Process Representation Meets Operational Realization:

An Architecture for Data-Driven Process Ontology Application Through Process Mining

UNIVERSITY OF TORONTO

Riley Moher & Michael Gruninger

12.09.2025

FOIS 2025

Semantic Technologies Lab

Processes and Data

- Business processes drive analysis and optimization

Processes and Data

- Business processes drive analysis and optimization
- Process mining is the primary analysis engine

Process Mining

Timestamp	Event	Patient
12:02	Patient Intake	John Smith
12:05	Patient Intake	John Smith
12:06	Diagnostic	John Smith

- Issue: Patient Intake should only occur once for the same patient.

Process Mining

Timestamp	Event	Patient
12:02	Patient Intake	John Smith
12:05	Patient Intake	John Smith
12:06	Diagnostic	John Smith

- **Issue**: Patient Intake should only occur once for the same patient.
- **Solution**: The event label was wrong

Process Mining

Timestamp	Event	Patient
12:02	Patient Intake	John Smith
12:05	Additional Screening	John Smith
12:06	Diagnostic	John Smith

- **Issue:** Patient Intake should only occur once for the same patient.
- **Solution:** The event label was wrong

Process Mining

Timestamp	Event	Patient
12:02	Patient Intake	John Smith
12:05	Additional Screening	John Smith
12:06	Diagnostic	John Smith

- **Issue**: Patient Intake should only occur once for the same patient.
- **Solution**: The event label was wrong

Process mining relies on process knowledge, at foundational and domain-specific levels

Process Mining: Events

- “Awaiting Assignment”
- “Document Under Review“
- “Review Document”
- “Wait User Status”

Process Mining: Events

- “Awaiting Assignment”
- “Document Under Review“
- “Review Document”
- “Wait User Status”

These are all simply “events” in an event log

Process Mining: Events

- “Awaiting Assignment”
- “Document Under Review“
- “Review Document”
- “Wait User Status”

These are all simply “events” in an event log

But they have very different process implications

Process Mining: Events

- “Awaiting Assignment”
- “Document Under Review“
- “Review Document”
- “Wait User Status”

These are all simply “events” in an event log

But they have very different process implications

Events are heavily overloaded and require interpretation for analysis

Process Mining

Process Data
(Event Logs)

Business Rules,
Process Knowledge

Cleaned Data,
Models,
Compliance Checks,
Insights

Process Mining

Process Data
(Event Logs)

Business Rules,
Process Knowledge

Cleaned Data,
Models,
Compliance Checks,
Insights

1. How does a process ontology fit?

Process Mining

Process Data (Event Logs)

Business Rules, Process Knowledge

Cleaned Data, Models, Compliance Checks, Insights

1. How does a process ontology fit?
2. What tools enable that fit?

Process Mining

Process Data
(Event Logs)

Business Rules,
Process Knowledge

Cleaned Data,
Models,
Compliance Checks,
Insights

1. How does a process ontology fit?
2. What tools enable that fit?
3. What is the impact to process ontologies?

Agenda

1. How does a process ontology fit?
2. What tools enable that fit?
3. What is the impact to process ontologies?

Agenda

1. How does a process ontology fit?
2. What tools enable that fit?
3. What is the impact to process ontologies?

Process Ontology Application: Challenges

- Traditional approaches focus on representation or access (OBDA), not data-driven usage

Process Ontology Application: Challenges

- Traditional approaches focus on representation or access (OBDA), not data-driven usage

- The ontology should be the engine of operational analysis

Process Ontology Application: Challenges

- Traditional approaches focus on representation or access (OBDA), not data-driven usage
- The ontology should be the engine of operational analysis
- Operational realization captures this notion

Operational Realization

- Different from representative, referential, or conceptual application

Operational Realization

- Different from representative, referential, or conceptual application
- Application of the ontology is rooted in datasets and problems in their analysis and interpretation

Achieving Operational Realization

- Process mining uses heterogenous data

Achieving Operational Realization

- Process mining uses heterogenous data
- Integration with existing data is ambiguous

Achieving Operational Realization

- Process mining uses heterogenous data
- Integration with existing data is ambiguous
- How do we represent and reason with different kinds of process knowledge and data?

