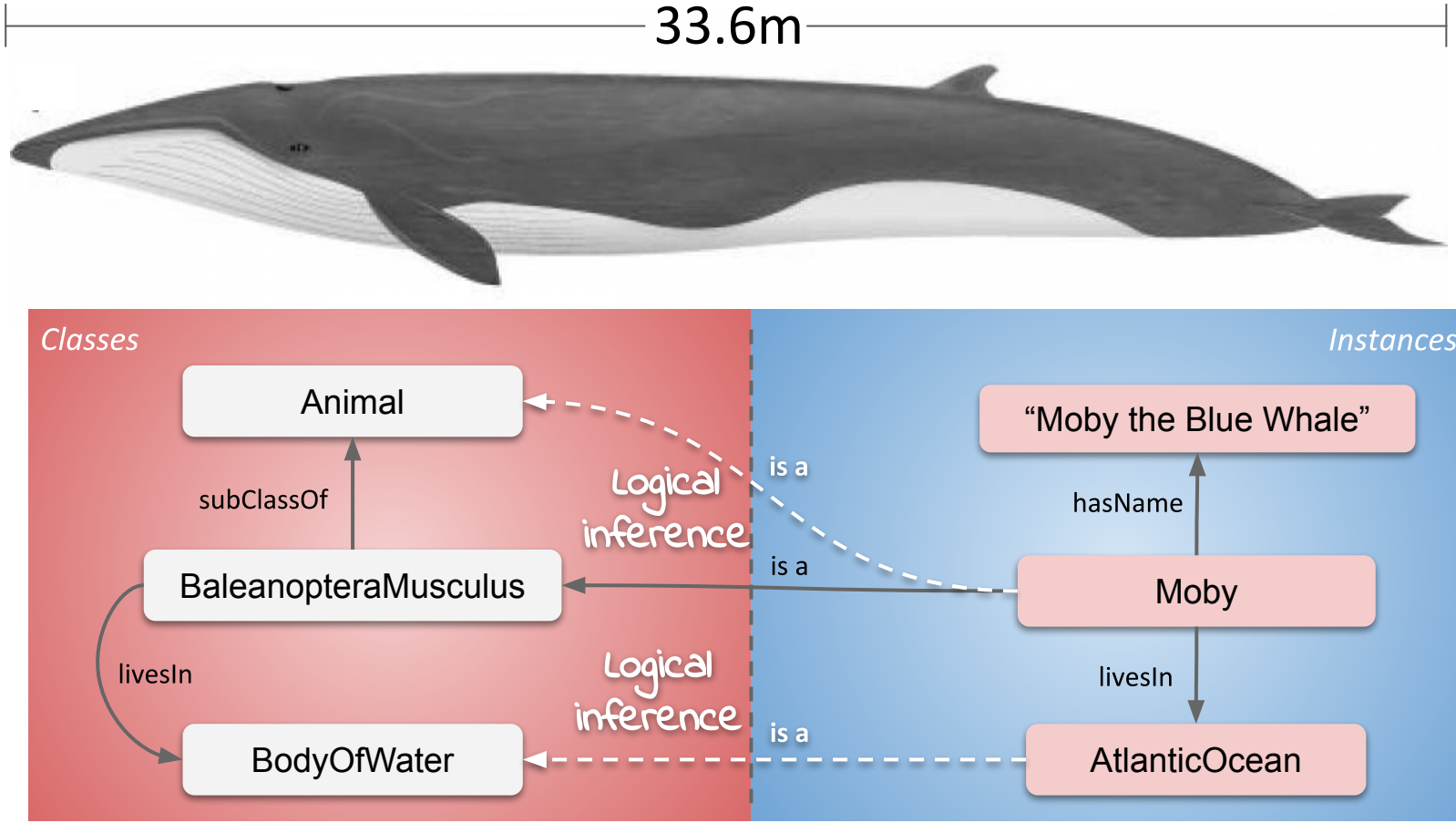
$\texttt{Animal} \sqcap \texttt{Plant} \equiv \varnothing$



$\texttt{BaleanopteraMusculus} \sqsubseteq \texttt{Animal} \sqcap \forall \texttt{maxLength}.\leq 33.6$

