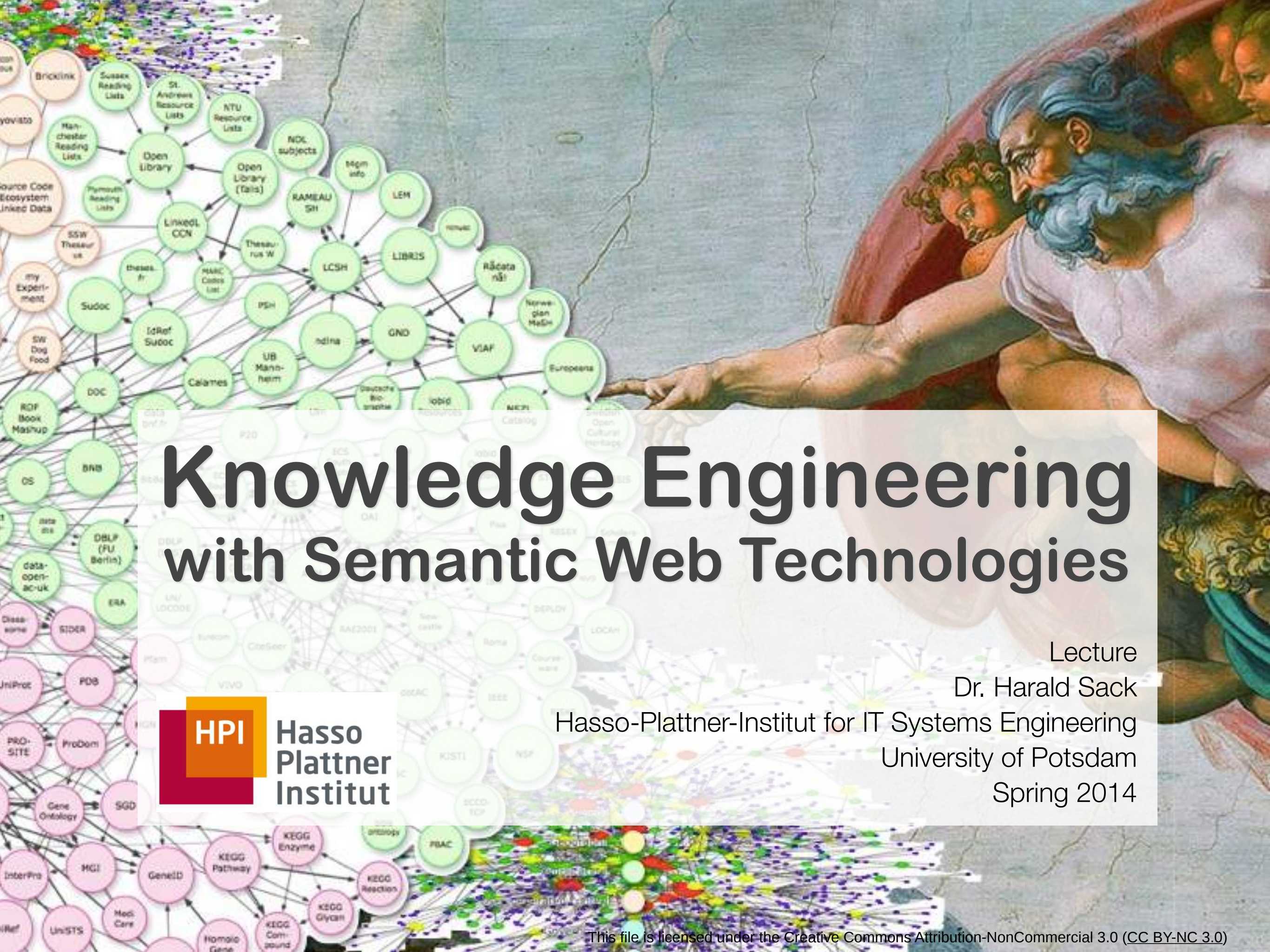




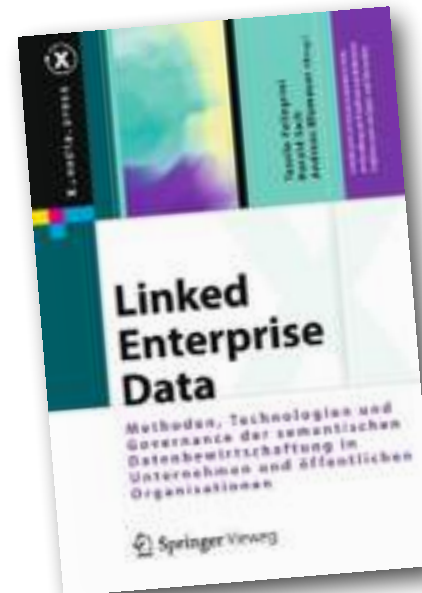
Knowledge Engineering with Semantic Web Technologies



Lecture
Dr. Harald Sack
Hasso-Plattner-Institut for IT Systems Engineering
University of Potsdam
Spring 2014

Knowledge Engineering with Semantic Web Technologies

Dr. Harald Sack



■ Senior Researcher at Hasso Plattner Institute (HPI)

■ Research Topics

- Semantic Web Technologies
- Knowledge Engineering
- Information Retrieval
- Multimedia Analysis & Retrieval
- Social Networking
- Data/Information Visualization