Agenda

1. How does a process ontology fit?
2. What tools enable that fit?
3. What is the impact to process ontologies?

Ontology-Driven Process Mining

Raw Event Data	Knowledge Base
Ad-Hoc Interpretations	Data Theory
Business Questions	Logical Queries
Business Validations	Logical Proofs

Ontology-Driven Process Mining

Raw Event Data

Ad-Hoc Interpretations

Business Questions

Business Validations

Map

Formalize

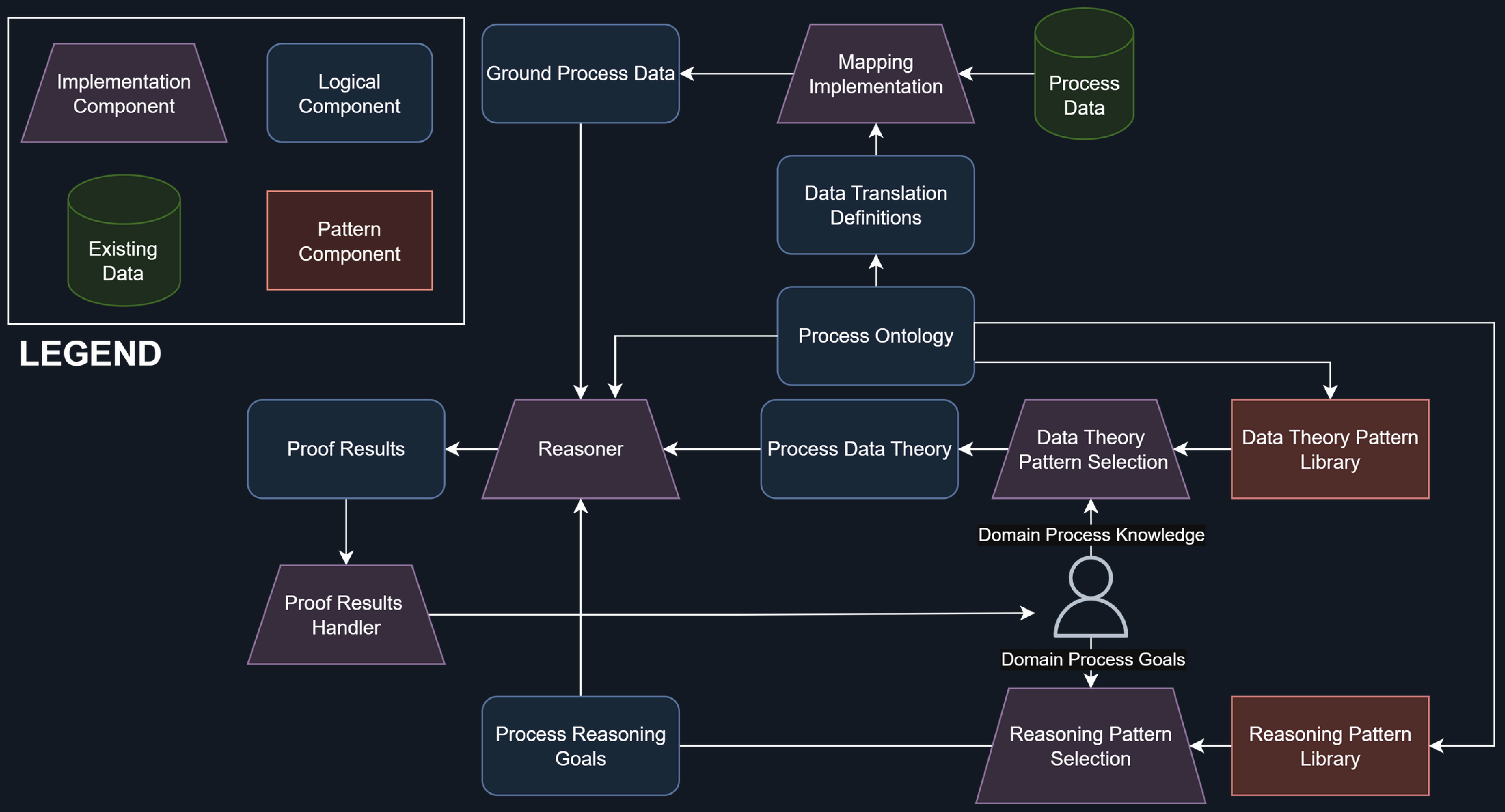
Analyze

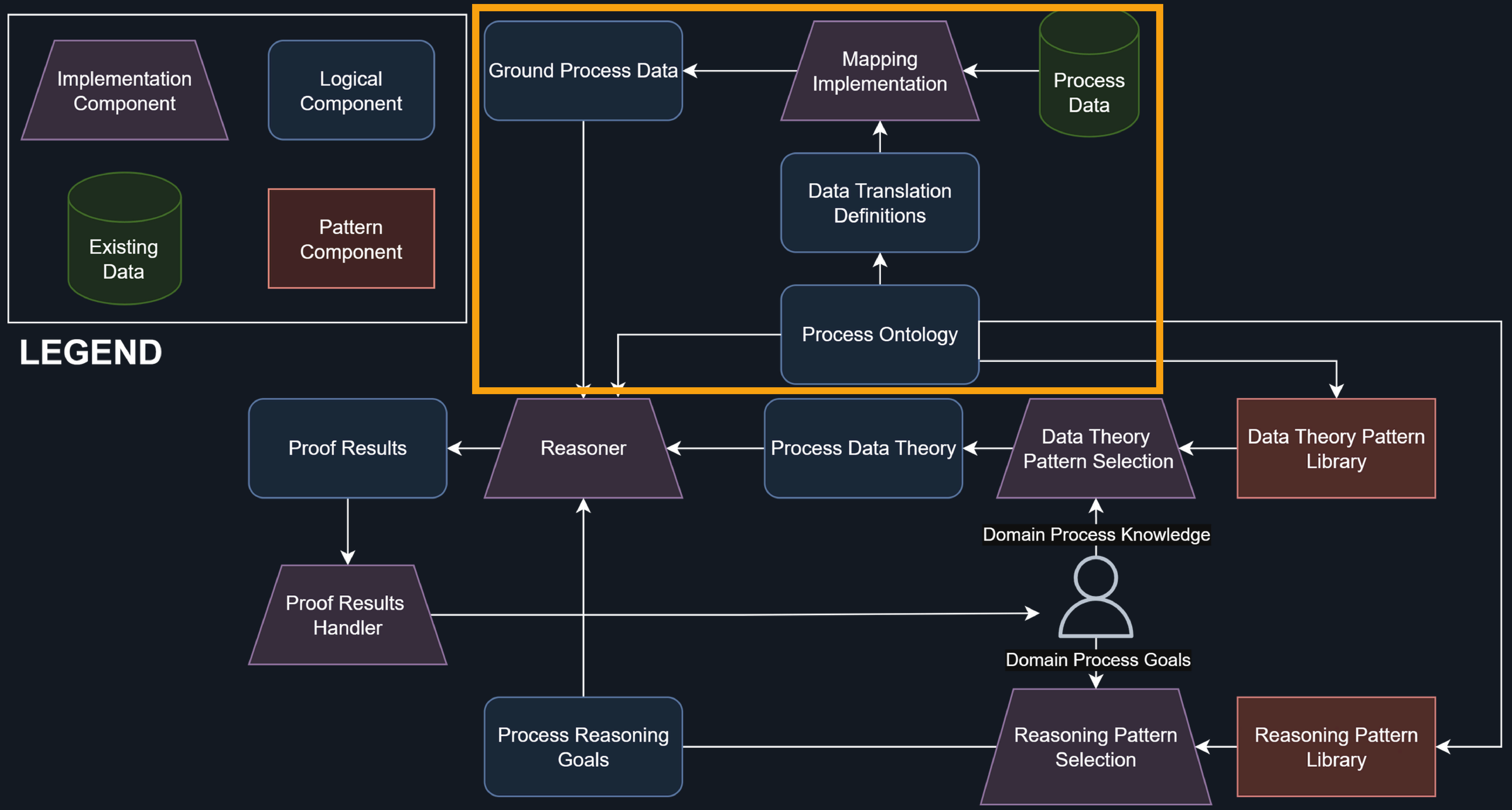
Knowledge Base

Data Theory

Logical Queries

Logical Proofs


Implementation Component
Logical Component
Existing Data
Pattern Component
LEGEND
Ground Process Data
Mapping Implementation
Process Data
Data Translation Definitions
Process Ontology
Proof Results
Reasoner
Process Data Theory
Data Theory Pattern Selection
Data Theory Pattern Library
Domain Process Knowledge
Proof Results Handler
Domain Process Goals
Process Reasoning Goals
Reasoning Pattern Selection
Reasoning Pattern Library