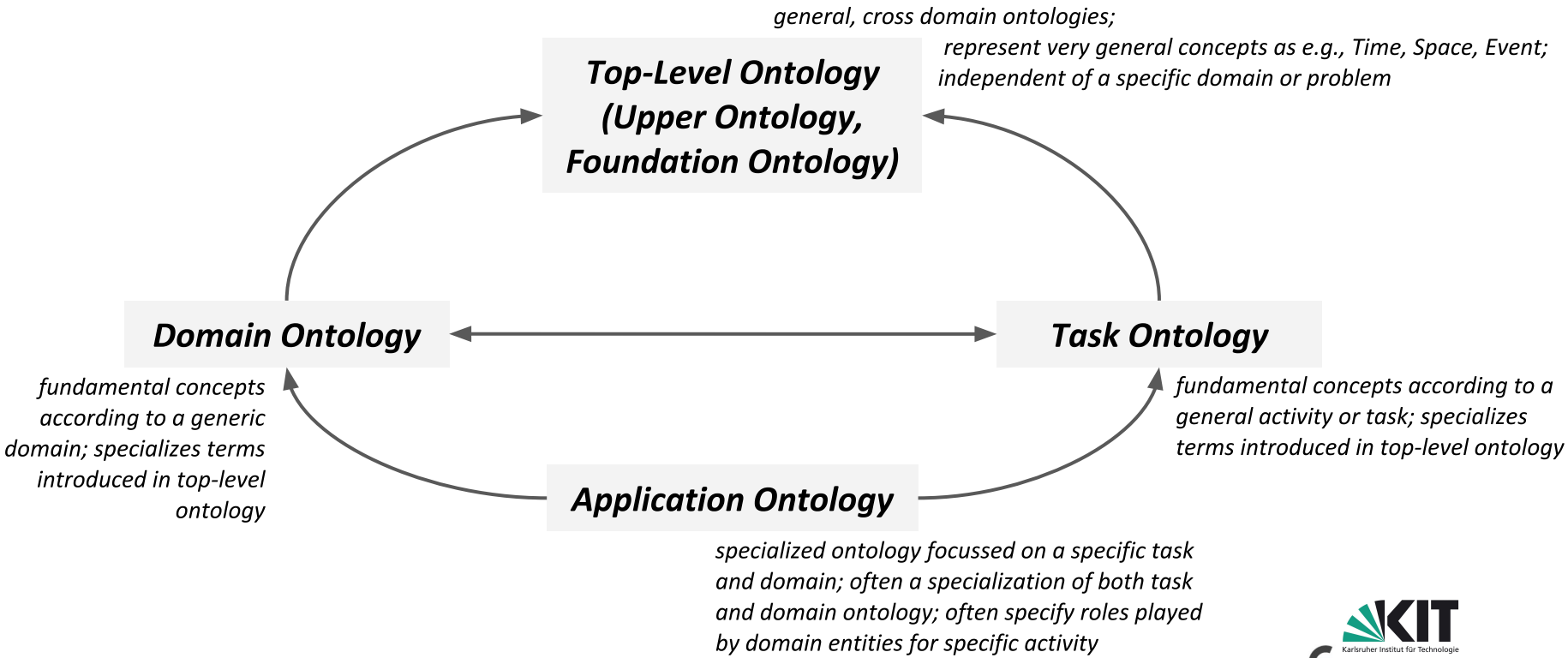
Class Class Property Constraint

Miniature Example Knowledge Base



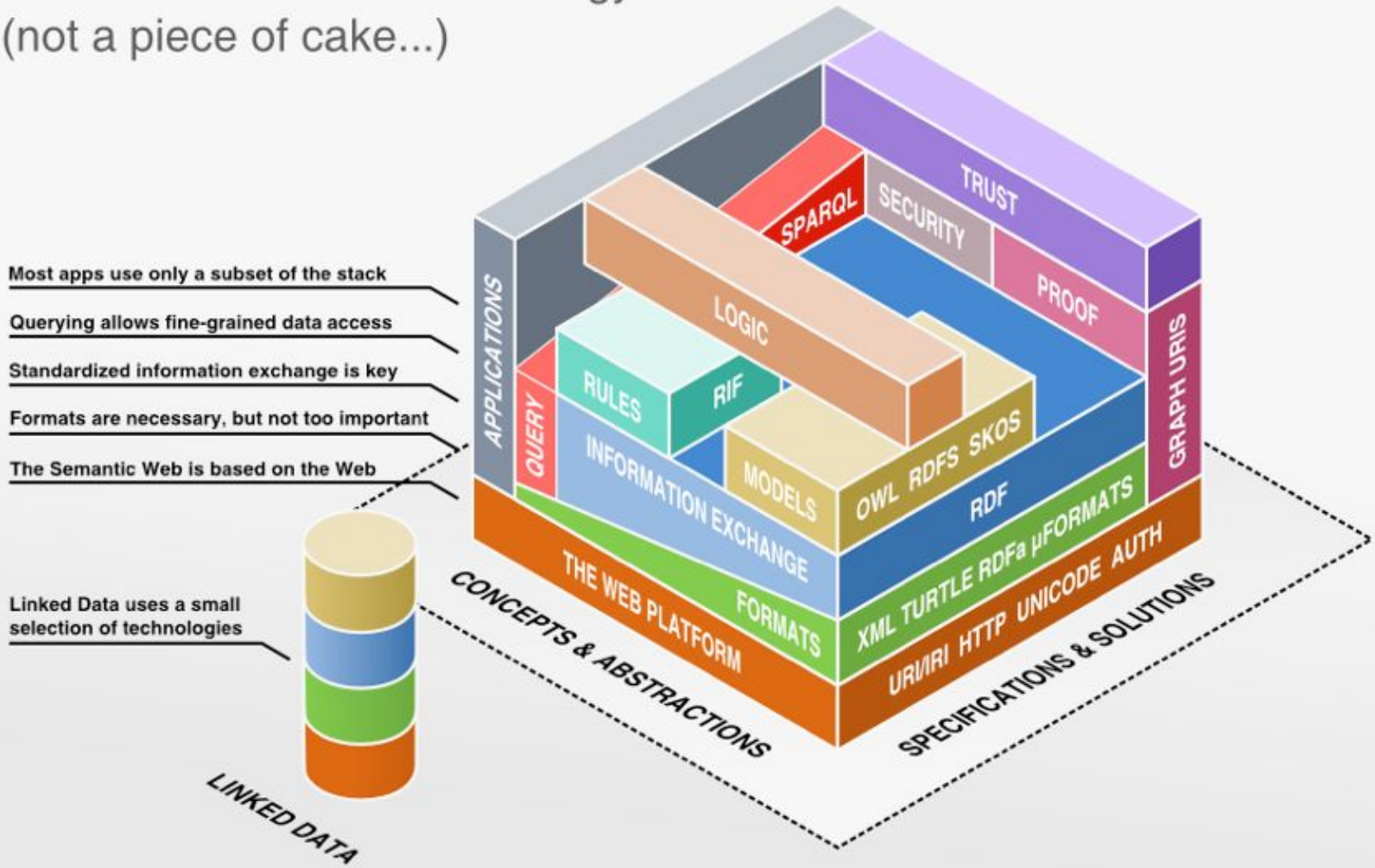
Ontology Types and Categories