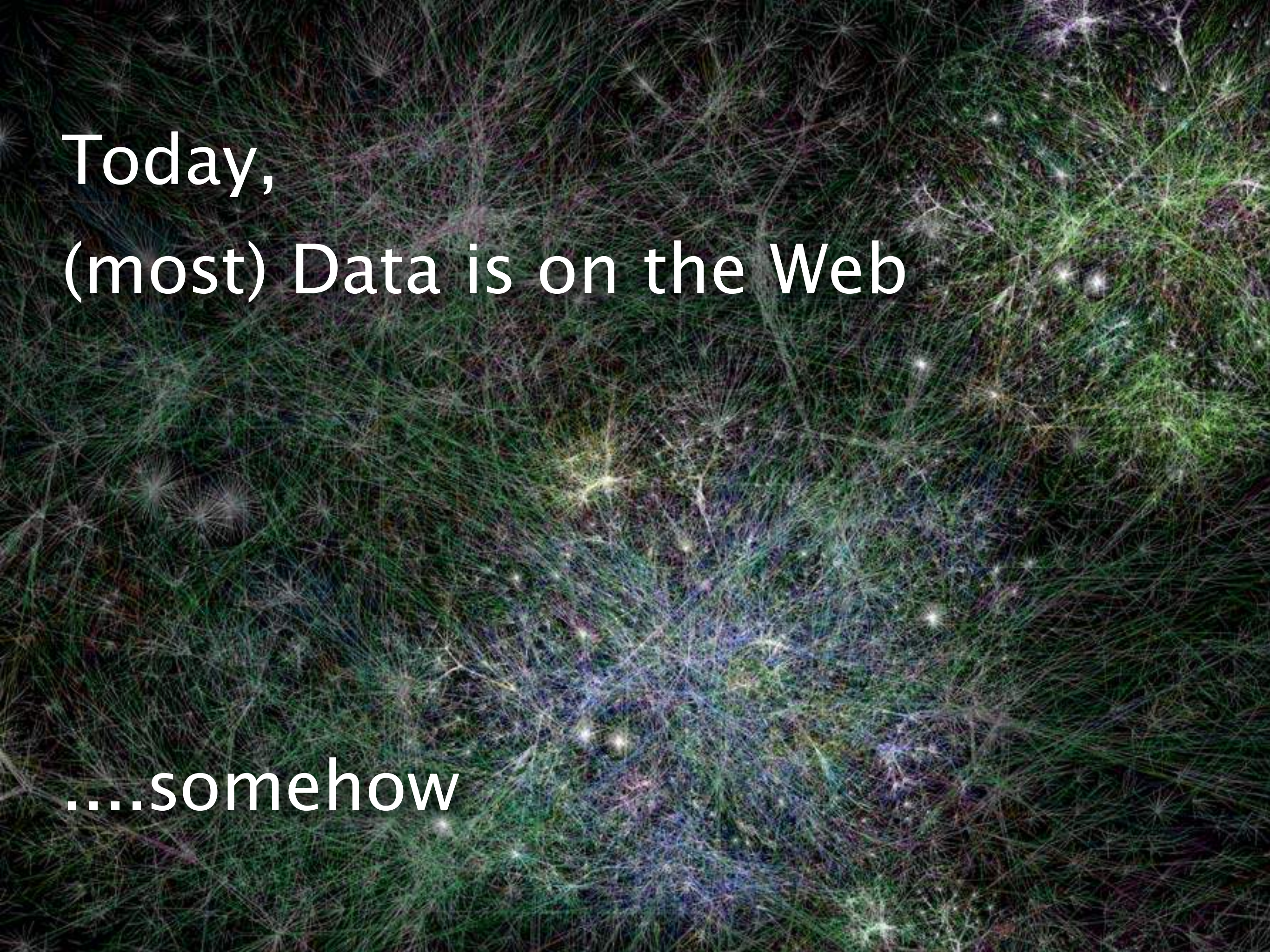
Knowledge Engineering with Semantic Web Technologies

Semantic Technologies Research Group @ HPI



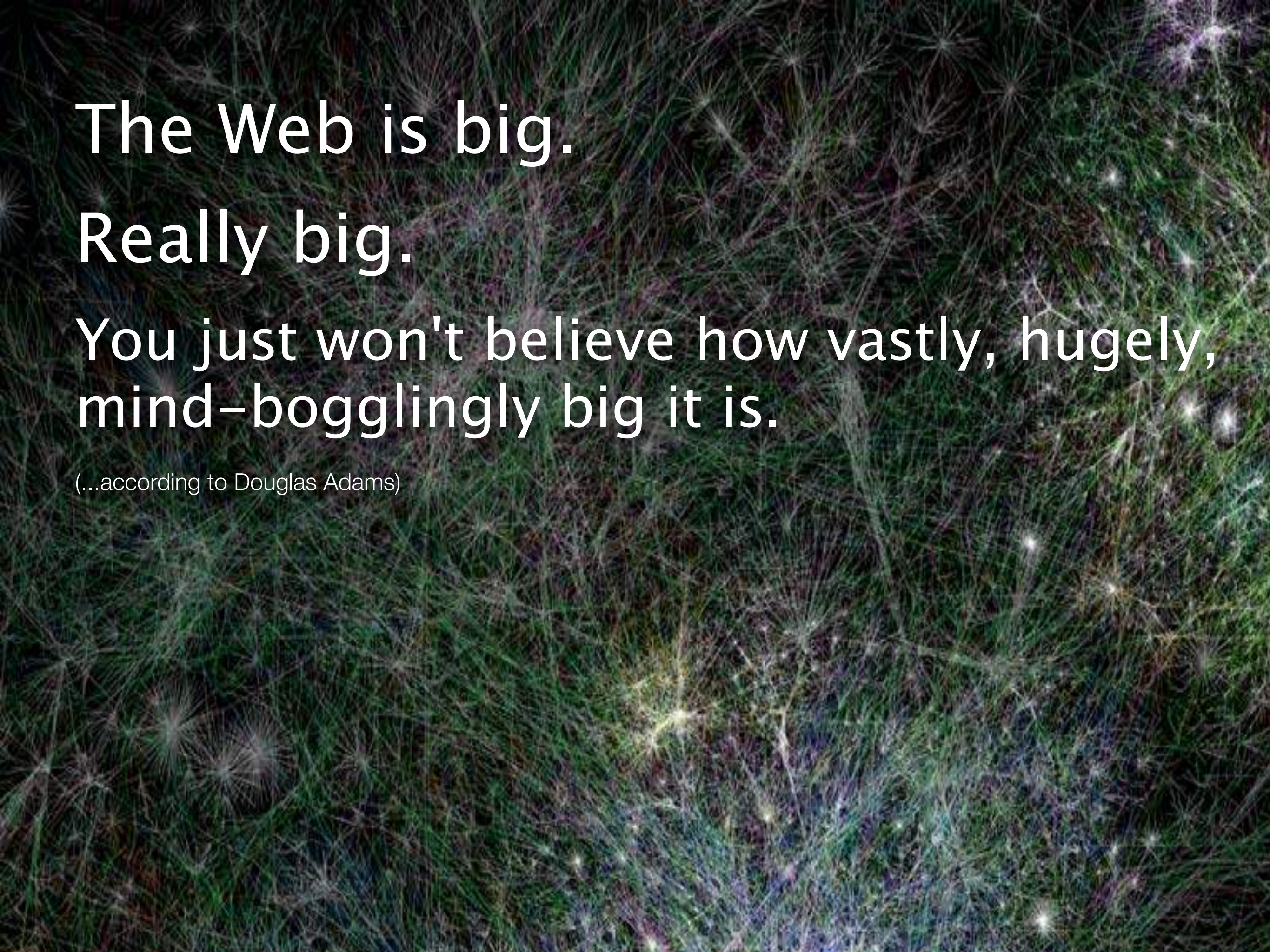


From Data to Knowledge



Today,
(most) Data is on the Web

....somehow



The Web is big.