Implementation Component
Logical Component
Existing Data
Pattern Component
LEGEND
Ground Process Data
Mapping Implementation
Process Data
Data Translation Definitions
Process Ontology
Proof Results
Reasoner
Process Data Theory
Data Theory Pattern Selection
Data Theory Pattern Library
Domain Process Knowledge
Proof Results Handler
Domain Process Goals
Process Reasoning Goals
Reasoning Pattern Selection
Reasoning Pattern Library

Mapping Event Log Data

- Tabular data from event logs becomes ground facts

Timestamp	Event	Lifecycle Transition
14:20	Process Application	Start
15:30	Credit Check	Complete
15:38	Credit Check	Start

Mapping Event Log Data

- Tabular data from event logs becomes ground facts

Timestamp	Event	Lifecycle Transition
14:20	Process Application	Start
15:30	Credit Check	Complete
15:38	Credit Check	Start

Event(e2)
hasActivity(e0, creditCheck)
hasTransition(e2, complete)

Mapping Event Log Data

- Tabular data from event logs becomes ground facts

Timestamp	Event	Lifecycle Transition
14:20	Process Application	Start
15:30	Credit Check	Complete
15:38	Credit Check	Start

Event(e2)
hasActivity(e0, creditCheck)
hasTransition(e2, complete)

Ground Data

Event(e3)
hasActivity(e3, creditCheck)
hasTransition(e3, start)

Mapping Event Log Data

- Tabular data from event logs becomes ground facts (via RML mappings)

Timestamp	Event	Lifecycle Transition
14:20	Process Application	Start
15:30	Credit Check	Complete
15:38	Credit Check	Start

Event(e2)
hasActivity(e0, creditCheck)
hasTransition(e2, complete)

Ground Data

Event(e3)
hasActivity(e3, creditCheck)
hasTransition(e3, start)

Mapping Event Log Data

Ground Data

Event(e2)
hasActivity(e0, creditCheck)
hasTransition(e2, complete)

Event(e3)
hasActivity(e3, creditCheck)
hasTransition(e3, start)

Mapping Event Log Data

Ground Data

Event(e2)
hasActivity(e0, creditCheck)
hasTransition(e2, complete)

Event(e3)
hasActivity(e3, creditCheck)
hasTransition(e3, start)

Ontology Translations

"Two events sharing an activity each with a start and end transition indicates an acitivty occurrence"

$$\forall e_s \, \forall e_e \, \forall a \, \Big(\texttt{Event}(e_s) \wedge \texttt{Event}(e_e) \wedge$$
$$\texttt{hasActivity}(e_s, a) \wedge \texttt{hasActivity}(e_e, a) \wedge$$
$$\texttt{hasTransition}(e_s, \mathsf{start}) \wedge \texttt{hasTransition}(e_e, \mathsf{complete}) \rightarrow$$
$$\exists o \, \big(\texttt{activity\_occurrence}(o) \wedge$$
$$\texttt{occurrence\_of}(o, a) \wedge$$
$$\texttt{beginOf}(o, e_s) \wedge$$
$$\texttt{endOf}(o, e_e) \big) \Big).$$

Mapping Event Log Data

Ground Data

Event(e2)
hasActivity(e0, creditCheck)
hasTransition(e2, complete)

Event(e3)
hasActivity(e3, creditCheck)
hasTransition(e3, start)