according to their level of Generality



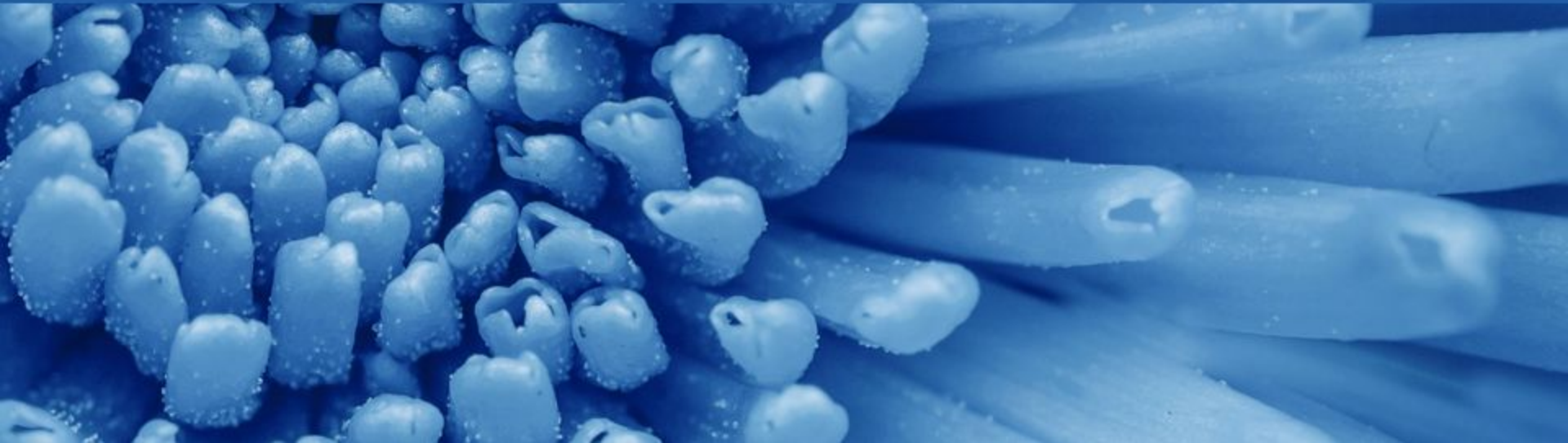
(according to Guarino: Formal Ontology in Information Systems, 1998)