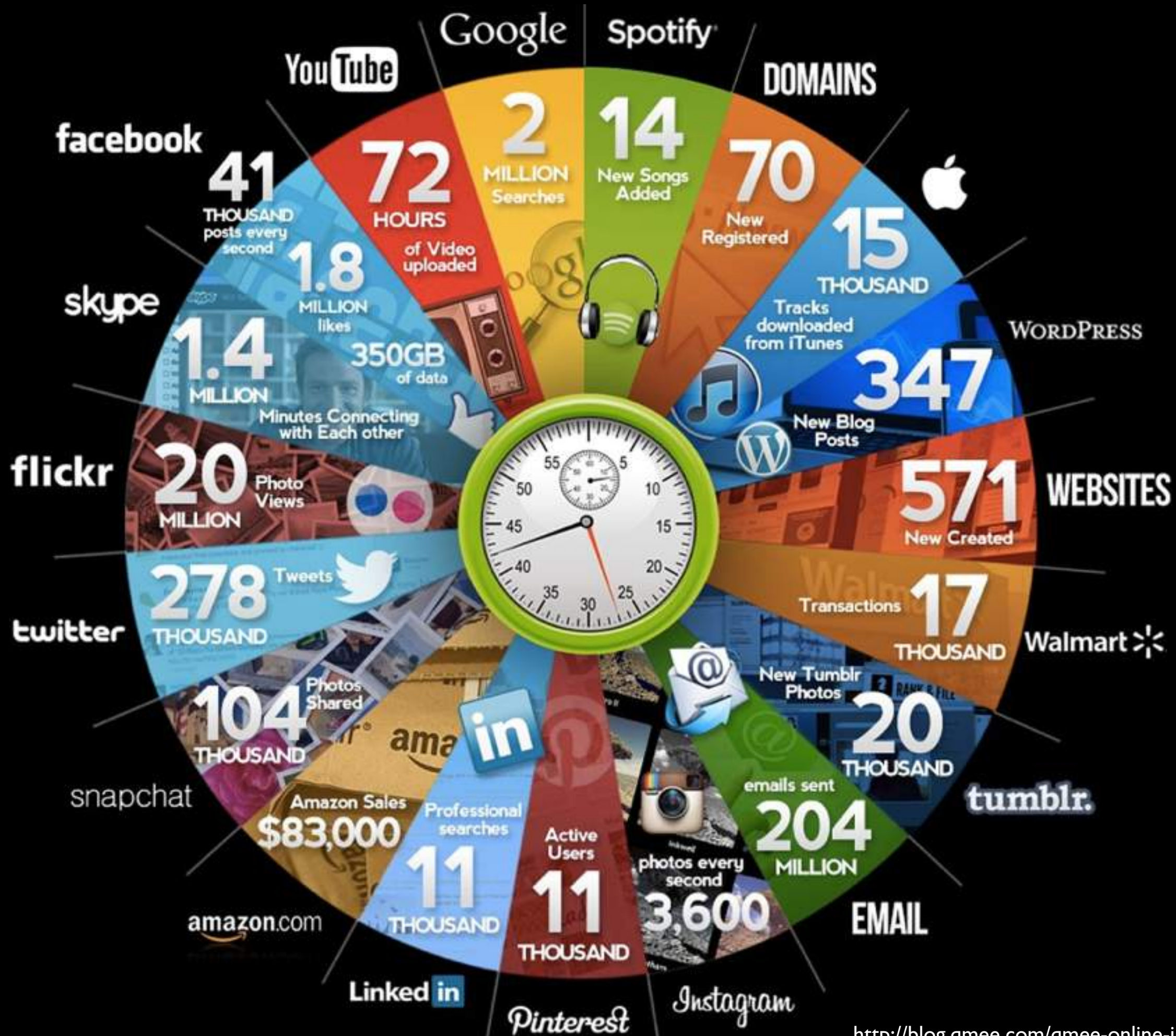
Really big.

You just won't believe how vastly, hugely,
mind-bogglingly big it is.

(...according to Douglas Adams)

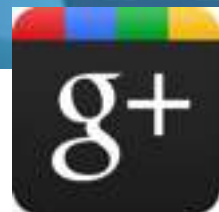
... and it's getting bigger
by the minute!





THE INTERNET OF THINGS

flickr™



facebook

twitter™

www.com
YouTube g

How can we find something on the Web?

10





One Therapist, \$4 Million in Medicare Bills in One Year

By JULIE CRESWELL and ROBERT GEBELOFF

Physical therapy is a big recipient of national Medicare dollars — and physical therapists in Brooklyn are among the biggest billers of all, government data indicates.

33 Comments

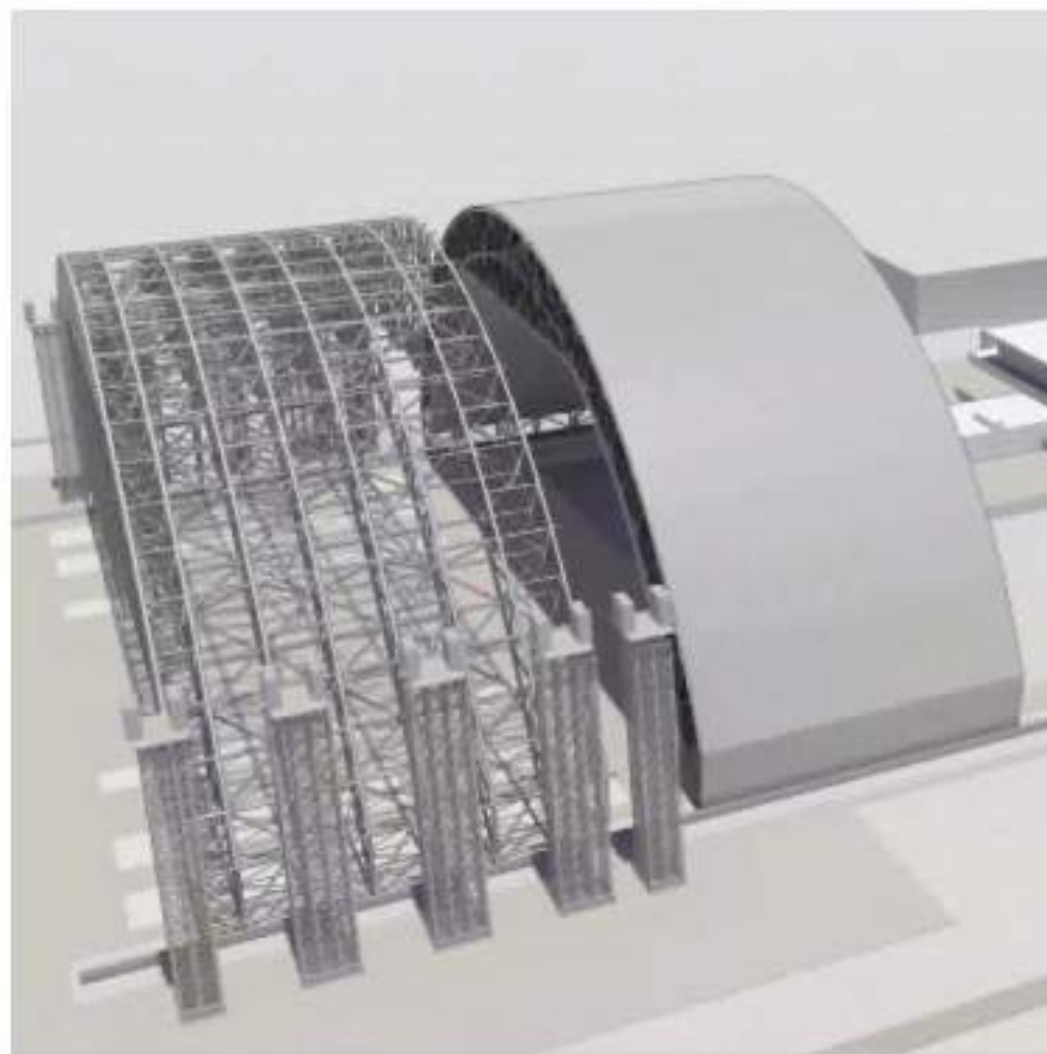
Obama Says More Sanctions Against Russia Are Coming

By MARK LANDLER and PETER BAKER 7 minutes ago

On the last stop of his Asian trip, President Obama said the United States would impose additional sanctions on Russia and freeze some exports of military technology.