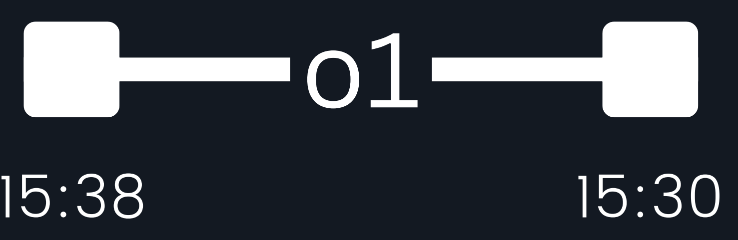
Ontology Translations

"Two events sharing an activity each with a start and end transition indicates an acitivty occurrence"

$$\forall e_s \forall e_e \forall a \Big(\texttt{Event}(e_s) \land \texttt{Event}(e_e) \land$$
$$\texttt{hasActivity}(e_s, a) \land \texttt{hasActivity}(e_e, a) \land$$
$$\texttt{hasTransition}(e_s, \text{start}) \land \texttt{hasTransition}(e_e, \text{complete}) \rightarrow$$
$$\exists o \; (\texttt{activity\_occurrence}(o) \land$$
$$\texttt{occurrence\_of}(o, a) \land$$
$$\texttt{beginOf}(o, e_s) \land$$
$$\texttt{endOf}(o, e_e))\Big).$$

Translated Data

occurrence(o1)
beginOf(o1, 15:38)
endOf(o1, 15:30)



Mapping Event Log Data

Ground Data

Event(e2)
hasActivity(e0, creditCheck)
hasTransition(e2, complete)

Event(e3)
hasActivity(e3, creditCheck)
hasTransition(e3, start)

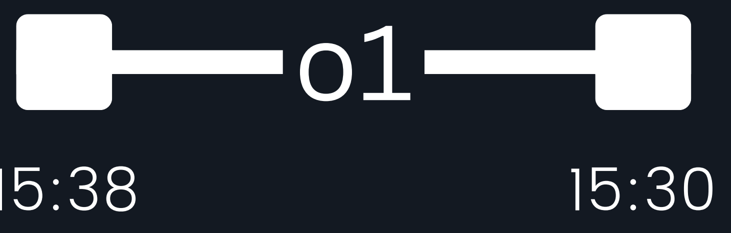
Ontology Translations

“Two events sharing an activity each with a start and end transition indicates an acitivty occurrence”

$$\forall e_s \, \forall e_e \, \forall a \, \Big(\texttt{Event}(e_s) \wedge \texttt{Event}(e_e) \wedge \\ \texttt{hasActivity}(e_s, a) \wedge \texttt{hasActivity}(e_e, a) \wedge \\ \texttt{hasTransition}(e_s, \text{start}) \wedge \texttt{hasTransition}(e_e, \text{complete}) \rightarrow \\ \exists o \, (\texttt{activity\_occurrence}(o) \wedge \\ \texttt{occurrence\_of}(o, a) \wedge \\ \texttt{beginOf}(o, e_s) \wedge \\ \texttt{endOf}(o, e_e)) \Big).$$

Translated Data

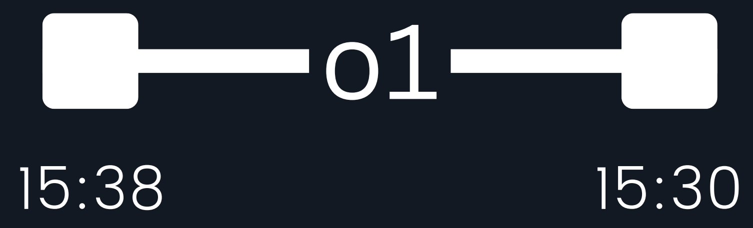
occurrence(o1)
beginOf(o1, 15:38)
endOf(o1, 15:30)



Event Log Data Quality

Translated Data

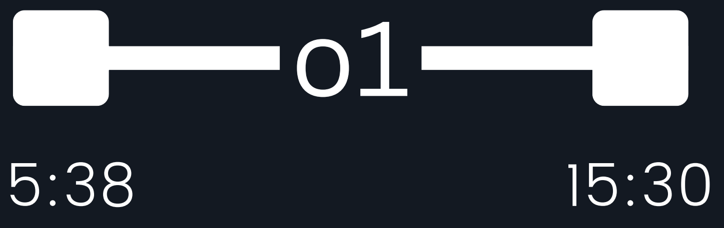
occurrence(o1)
beginOf(o1, 15:38)
endOf(o1, 15:30)