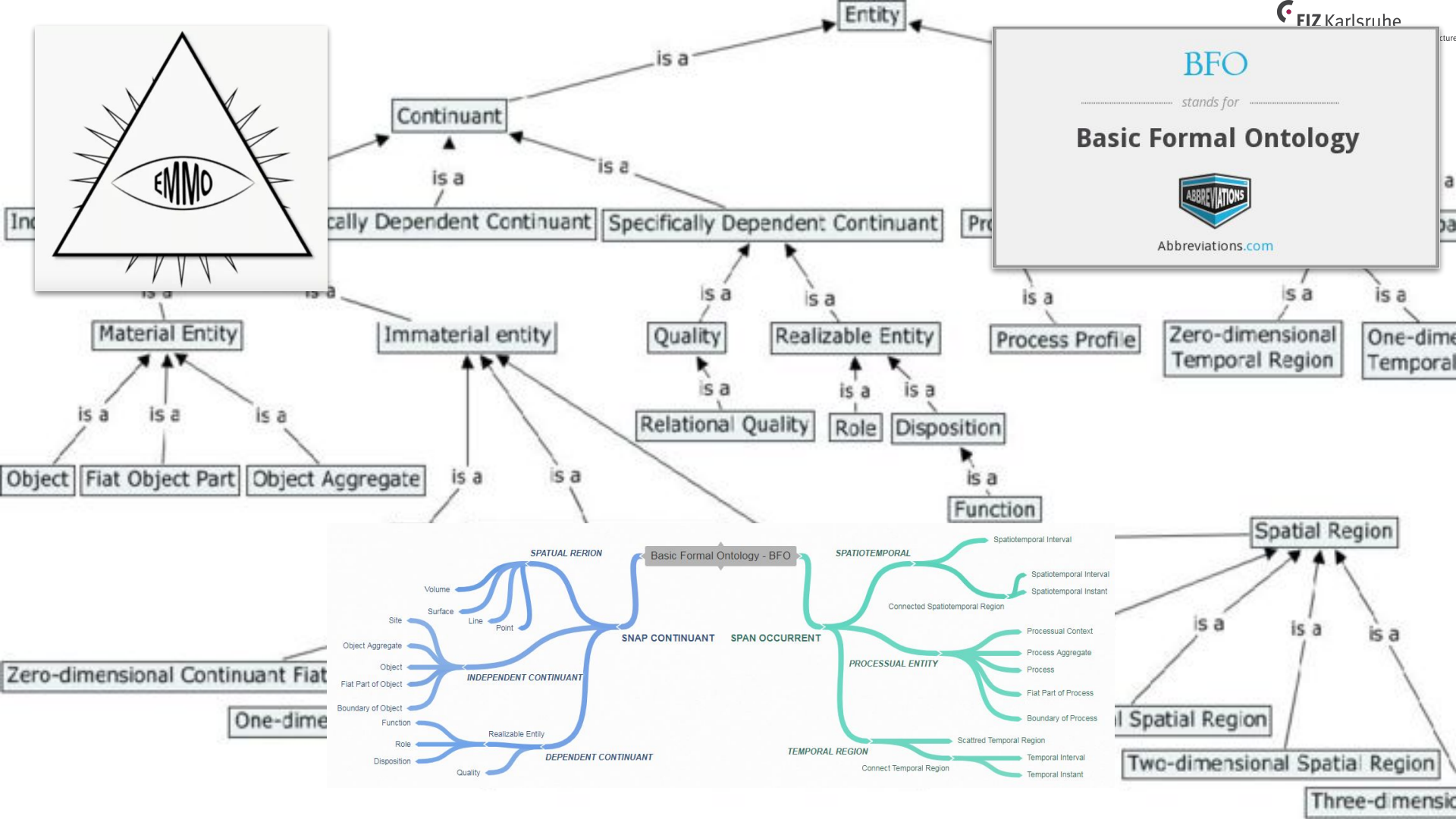
The Semantic Web Technology Stack
(not a piece of cake...)



