Ontology engineering approach to support process of model and data integration

Linsey Koo, Edlira Kalemi, Franjo Cecelja

Background

- Increasing numbers of modelling methods using heterogeneous tools
  - Models remain implicit to whom built them
  - Limits the potential of reusability
  - Time consuming & redundant work

- Lack of complete libraries of bio-chemical processes

  ➢ To retain the valuable models and data in biorefining
    ➢ Systematic approach to identifying, capturing, retrieving, sharing and effectively reusing these models and data

  ➢ Need to build new models or integrate existing ones
Existing Framework: CAPE-OPEN

- **Concept**
  - Standardisation of interfaces to enable interoperability between simulator software components from different sources
  - Integration of models and tools to take advantage of characteristics that vary between simulation environments

- **Middleware Architecture**

- Widely **supported** by existing software
  - Rely on user intervention
  - Limits possibility of reusing existing models

Message from CO-LaN - the CAPE-OPEN Laboratories Network:
‘... speaking **negatively** on CAPE-OPEN ...’

We **benefit** from CAPE-OPEN achievements **greatly** and only propose to build upon it:
- Introducing higher level of **flexibility** in model/data integration by introducing knowledge modelling - ontology (repository)
- To assist user decision by introducing **partial matching** and relevance **ranking**
- Benefitting from well-proven (service) **integration** technologies such as SWS
Ontology Engineering Approach

- Build on previous work – CAPE-OPEN Framework
- Semantic model/data integration

Model Registration
- Web Based Repository
- Parsing Ontology

Request Formulation
- User specification
- Input specification

Model Discovery
- Input-Output Matching
- Partial Matching

Selection
- Match ranking
- Best match to the request

Composition (orchestration) & Model Integration
- Ontology Web Language for Services Framework

Model Representation/Description

Generic model representation with semantic description using ontologies: the semantic web service (SWS) description in OWL-S framework

Model
- Model Profile
- Model Discovery
- Model Grounding
- Model Invocation
- Model Operation
- Model Composition & Interoperation

Model Preconditions (environment)
- Inputs
- Outputs (effects)

Model Operation:
- Composition of descriptions from multiple elements to perform a specific task.
- Coordination of activities of different participants.
- Monitoring the execution.
Data Representation/Description

Data uses similar principle to model representation and SWS description:

InterCAPEmodel Ontology

- **Knowledge Representation** in the domain of biorefining
  - Expand knowledge of process models & data
  - Provide a classification & characterisation of models & data
  - Derive implicit information through the analysis of explicit knowledge

- **Classification** of Models
  - Functionality, (biorefining) Platform, Characteristics, Input & Output

- **Inputs & Outputs** of Models
  - No. of Inputs & Outputs
  - Type of Inputs & Outputs (i.e. material, energy, etc.)
  - Parameters of Inputs & Outputs
# InterCAPEmodel Ontology

## Classification of Models

<table>
<thead>
<tr>
<th>Modelling Functionality</th>
<th>ModellingScope</th>
<th>ModellingCharacteristics</th>
<th>ModellingInputType</th>
<th>ModellingOutputType</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reaction</td>
<td>SimulatePlatform</td>
<td>ReactionSimulation</td>
<td>FeedstockType</td>
<td>ProductType</td>
</tr>
<tr>
<td>BiochemicalReaction</td>
<td>BiochemicalPlatform</td>
<td>BiochemicalSimulation</td>
<td>VirginResource</td>
<td>BiochemicalProduct</td>
</tr>
<tr>
<td>ChemicalReaction</td>
<td>ChemicalPlatform</td>
<td>ChemicalSimulation</td>
<td>WasteResource</td>
<td>Biofuel</td>
</tr>
<tr>
<td>Cooling</td>
<td>CoolingPlatform</td>
<td>CoolingSimulation</td>
<td>PrimaryResidue</td>
<td>CommunicationSector</td>
</tr>
<tr>
<td>PressureChanger</td>
<td>PressurePlatform</td>
<td>PressureSimulation</td>
<td>ChemicalComponent</td>
<td>HealthAndHygieneSector</td>
</tr>
<tr>
<td>IncreasePressure</td>
<td>IncreasePressurePlatform</td>
<td>IncreasePressureSimulation</td>
<td>EnergyInput</td>
<td>Heating</td>
</tr>
<tr>
<td>DecreasePressure</td>
<td>DecreasePressurePlatform</td>
<td>DecreasePressureSimulation</td>
<td>Cooling</td>
<td>Sector</td>
</tr>
</tbody>
</table>