Event Log Data Quality

Translated Data

occurrence(o1)
beginOf(o1, 15:38)
endOf(o1, 15:30)



Ontology

"Activity occurrences start points are less than or equal to their end points"

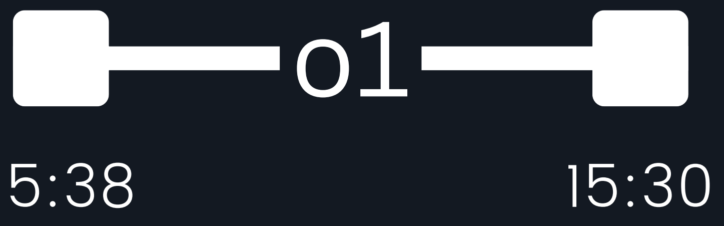
$$(\forall o\, \mathrm{activity\_occurrence}(o) \implies \exists t_1, t_2\, (\mathrm{begin\_of}(o) = t_1 \wedge \mathrm{end\_of}(o) = t_2) \wedge t_1 \leq t_2))$$

$$t_1 \leq t_2$$

Event Log Data Quality

Translated Data

occurrence(o1)
beginOf(o1, 15:38)
endOf(o1, 15:30)



Ontology

"Activity occurrences start points are less than or equal to their end points"

$$(\forall o \, \text{activity\_occurrence}(o) \implies \exists t_1, t_2 \, (\text{begin\_of}(o) = t_1 \wedge \text{end\_of}(o) = t_2) \wedge t_1 \leq t_2))$$