“Technology presumes there's just one right way to do things and there never is”

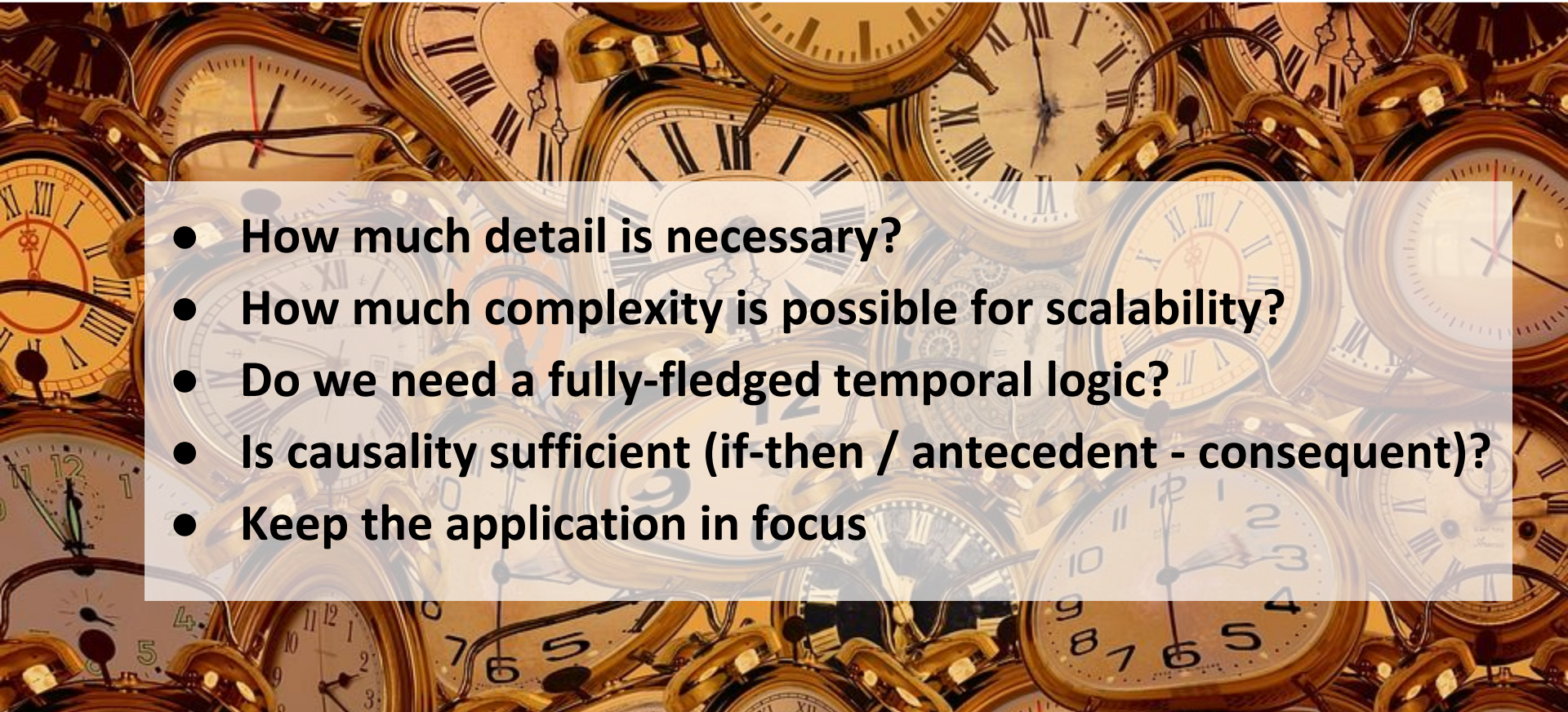
Robert M. Pirsig, Zen and the Art of Motorcycle Maintenance (1974)




