Summer 8-6-2024

Knowledge Management and Semantic Reasoning: Ontology and Information Theory Enable the Construction of Knowledge Bases and Knowledge Graphs

Quynh D. Tran
Case Western Reserve University, qdt@case.edu

Ozan Dernek
Case Western Reserve University, oxd37@case.edu

Erika I. Barcelos
Case Western Reserve University, eib14@case.edu

Laura S. Bruckman
Case Western Reserve University, lsh41@case.edu

Roger H. French
Case Western Reserve University, rxf131@case.edu

Author(s) ORCID Identifier:
Quynh D. Tran
Ozan Dernek
Erika I. Barcelos
Laura S. Bruckman
Roger H. French

Materials Data Science for Stockpile Stewardship

COE: US-Department of Energy-NNSA Award

MDS³ CENTER OF EXCELLENCE

Lawrence Livermore National Laboratory · Los Alamos NATIONAL LABORATORY EST.1943 · Sandia National Laboratories · KANSAS CITY NSC National Security Campus · The Nuclear Security Enterprise NNSA

CASE WESTERN RESERVE UNIVERSITY EST. 1826 think beyond the possible™

Knowledge Management and Semantic Reasoning: Ontology and Information Theory Enable the Construction of Knowledge Bases and Knowledge Graphs

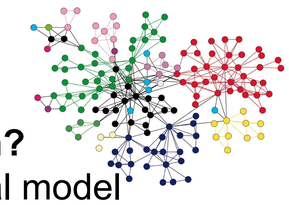
Quynh D. Tran[1,2], Ozan Dernek[1,2], Erika I. Barcelos[1,2], Laura S. Bruckman[1,2], Roger H. French[1,2]

[1]Materials Data Science for Stock Stewardship Center of Excellence, Cleveland OH, USA

[2]Case Western Reserve University, Cleveland OH, USA

INTRODUCTION OF CONCEPTS

❖ **What even is knowledge?**
- ➢ Facts & statements about entities
- ➢ How to do something (skills)
- ➢ What happened or did not happen



❖ **What is a graph?**
- ➢ A mathematical model
- ➢ A set of nodes and edges

<u>Knowledge Graph</u>: a set of statements/facts/results about an entity or entities that can be represented as a set of nodes and edges and can be reasoned on to infer new knowledge

❖ **Ontology** unifies the logical relations & definitions of domain information
- ➢ Provides a schema layer to KGs that enables **semantic reasoning**
- ➢ <u>Must be connected to top and mid-level ontologies for interoperability</u>
- ➢ Enables more understanding of workings “under-the-hood” for humans

KNOWLEDGE GRAPH ≠ **GRAPH NEURAL NETWORK**

❖ **GNN lacks semantics supplied by ontology**

❖ **Knowledge base** or **knowledge repository** is required
- ➢ Storage of large number of **RESULTS** for KG semantic reasoning

❖ Knowledge graph is capable of extracting insights from hundreds or more results in high dimensional design space
- ➢ Humans have traditionally done the reasoning (cognition);
  - ■ However, we are limited to low dimensional space
  - ■ Deductive, inductive, and abductive reasoning.

❖ **Information Theory (IT)** provides model-independent
- ➢ Uncertainty quantification, noise rejection
- ➢ And compact knowledge representation, latent space manifolds

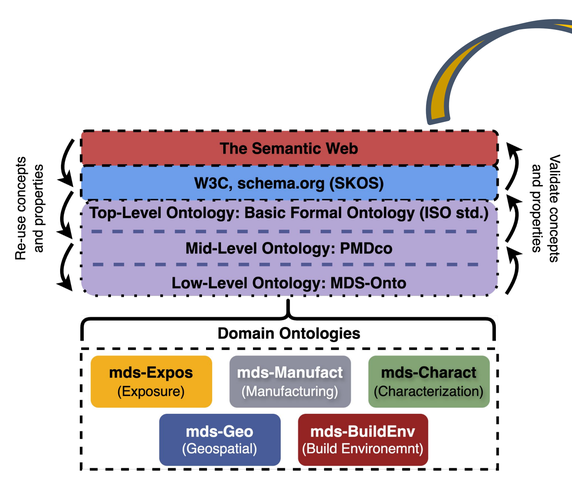


Fig1: MDS-Onto, ontologies of subdomains at MDS³, is connected to mid-level and top-level (BFO) ontologies to ensure interoperability.