• Mayor of Ukraine's Second-Largest City Is Shot



Putting a 32,000-Ton Cap on Chernobyl

By HENRY FOUNTAIN

Chernobyl, Ukraine, is building a \$1.5 billion arch to all but eliminate the risk of further contamination at the site of the 1986 nuclear reactor explosion.

■ Photos: Years After Disaster, Building Progress

String of Deadly Storms Hits Midwest and South

By ALAN BLINDER and ASHLEY SOUTHALL 24 minutes ago

A violent weather system lashed the Midwest and



The Opinion Pages

Smartphones and the 4th Amendment

By THE EDITORIAL BOARD

The police should need a warrant to search mobile phones, because they contain enormous amounts of personal information.

MORE IN OPINION

- Op-Ed: The War on Truth in Ukraine
- ■ Op-Docs: 'Verbatim: What Is a Photocopier?'

OPINIONATOR | THE GREAT DIVIDE No Accounting Skills? No Moral Reckoning

By JACOB SOLL

Our financial illiteracy is costing us much more than we realize.



- Room for Debate: Income and Affirmative Action
- Krugman: High Plains Moochers
- Gessen: Salon of the Exiled

Today's Times Insider

Behind the scenes of The New York Times

- Thinking of Wine as Food With Eric Asimov
- Introducing Times Insider



MARKETS »

At 9:09 AM ET

BUSINESS DAY »

Britain	Germany	France
FTSE 100	DAX	CAC 40
6,690.63	9,422.85	4,446.23
+4.94	+21.30	+2.60
+0.07%	+0.23%	+0.06%

Data delayed at least 15 minutes

Get Quotes | My Portfolios »

THE UPSHOT

Why the Housing Market Is Still Stalling the Economy

Millions are living with parents or friends and just not setting up households.



Understanding Language is difficult...

12

Non-standard English

Great Job @justinbieber! Were soo PROUD of what youve accomplished! U taught us 2 #neversyanever...

segmentation problems

While hunting in Africa, I shot an elephant in my pajamas.

While hunting in Africa, I shot an elephant in my pajamas.

idioms

kick the bucket
pulling my leg
to get cold feet
lose face
raining cats and dogs

neologisms

retweet
webinar
Obamacare
selfie
Brangelina

world knowledge

Mary and Sue are sisters.
Mary and Sue are mothers.

tricky entity names

Let it Be was recorded...
the New York John F. Kennedy airport



WHAT IS THE MEANING G

Approach 1:

Natural Language Processing

Artificial Intelligence

Probabilistic Models

Information Retrieval

Machine Learning



Hasso
Plattner
Institut

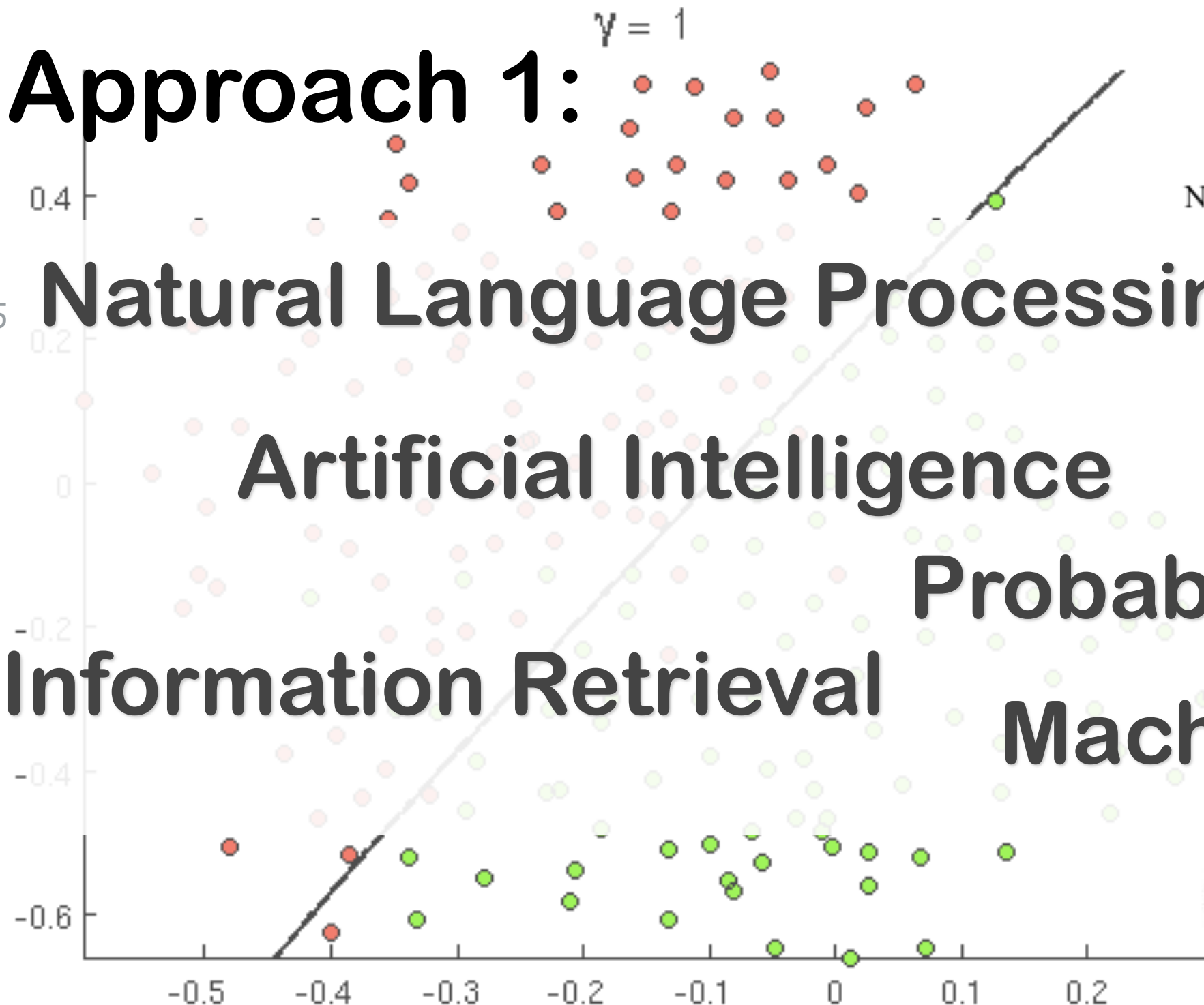
$NP' = \text{Name}'$
 $= \lambda PP(\text{Thetis})$
 $\text{Name}' = \lambda PP(T$