FIZ Karlsruhe
EMMO
BFO
stands for
Basic Formal Ontology
ABBREVIATIONS
Abbreviations.com
Entity
is a
Continuant
is a
is a
cally Dependent Continuant
Specifically Dependent Continuant
Material Entity
Immaterial entity
Quality
Realizable Entity
Process Profile
Zero-dimensional Temporal Region
Object
Fiat Object Part
Object Aggregate
Relational Quality
Role
Disposition
Function
Basic Formal Ontology - BFO
SPATUAL RERION
Volume
Surface
Line
Point
Site
Object Aggregate
Object
Fiat Part of Object
Boundary of Object
INDEPENDENT CONTINUANT
Function
Role
Disposition
Realizable Entity
Quality
DEPENDENT CONTINUANT
SNAP CONTINUANT
SPAN OCCURRENT
SPATIOTEMPORAL
Spatiotemporal Interval
Spatiotemporal Interval
Spatiotemporal Instant
Connected Spatiotemporal Region
PROCESSUAL ENTITY
Processual Context
Process Aggregate
Process
Fiat Part of Process
Boundary of Process
TEMPORAL REGION
Scattred Temporal Region
Temporal Interval
Temporal Instant
Connect Temporal Region
Spatial Region
Zero-dimensional Continuant Fiat
One-dime
Spatial Region
Two-dimensional Spatial Region
Three-d mensio