### Inputs & Outputs of Model (as data/object properties and as ontology entities)

<table>
<thead>
<tr>
<th>Input</th>
<th>Output</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>hasInput</td>
<td>hasOutput</td>
<td>Define direction of flow</td>
</tr>
<tr>
<td>hasNumberOfInputs</td>
<td>hasNumberOfOutputs</td>
<td>Define number of ports required for the model by type of inputs and outputs</td>
</tr>
<tr>
<td>hasNumberOfMaterialInputs</td>
<td>hasNumberOfMaterialOutputs</td>
<td>Define number of material inputs and outputs for the model</td>
</tr>
<tr>
<td>hasMaterialInputs</td>
<td>hasMaterialOutputs</td>
<td>Define value of material composition for each stream</td>
</tr>
<tr>
<td>hasEnergyInputs</td>
<td>hasEnergyOutputs</td>
<td>Define value of energy composition for each stream</td>
</tr>
<tr>
<td>hasInputParameters</td>
<td>hasOutputParameters</td>
<td>Define parameters of input/output and set values for each parameter in SI units</td>
</tr>
<tr>
<td>hasInputFluxrate</td>
<td>hasOutputFluxrate</td>
<td></td>
</tr>
<tr>
<td>hasMassFluxrate</td>
<td>hasMassFluxrate</td>
<td></td>
</tr>
<tr>
<td>hasMolarFluxrate</td>
<td>hasMolarFluxrate</td>
<td></td>
</tr>
<tr>
<td>hasVolumetricFluxrate</td>
<td>hasVolumetricFluxrate</td>
<td></td>
</tr>
<tr>
<td>hasPhaseFraction</td>
<td>hasPhaseFraction</td>
<td></td>
</tr>
<tr>
<td>hasTemperature</td>
<td>hasTemperature</td>
<td></td>
</tr>
<tr>
<td>hasPressure</td>
<td>hasPressure</td>
<td></td>
</tr>
</tbody>
</table>
Registration path is created by parsing ontology – subsumption and object properties, inferred ontology

Incorporate explicit knowledge of each model/data in public repository(ies)

Increase share & reuse of existing models/datasets

Model Reg. by Ontology Parsing

Fatty Acid Methyl Esters (FAME) Purification Model

Aspen Plus Software
Restriction (Material): Presence of Glycerol

Material Input 1
Compositions:
• FAME
• Methanol
Parameters:
• Concentration, x
• Flowrate, m
• Temperature, t
• Pressure, p

Model Description
• Functionality: Separation
• Equipment Type: Distillation column with 5 equilibrium stages
• Scale: Functional Process Unit
• Model Method: NRTL
• Model Type: Simulation
• Purity of product: 99.8%
• Number of Input Data: 1
• Number of Output Data: 2

Material Output 1
Composition
• FAME
• Methanol
 Parameters
• x, m, t, p

Material Output 2
Composition
• Methanol
 Parameters
• m, t, p

Q) Functionality of the model?
→ Choose from available options

→ Associated equipment type for separation

→ # of equilibrium stages (data property)
Registration of Model’s Input & Output

Precondition & Restriction

Precondition \(\rightarrow\) modelling software \(\rightarrow\) AspenPlus

- This definition states FAME Purification model is a model with an input that is not Glycerol. Thus any instance of this class must have any inputs that of any class, except Glycerol.
Model Discovery

- Process of finding a candidate model/dataset which is to be integrated with the requesting model
  - Fully or partially satisfy the requested functionality and the output properties
    - **Functionality**: distance measurement matching & similarity measure
    - **Input/Output property**: data property matching & similarity measure
  - Requested parameter may come from different models or data, or more than one model or data