$\epsilon NP'(\lambda$
 $\text{rtal}(z)$

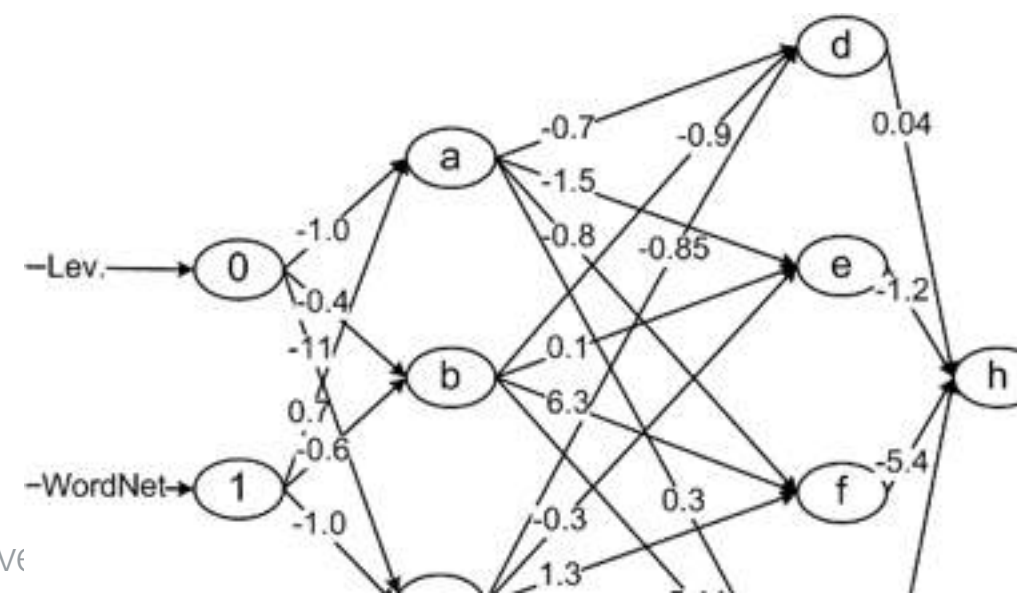
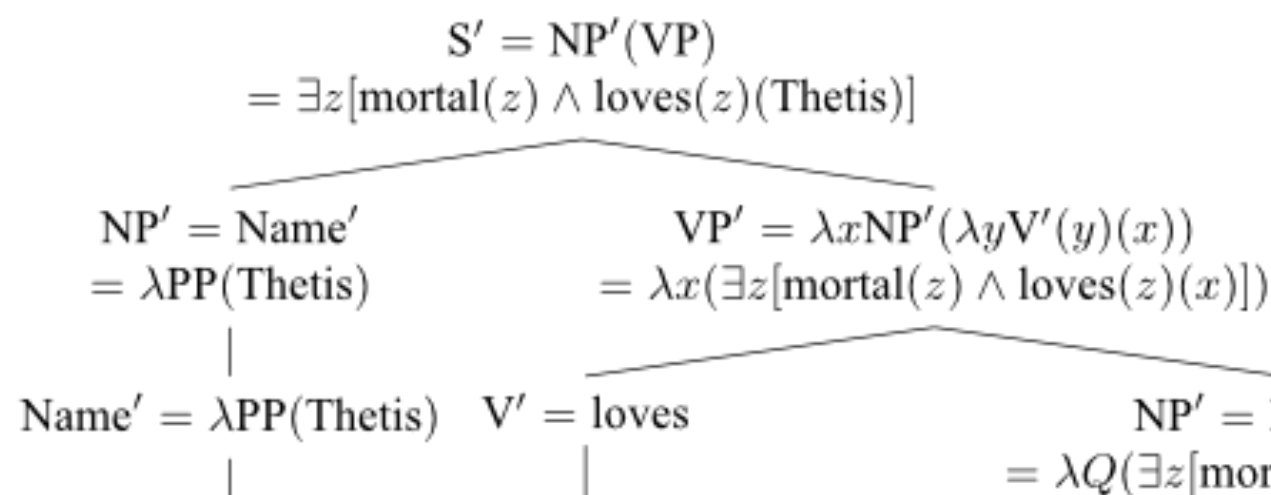
$\text{Def}' = \lambda P \lambda$