Representation of Time

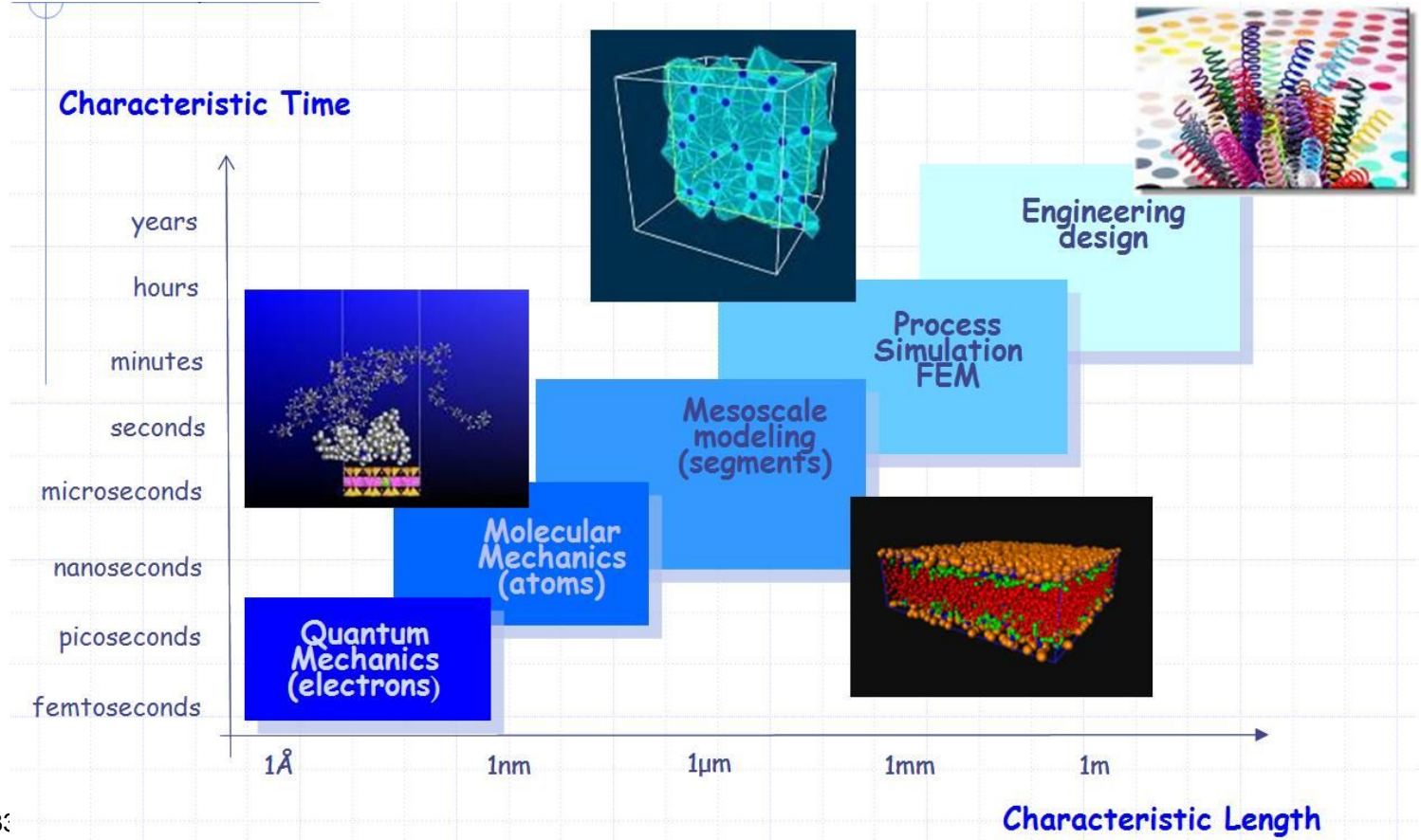
Challenges for Material Sciences



- How much detail is necessary?
- How much complexity is possible for scalability?
- Do we need a fully-fledged temporal logic?
- Is causality sufficient (if-then / antecedent - consequent)?
- Keep the application in focus