Model Selection

- Matching process
  - Process of reducing redundant matching to avoid performance deficiency
  - Key components required for the input of the requesting model is considered as elimination criterion
  - Still refining elimination process
  - Process of quantifying the semantic relevance between the requesting model and models residing in the repository
    - **Distance measurement** between respective concepts representing tacit knowledge which is measured along the hierarchical and object property relationship in the ontology
    - **Property similarity** that calculates values of properties that characterise explicit knowledge in the form of vectors
  - Performance Ranking
    - Based on the level of match, the discovered models that fully or partially satisfy matching conditions are ranked
Input/Output Matching

- Graphical Method
- Functionality & Input/Output Type
- Process Synthesis Logic
- Property Matching
- Explicit Knowledge
- Cosine Similarity
  \[ h^{V,C}_k = \frac{\mathbf{p}_r \cdot \mathbf{p}_l}{\|\mathbf{p}_r\| \|\mathbf{p}_l\|} = \frac{\sum_{i=1}^{n} p_{r,i} \times p_{l,i}}{\sqrt{\sum_{i=1}^{n} (p_{r,i})^2} \times \sqrt{\sum_{i=1}^{n} (p_{l,i})^2}} \]
- Euclidean Similarity
  \[ h^{V,E}_k = 1 - \frac{\sqrt{\sum_{i=1}^{n} (p_{r,i} - p_{l,i})^2}}{\max \sum_{i=1}^{n} (p_{r,i} - p_{l,i})^2} \]
- Aggregated Similarity
  \[ h^V_k = \frac{h^{V,C}_k + h^{V,E}_k}{2} \]

➢ Partial Matching Allowed

Model Integration

- Candidate models and data are ranked based on semantic relevance
- Best matches that satisfy the requestor's functionality and output property are proposed – the user makes decision
- Enable a (semantic) flexible and user customised model integration fully coordinated by SWS

### Demonstration

**Process Model**

<table>
<thead>
<tr>
<th>Process Model</th>
<th># Input</th>
<th>Required Input Components</th>
<th>FAME %</th>
<th># Output</th>
<th>Main Output Components</th>
<th>FAME %</th>
</tr>
</thead>
<tbody>
<tr>
<td>FAME Purification</td>
<td>1</td>
<td>S1. FAME, Oil</td>
<td>.852 – .969</td>
<td>2</td>
<td>S1. FAME</td>
<td>.997 - .998</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>S2. Oil</td>
<td></td>
</tr>
<tr>
<td>Hexane Extraction</td>
<td>2</td>
<td>S1. FAME, Glycerol</td>
<td>.830 - .852</td>
<td>2</td>
<td>S1. FAME</td>
<td>.852 – .969</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S2. Hexane</td>
<td></td>
<td></td>
<td>S2. Glycerol</td>
<td></td>
</tr>
<tr>
<td>Water Washing</td>
<td>2</td>
<td>S1. FAME, Glycerol</td>
<td>.830 - .852</td>
<td>2</td>
<td>S1. FAME</td>
<td>.852 – .969</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S2. Water</td>
<td></td>
<td></td>
<td>S2. Glycerol</td>
<td></td>
</tr>
<tr>
<td>Glycerol Separation</td>
<td>1</td>
<td>S1. FAME</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methanol Recovery</td>
<td>1</td>
<td>S1. Methanol, FAME, Glycerol</td>
<td>.719 - .830</td>
<td>2</td>
<td>S1. Methanol (Recycle)</td>
<td>.719 – .830</td>
</tr>
<tr>
<td></td>
<td></td>
<td>S2. FAME, Glycerol</td>
<td></td>
<td></td>
<td>S2. FAME, Glycerol</td>
<td></td>
</tr>
<tr>
<td>Transesterification 1</td>
<td>3</td>
<td>S1. Waste Oil, Methanol</td>
<td>.000</td>
<td>2</td>
<td>S1. FAME, Glycerol, Methanol</td>
<td>.719 - .830</td>
</tr>
<tr>
<td>Transsesterification 2</td>
<td>3</td>
<td>S3. Catalyst: T1 - H₂SO₄</td>
<td></td>
<td></td>
<td>S2. Catalyst (Recycle)</td>
<td>.830</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T2 - H₂SO₄, NaOH</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transsesterification 3</td>
<td>3</td>
<td>S1. Virgin Oil, Methanol</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transsesterification 4</td>
<td>3</td>
<td>S3. Catalyst: T1 - NaOH,</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>T2 - Ca₅La₁</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Demonstration**