Thetis

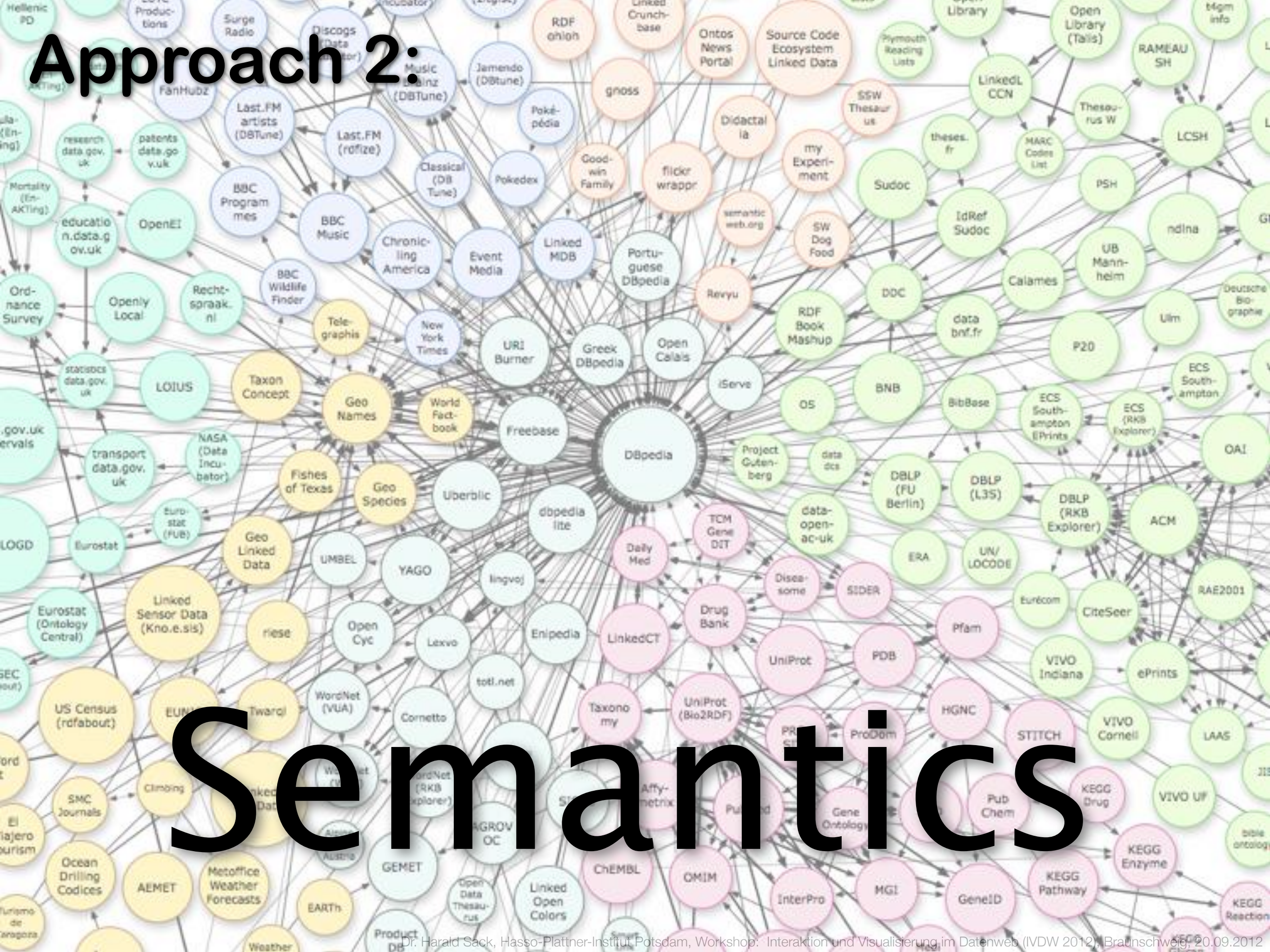
loves



$$TF'(t, D) = \left(\sum_{f \in S} TF_f(t, D) \cdot w_f \cdot \frac{1}{(1 - b_f) + b_f \cdot \frac{DL_f(D)}{AVDL_f}} \right)$$



Approach 2:



Semantics

P A R E N T A L

A D V I S O R Y

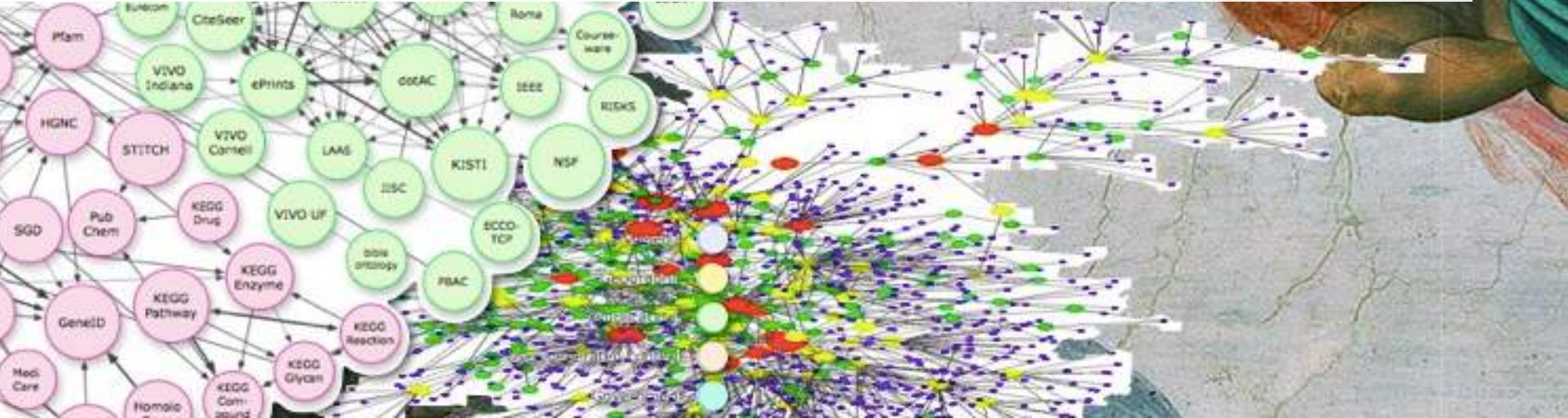
E X P L I C I T S E M A N T I C S

[illegible]

Knowledge Engineering with Semantic Web Technologies

Content

- Lecture 1: Knowledge Engineering and the Web of Data
- Lecture 2: Semantic Web Technologies - Part 1
- Lecture 3: Semantic Web Technologies - Part 2
- Lecture 4: Knowledge Representation - Part 1
- Lecture 5: Knowledge Representation - Part 2
- Lecture 6: Knowledge Engineering

The bottom portion of the slide features a large, intricate network graph. The nodes are represented by circles of various colors (pink, green, yellow, red, blue) and sizes, some containing text labels such as 'HGNC', 'KEGG Drug', 'KEGG Enzyme', 'KEGG Pathway', 'KEGG Reaction', 'KEGG Glycan', 'KEGG Compound', 'Homolo', 'Medi Care', 'GenelD', 'Pub Chem', 'SGD', 'STITCH', 'VIVO Cornell', 'VIVO Indiana', 'CiteSeer', 'ePrints', 'LAMS', 'IISC', 'KISTI', 'NSF', 'BCCO-TOP', 'PBAC', 'bioRxiv', 'KEGG Reaction', 'KEGG Glycan', 'KEGG Compound', 'Homolo', 'Medi Care', 'GenelD', 'Pub Chem', 'SGD', 'STITCH', 'VIVO Cornell', 'VIVO Indiana', 'CiteSeer', 'ePrints', 'LAMS', 'IISC', 'KISTI', 'NSF', 'BCCO-TOP', 'PBAC', 'bioRxiv'. The edges are thin black lines connecting these nodes, forming a dense web of relationships. The background of the slide is a light gray with faint, larger-scale network patterns. On the right side, there is a vertical strip showing a classical painting of a person's face and shoulders.

Knowledge Engineering with Semantic Web Technologies

Content

Lecture 1: Knowledge Engineering and the Web of Data

- The history of the Web
- Limits of the (current) Web
- The Web becomes „intelligent“
- The Web of Data and how to make use of it

Ceci n'est pas une pomme.

Knowledge Engineering with Semantic Web Technologies

Content



Lecture 2: Semantic Web Technologies - Part 1

- *How to name things?*
Uniform Resource Identifier - URI
- *How to represent simple facts?*
Resource Description Framework - RDF
- *How to represent more complicated models?*
RDF Schema

Knowledge Engineering with Semantic Web Technologies

Content

Lecture 3: Semantic Web Technologies - Part 2

- *How to ask for (semantic) data?*
Query Language SPARQL
- *How to store (semantic) data?*
RDF Databases - Tripel Stores
- *How to make the Web smarter?*
RDFa and microdata

Knowledge Engineering with Semantic Web Technologies

Content

22

Lecture 4: Knowledge Representation - Part 1

- *Ontology One on One*
from Philosophy to Computer Science
- *The Foundations of Mathematical Logic*
and why do we need it
- *Propositional Logic and First Order Logic*
a short recapitulation

Knowledge Engineering with Semantic Web Technologies

Content

23

Lecture 5: Knowledge Representation - Part 2

- *How to describe the World in a formal way?*
Description Logics
- *Logics becomes part of the Web*
OWL and OWL-Semantics
- *and what if things gets more complicated...?*
Axiomatic Knowledge with Rules

