Automatic Knowledge Discovery and Evolution: An Ontology for supporting Knowledge Discovery and Evolution

Introduction/Problem

Considering Knowledge Discovery and Evolution (KDE) from Philosophy of Science, Haig recently proposed the Abductive Theory of method (ATOM) that uses abduction for generating theories to explain phenomena [1]. We analyse ATOM and align it with two perspectives; Knowledge Representation and Reasoning which focuses on theory construction and Machine Learning and Data mining which are two different areas that are grouped together in this context because they are both data driven and offer support for phenomenon detection from data. These perspectives not only have different research cultures and practices which makes collaboration and interaction difficult but also use terminologies in different ways which makes it hard to work with the two perspectives in tandem; for instance in agent based KDE. We propose an algorithm and a KDE ontology that harmonises vocabulary for supporting agent based knowledge discovery and evolution based on ATOM.

The Abductive Theory of Method

Algorithm 1: Basic algorithm for the abductive theory of method (ATOM)

```
input: Data D
output: best explanatory theory t

procedure detectPhenomena(D):
perform initial data analysis on D to assess data quality
repeat until phenomena detected
suggest pattern using exploratory data analysis
confirm pattern through close replication e.g. cross validation
generalize pattern through constructive replication
if stable pattern found
p ← generalised pattern;
end repeat
return p

procedure constructTheory(p):
generate plausable theories T
develop theories using analogical modeling
assess and rank competing theories
t ← best theory as explanation for p;
return t

main:
p=detectPhenomena(D)
t=constructTheory(p)
return t
```

Example Competency Questions

CQ1(What theories exists for [phenomenon]?)

```
SELECT DISTINCT ?theory
WHERE { ?theory kdeontology:was_Influenced_By
kdeontology:h's consumption aligns more closely with C2 than C1}
```

CQ2 (From what data was a given pattern detected?)

```
SELECT DISTINCT ?Data
WHERE  kdeontology:h's consumption aligns more closely with C2 than C1}
kdeontology:was_detected_from
?Data
```

Conclusion

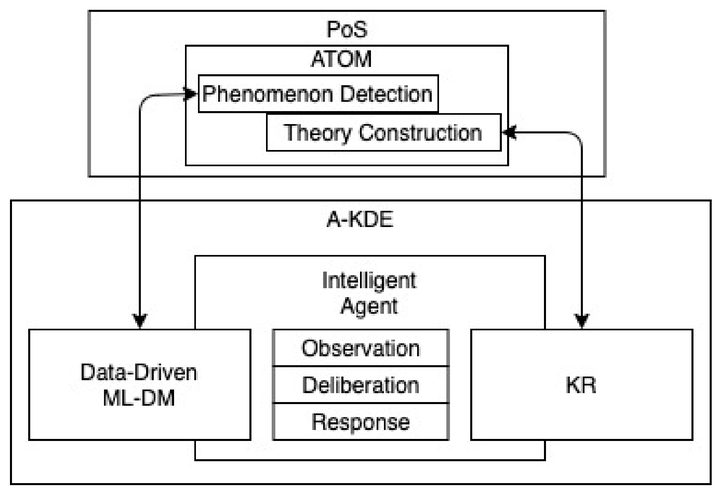
The proposed KDE ontology harmonises vocabulary that would be used by an agent based system that applies ML-DM and KR tools and techniques in tandem to a given use case in order to detect phenomena and construct explanatory theories for the phenomena. This would enable the representation and communication of generated knowledge.