**Fatty acid methyl esters (FAME)**

**Purification (Requesting Model)**

- FAME + Methanol → FAME@.997
- Methanol

**Methanol Recovery**

- FAME, Methanol, Glycerol, Others
- Methanol → FAME + Glycerol

**Glycerol Separation**

- FAME + Glycerol → Hexane Extraction → Glycerol
- Water Washing → FAME
- Glycerol Separation

**Transesterification**

- Waste/Virgin Oil + Methanol + Catalyst
- Waste Oil + Acid Catalyst
- Virgin Oil + Alkal Catalyst
- FAME + Glycerol + Methanol → Catalyst
Demonstration

FAME Purification
(Requesting Model)

Methanol Recovery

Glycerol Separation

Transesterification

Demonstration: Backward Matching

Restriction to the FAME Purification: Presence of Glycerol

FAME + Methanol FAME@.997 Methanol

FAME, Methanol, Glycerol, Others

Methanol

FAME + Glycerol

Waste/Virgin Oil

Catalyst

FAME + Glycerol + Methanol

Catalyst

FAME, Methanol, Glycerol, Others

Methanol

FAME + Glycerol

Waste Oil

Acid Catalyst

Waste Oil

Acid + Alkali

Virgin Oil

Alkali Catalyst

Virgin Oil

Heterogeneous Catalyst

FAME + Methanol

FAME + Glycerol

Hexane Extraction

Glycerol

Water Washing

FAME

Glycerol Separation

Restriction to the FAME Purification: Presence of Glycerol

FAME + Methanol

FAME@.997

Methanol

FAME + Methanol

FAME + Glycerol

Hexane Extraction

Glycerol

Water Washing

FAME

Glycerol Separation

Restriction to the FAME Purification: Presence of Glycerol

FAME + Methanol

FAME@.997

Methanol

FAME + Methanol

FAME + Glycerol

Hexane Extraction

Glycerol

Water Washing

FAME

Glycerol Separation

Restriction to the FAME Purification: Presence of Glycerol

FAME + Methanol

FAME@.997

Methanol

FAME + Methanol

FAME + Glycerol

Hexane Extraction

Glycerol

Water Washing

FAME

Glycerol Separation

Restriction to the FAME Purification: Presence of Glycerol

FAME + Methanol

FAME@.997

Methanol

FAME + Methanol

FAME + Glycerol

Hexane Extraction

Glycerol

Water Washing

FAME

Glycerol Separation

Restriction to the FAME Purification: Presence of Glycerol

FAME + Methanol

FAME@.997

Methanol

FAME + Methanol

FAME + Glycerol

Hexane Extraction

Glycerol

Water Washing

FAME

Glycerol Separation

Restriction to the FAME Purification: Presence of Glycerol

FAME + Methanol

FAME@.997

Methanol

FAME + Methanol

FAME + Glycerol

Hexane Extraction

Glycerol

Water Washing

FAME

Glycerol Separation

Restriction to the FAME Purification: Presence of Glycerol

FAME + Methanol

FAME@.997

Methanol

FAME + Methanol

FAME + Glycerol

Hexane Extraction

Glycerol

Water Washing

FAME

Glycerol Separation

Restriction to the FAME Purification: Presence of Glycerol

FAME + Methanol

FAME@.997

Methanol

FAME + Methanol
Demonstration

- Biorefinery supply chain model

Inputs:
- Material Inputs
- Capacity
- Biorefinery Types
- Transportation Types
- Conversion Rates
- Geo Location
- Cost
- Demand
- Selling Price