$$t_1 \leq t_2$$

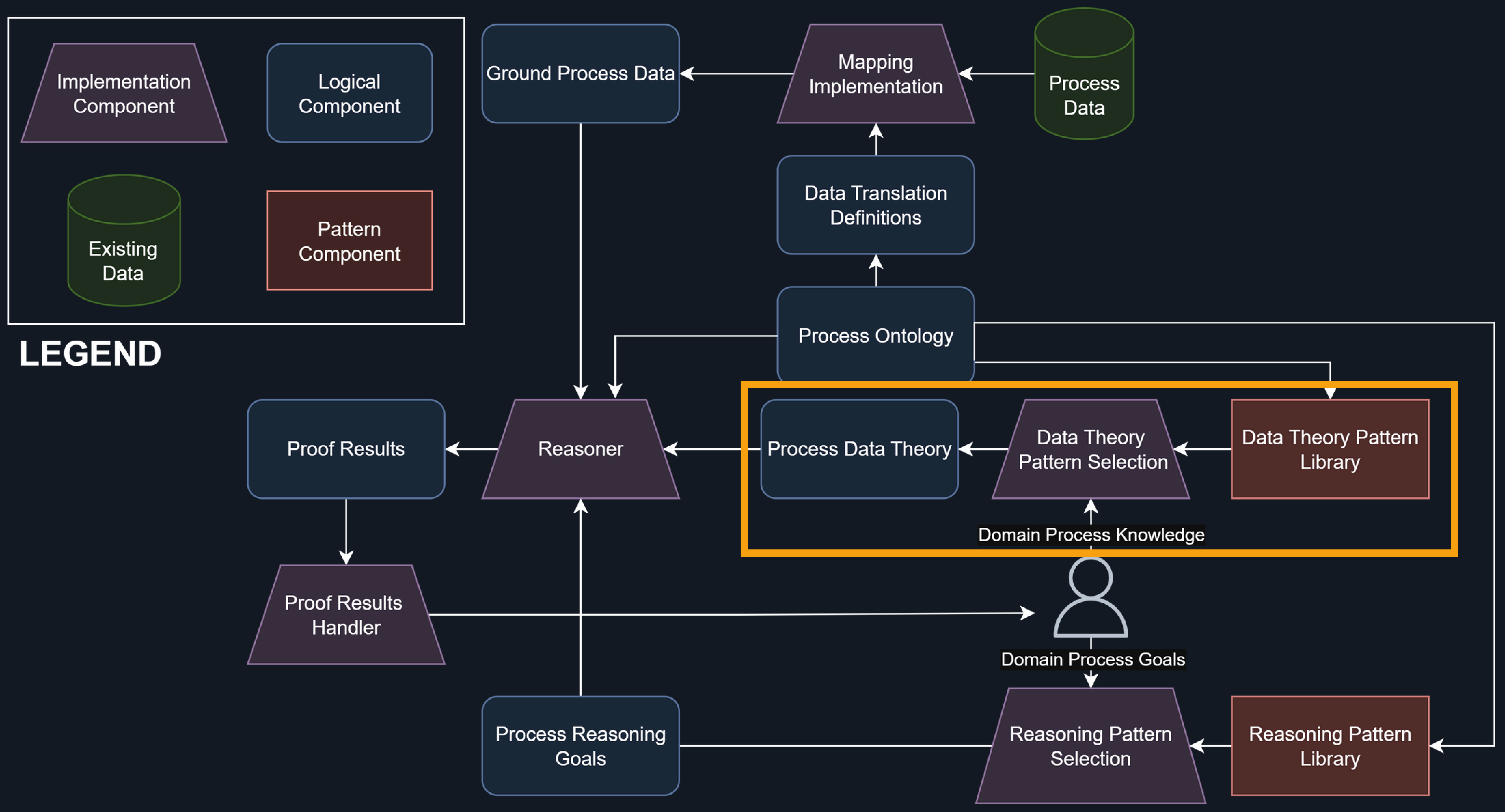
Proof of Inconsistency

$$t_1 = 15:38$$

$$t_2 = 15:30$$

$$\boxed{\begin{matrix} t_1 > t_2 \\ t_1 \leq t_2 \end{matrix}}$$

What about domain-specific analysis?


Implementation Component
Logical Component
Existing Data
Pattern Component
LEGEND
Ground Process Data
Mapping Implementation
Process Data
Data Translation Definitions
Process Ontology
Proof Results
Reasoner
Process Data Theory
Data Theory Pattern Selection
Data Theory Pattern Library
Domain Process Knowledge
Proof Results Handler
Domain Process Goals
Process Reasoning Goals
Reasoning Pattern Selection
Reasoning Pattern Library

Knowledge Patterns

- “When a fragile object is dropped, it breaks”
- “Patient Intake should only occur once for the same patient.”

Knowledge Patterns

- “When a fragile object is dropped, it breaks”
- “Patient Intake should only occur once for the same patient.”

Knowledge Patterns

- “When a fragile object is dropped, it breaks”

State Based Effect

- “Patient Intake should only occur once for the same patient.”

Occurrence Constraint

Knowledge Patterns

- “When a fragile object is dropped, it breaks”

State Based Effect

- “Patient Intake should only occur once for the same patient.”

Occurrence Constraint

Knowledge Patterns

- State-Based Effects (SBE)
- “When a fragile object is dropped, it breaks”
- While some initial condition holds, and an occurrence happens, some resulting condition holds afterwards

Knowledge Patterns

- State-Based Effects (SBE)
- "When a fragile object is dropped, it breaks"
- While some initial condition holds, and an occurrence happens, some resulting condition holds afterwards
- Patterns abstract common process knowledge
- SBE(c1, a, c2)

Knowledge Patterns

- State-Based Effects (SBE)

$$SBE(a, f_1, f_2) := (\forall o)\ occurrence\_of(o, a) \wedge prior(f_1, o) \implies holds(f_2, o)$$

$$(\forall o)\ occurrence\_of(o, drop) \wedge prior(fragile, o) \implies holds(broken, o)$$

Knowledge Patterns

- State-Based Effects (SBE)

$$SBE(a, f_1, f_2) := (\forall o)\ occurrence\_of(o, a) \land prior(f_1, o) \implies holds(f_2, o)$$

$$(\forall o)\ occurrence\_of(o, drop) \land prior(fragile, o) \implies holds(broken, o)$$