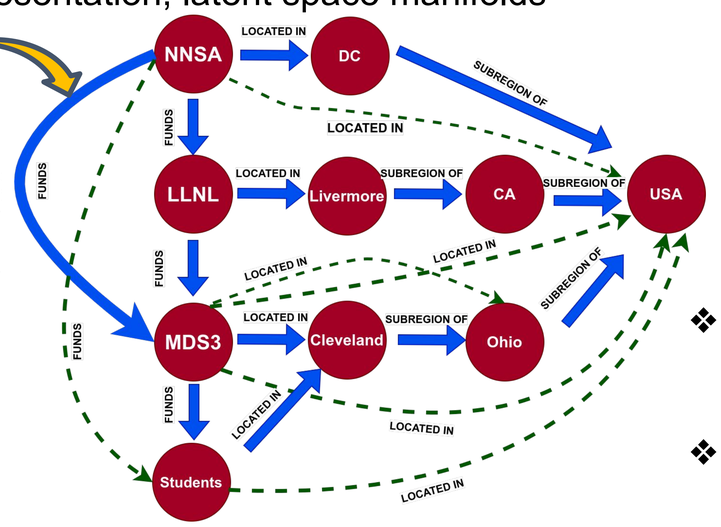


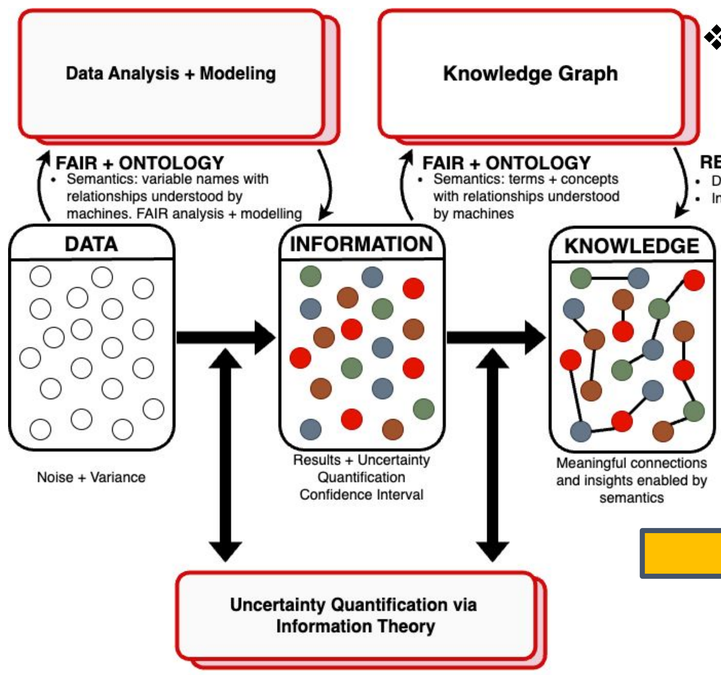
Fig.2: An example of a knowledge graph describing the funding source and location of NNSA, LLNL, MDS³-COE and its students

DYNAMIC DATABASE WITH KNOWLEDGE GRAPH AND INFORMATION THEORY

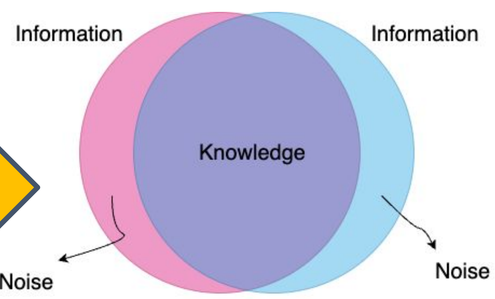
❖ Dynamic Knowledge Graphs facilitate forming relationships,
- ➢ increasing knowledge.

❖ New knowledge inferred from KG
- ➢ Can be stored back into the knowledge base

❖ Uncertainty quantification can be determined from hundreds of results



❖ **Information Theory** can further improve reasoning & accumulation of knowledge.

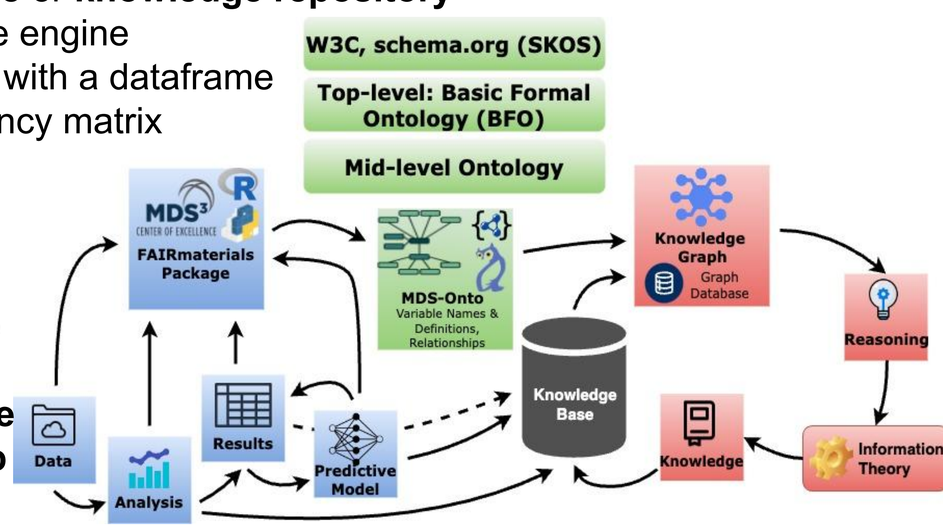


CONCLUSIONS

❖ To build a knowledge graph, these are the essential components:
- ➢ **FAIR** (**F**indable, **A**ccessible, **I**nteroperable, **R**eusable) data practices ensure proper assembly of large <u>datasets, analyses, models, & results</u>
- ➢ **Ontology** equips semantics for organization and management
- ➢ **Knowledge base** or **knowledge repository**
- ➢ A graph compute engine
  - ■ For computing with a dataframe and an adjacency matrix

❖ Knowledge base, ontology, & FAIR are instrumental

❖ To **Data Governance** & **Data Stewardship**



References

[1] James V. Stone, Information Theory: A Tutorial Introduction, 2015, 1st Ed., Sebtel Press

[2] Alessandro Negro, Graph Powered Machine Learning, 2021, Manning Publications Co.

[3] Ehrlinger, L., & Wöß, W. (n.d.). *Towards a Definition of Knowledge Graphs*.

[4] Paulheim, H. (n.d.). *Automatic Knowledge Graph Refinement: A Survey of Approaches and Evaluation Methods*.

Acknowledgement

This material is based upon research in the Materials Data Science for Stockpile Stewardship Center of Excellence (MDS3-COE), and supported by the U.S. Department of Energy's National Nuclear Security Administration under Award Number(s) DE-NA0004104. Work supported by the Laboratory Directed Research and Development program at Sandia National Laboratories, a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy’s National Nuclear Security Administration under contract DE-NA0003525. This work made use of the High Performance Computing Resource in the Core Facility for Advanced Research Computing at Case Western Reserve University.