Tezira Wanyana*
Supervised by: A/Prof. Deshen Moodley

Department of Computer Science,
University of Cape Town

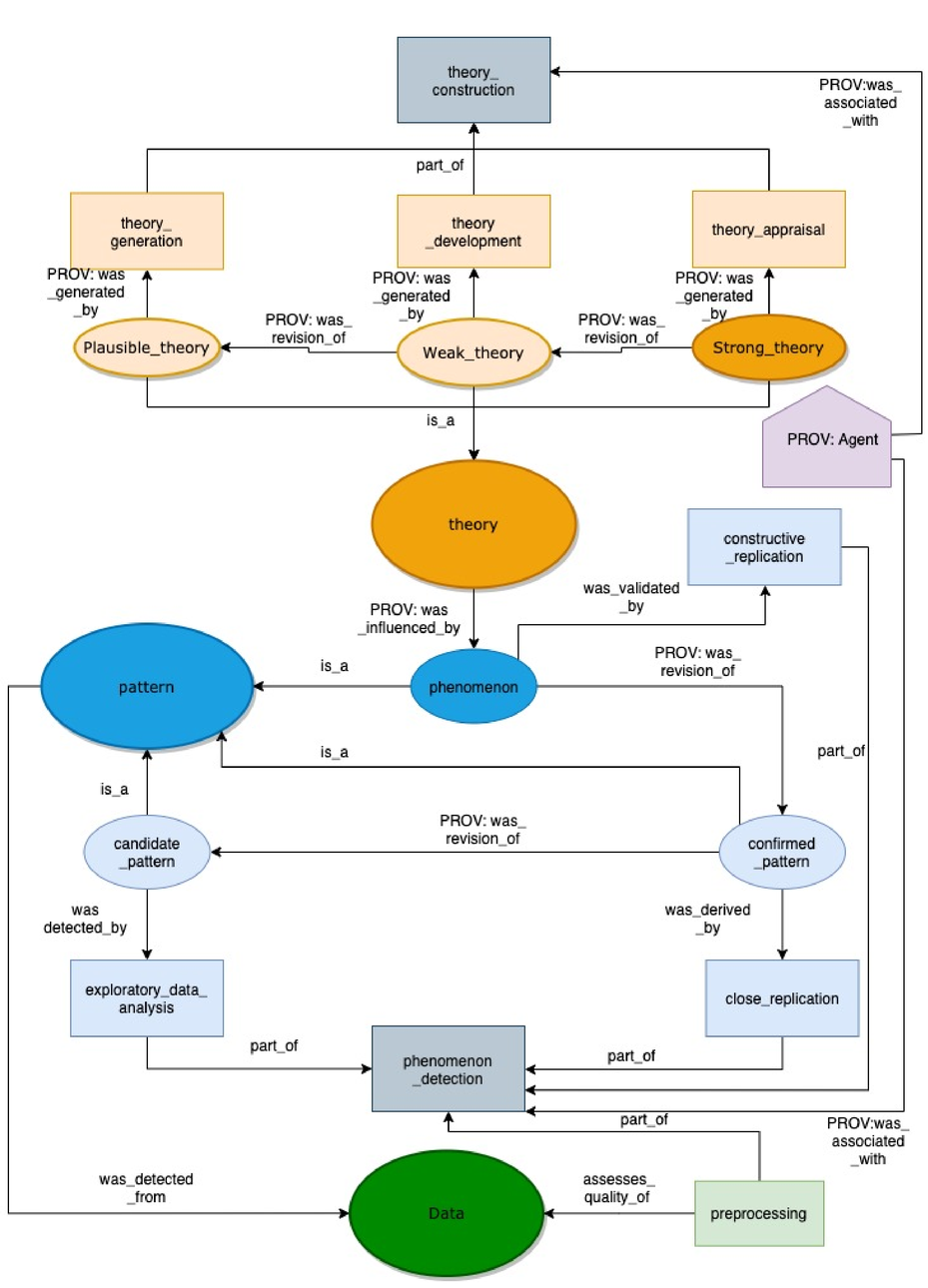
*twanyana@cs.uct.ac.za

A Unified Conceptual Model for KDE



KDE ontology - overview

The KDE ontology was designed using slight variation of the UPON methodology [2]. It is aligned to the W3C PROV standard [3]



https://sourceforge.net/projects/akde/

References

1. B. D. Haig, in Method Matters in Psychology (Springer, 2018), S. 35–64.
2. A. De Nicola u. a., in International Conference on Database and Expert Systems Applications, hrsg. von K. V. Andersen u. a., S. 655–664.
3. T. Lebo u. a., W3C recommendation **30** (2013).

Research School University of Cape Town
Powered by Hasso-Plattner-Institut
HPI
cair CENTRE FOR ARTIFICIAL INTELLIGENCE RESEARCH
SCHOOL OF IT