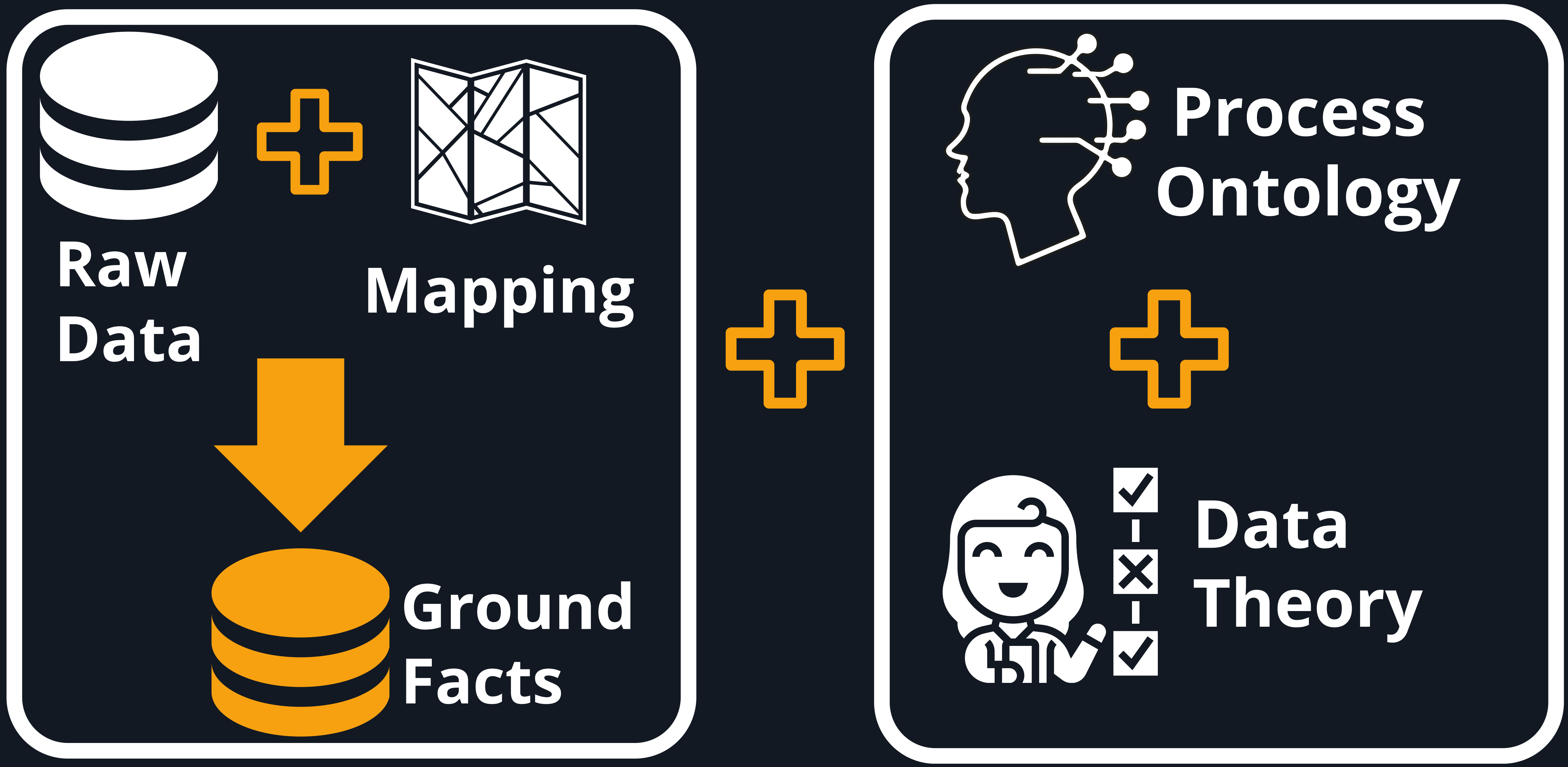
T-Box?

A-Box?

Domain Ontology?

Process Ontology

Data-Driven Process Ontology



Agenda

1. How does a process ontology fit?
2. What tools enable that fit?
3. What is the impact to process ontologies?

Process Ontology Impacts

- Clarifies how ontologies interact with enterprise data and systems
- Introduces data-driven approaches to benchmarking process ontologies
- Fosters engagement with the process mining community and its methods

Takeaways

- Operational realization: a novel paradigm for applying ontologies to process data
- Data theory: a new structure for domain-level, data-dependent knowledge expressed in the language of upper ontologies
- Demonstrated architecture as a methodological bridge for process ontology–data integration

Contact / More Details