Representation of Multiple Scales

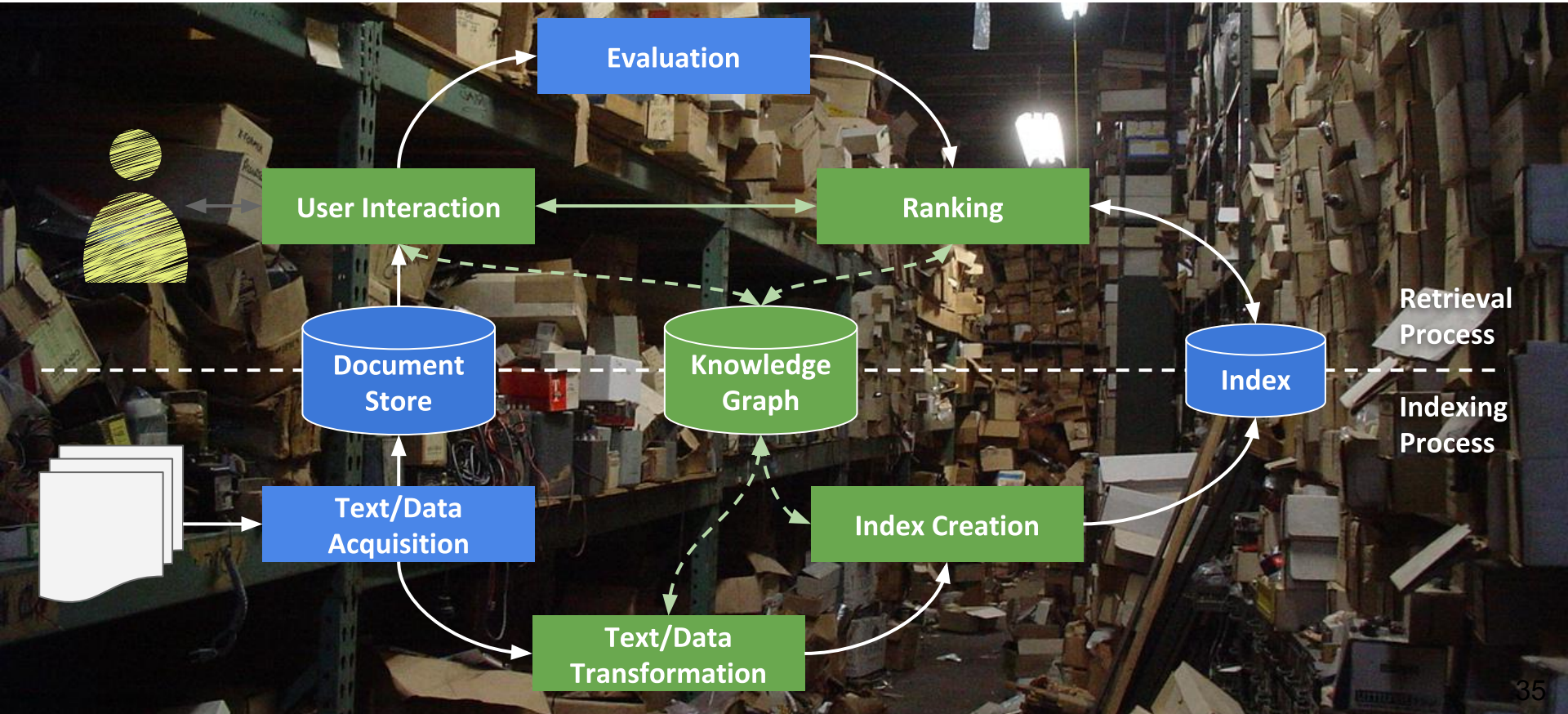
Challenges for Material Sciences



What will be the Application?

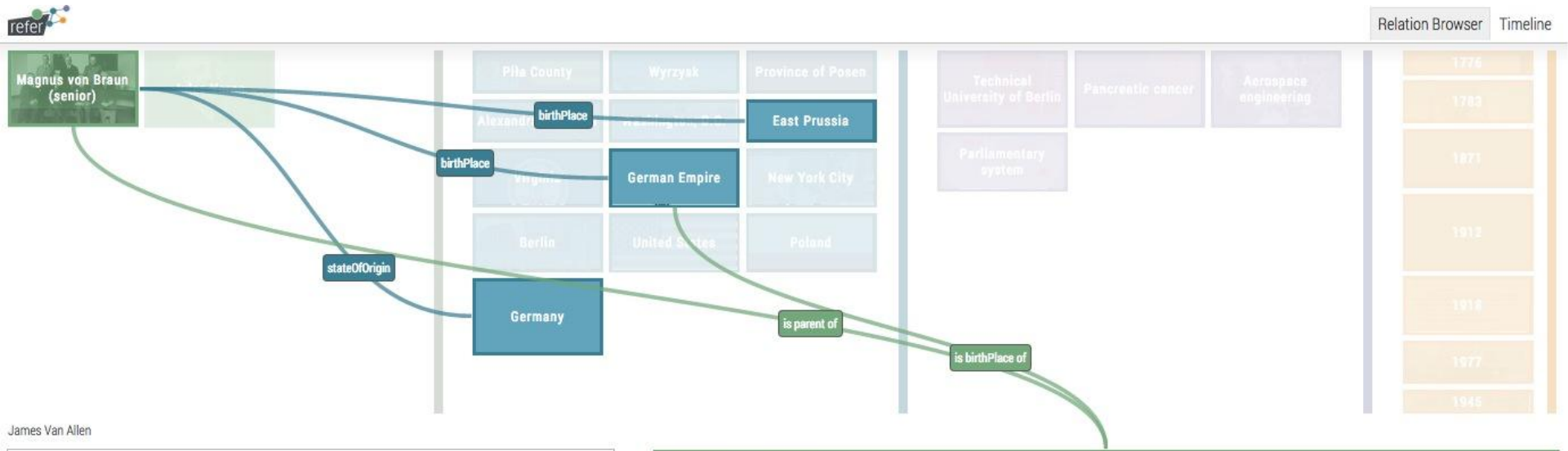
Semantic Search & Retrieval

Application and Purpose of the Material Sciences Ontology



Exploration & Recommendation

Application and Purpose of the Material Sciences Ontology



e.g. via refer.cx WordPress PlugIn at http://scihi.org/

Consistency Checking & Prediction

Application and Purpose of the Material Sciences Ontology

to be continued...

“Nature has given us the seeds of knowledge, not knowledge itself.”

(Lucius Annaeus Seneca, 60AD)

Prof. Dr. Harald Sack

Ontologies for Material Sciences

harald.sack@fiz-karlsruhe.de
twitter: lysander07

BMBF Technologiegespräch “MaterialDigital”
Dresden, 18. Sep. 2019