Knowledge Engineering with Semantic Web Technologies

The Handy Tool Set

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

Most apps use only a subset of the stack

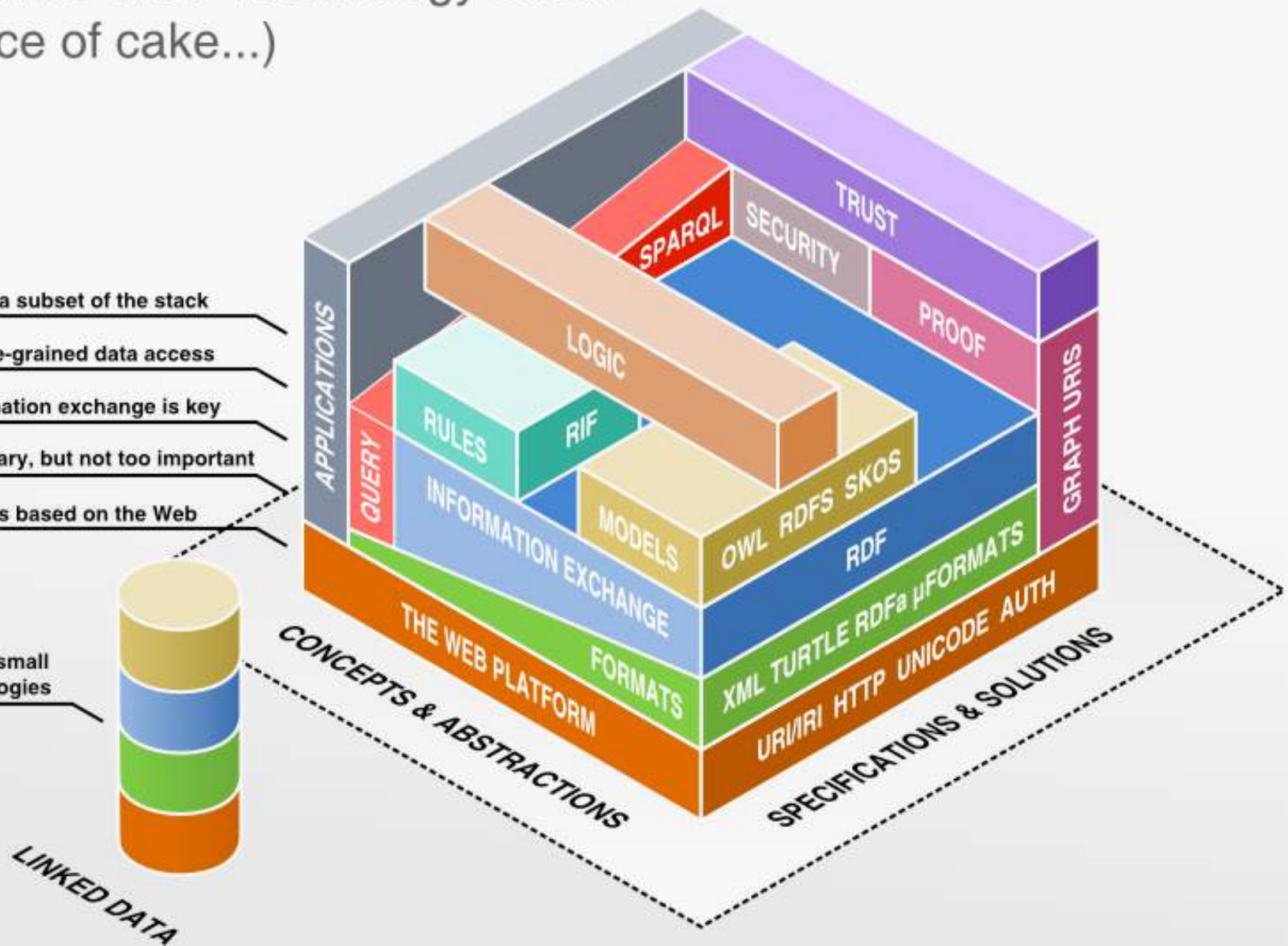
Querying allows fine-grained data access

Standardized information exchange is key

Formats are necessary, but not too important

The Semantic Web is based on the Web

Linked Data uses a small selection of technologies



Knowledge Engineering with Semantic Web Technologies

Content

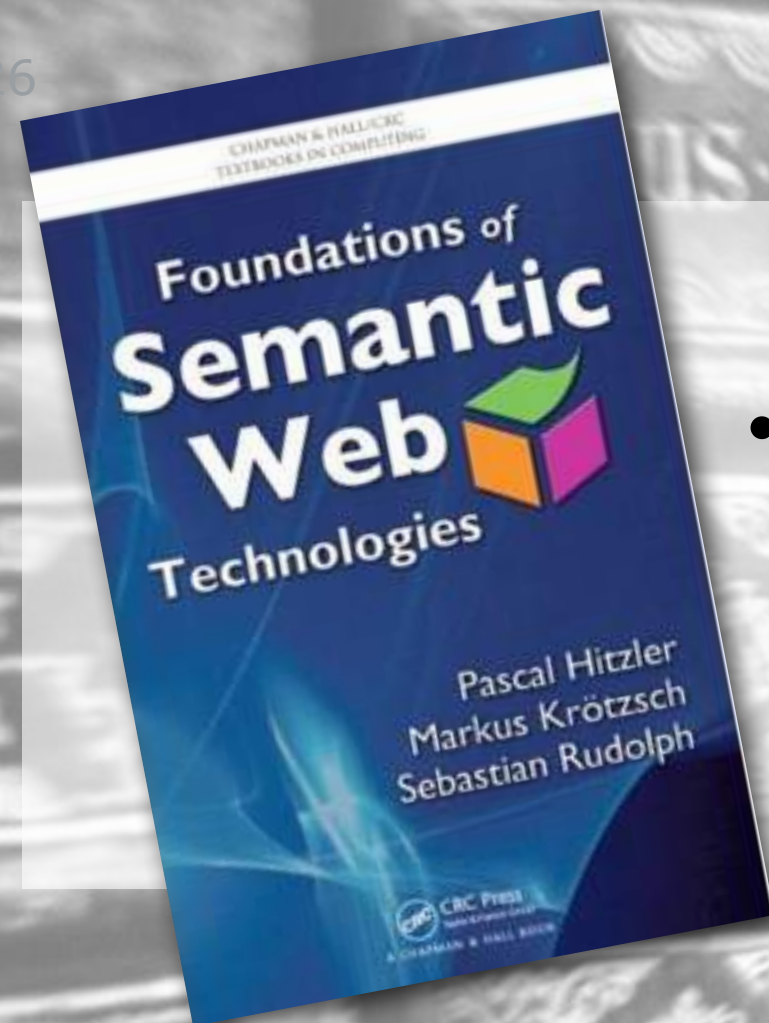
25

Lecture 6: Knowledge Engineering

- *How to handle knowledge?*
Ontological Engineering
- *The very Basics of Understanding*
Named Entity Recognition & Disambiguation
- *Knowledge Engineering in the Web of Data*
Linked Data Engineering
- *And what can all this be used for...?*
Search beyond Google

Knowledge Engineering with Semantic Web Technologies

Bibliography



- P. Hitzler, M. Krötzsch, S. Rudolph:
Foundations of Semantic Web Technologies,
CRC Press, 2009.

Knowledge Engineering with Semantic Web Technologies

Bibliography

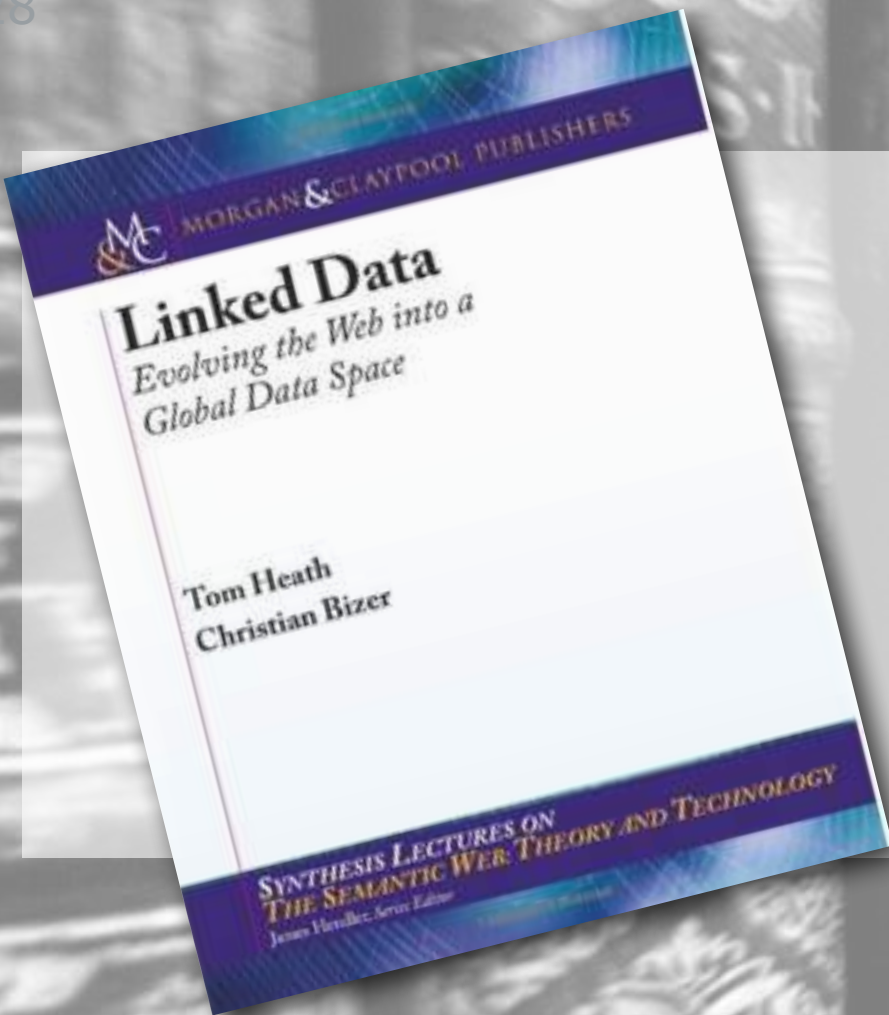


- U. Schöning:
Logic for Computer Scientists,
Birkhäuser, 2008.

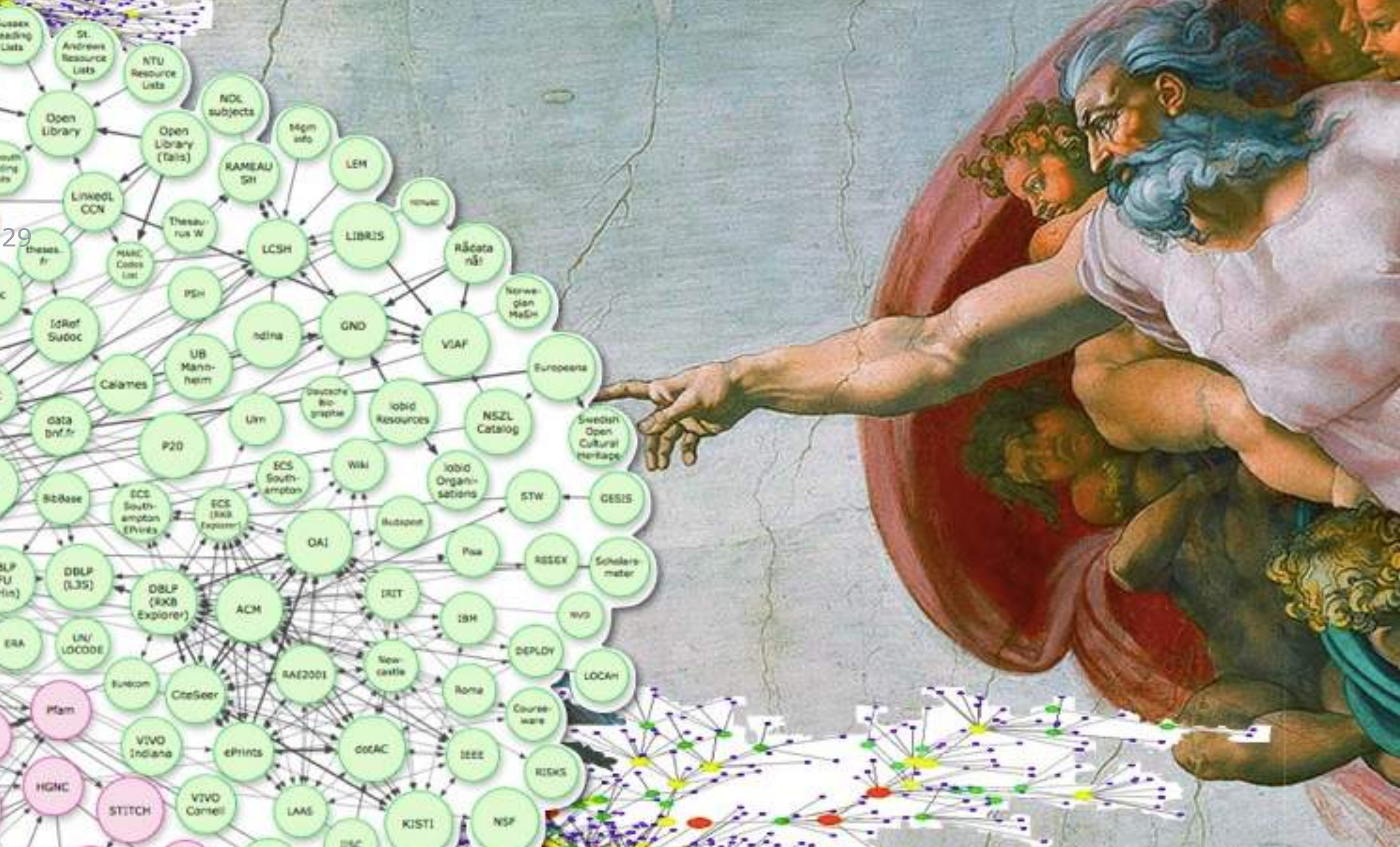
Knowledge Engineering with Semantic Web Technologies

Bibliography

8



- T. Heath, Ch. Bitzer
Linked Data - Evolving the Web into a Global Data Space,
Morgan & Claypool, 2011.



Knowledge Engineering with Semantic Web Technologies