Outputs:
- Profits
- Environmental Impact
- Material Outputs

Representation of supply chain network using taxonomy, attribute, and relation in InterCAPE Ontology
Demonstration

- **Key parameters** of Input/Output for supply chain

<table>
<thead>
<tr>
<th>Input</th>
<th>Output</th>
<th>Environment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cultivation</td>
<td>Material Input</td>
<td>Material Output</td>
</tr>
<tr>
<td></td>
<td>Yield of Biomass</td>
<td>Output Flowrate</td>
</tr>
<tr>
<td></td>
<td>Seasonality of Biomass</td>
<td>Geo. Location</td>
</tr>
<tr>
<td></td>
<td>Input Flowrate</td>
<td>Cost of Cultivation</td>
</tr>
<tr>
<td></td>
<td>Latent Configuration</td>
<td>Cost of Cultivation</td>
</tr>
<tr>
<td></td>
<td>Geo. Location</td>
<td>Unit Cost of Cultivation</td>
</tr>
<tr>
<td>Transportation</td>
<td>Material Input</td>
<td>Material Output</td>
</tr>
<tr>
<td></td>
<td>Max. Capacity</td>
<td>Output Flowrate</td>
</tr>
<tr>
<td></td>
<td>Input Flowrate</td>
<td>Geo. Location</td>
</tr>
<tr>
<td></td>
<td>Geo. Location</td>
<td>Cost of Transport</td>
</tr>
<tr>
<td></td>
<td>Unit Cost of Transport</td>
<td>Cost of Transport</td>
</tr>
<tr>
<td>Storage</td>
<td>Material Input</td>
<td>Material Output</td>
</tr>
<tr>
<td></td>
<td>Max. Capacity</td>
<td>Output Flowrate</td>
</tr>
<tr>
<td></td>
<td>Input Flowrate</td>
<td>Geo. Location</td>
</tr>
<tr>
<td></td>
<td>Geo. Location</td>
<td>Cost of Storage</td>
</tr>
<tr>
<td></td>
<td>Unit Cost of Storage</td>
<td>Cost of Storage</td>
</tr>
<tr>
<td>Process</td>
<td>Material Input</td>
<td>Material Output</td>
</tr>
<tr>
<td></td>
<td>Max. Capacity</td>
<td>Output Flowrate</td>
</tr>
<tr>
<td></td>
<td>Yield of Product</td>
<td>Geo. Location</td>
</tr>
<tr>
<td></td>
<td>Input Flowrate</td>
<td>Cost of Process</td>
</tr>
<tr>
<td></td>
<td>Geo. Location</td>
<td>Unit Cost (CAPEX &amp; OPEX) of Process</td>
</tr>
<tr>
<td>Demand Centre</td>
<td>Material Input</td>
<td>Profit</td>
</tr>
<tr>
<td></td>
<td>Min. Demand</td>
<td>Software</td>
</tr>
<tr>
<td></td>
<td>Selling Price</td>
<td>Software</td>
</tr>
<tr>
<td></td>
<td>Input Flowrate</td>
<td>Software</td>
</tr>
<tr>
<td></td>
<td>Geo. Location</td>
<td>Software</td>
</tr>
</tbody>
</table>

Case Study

- Demonstrate performance of ontological approach to coordinate interoperability using lignocellulosic biorefining supply chain model that maximises total profit

- Process 3 is a requesting model (separation process) that seeks different conversion processes producing ethanol

  Requesting criteria:
  - Functionality of Model
  - Material Input
  - Max Capacity
  - Input Flowrate
  - Yield of Product
  - Geo. Location
  - Cost
  - Software
Case Study

- **Requirement of Requesting Model**

<table>
<thead>
<tr>
<th>Model Functionality for Process</th>
<th>Complexity</th>
<th>Flowrate</th>
<th>Yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>Process 3</td>
<td>Co Fermentation</td>
<td>Shortcut</td>
<td>50,000 kg/hr</td>
</tr>
</tbody>
</table>

- **List of Models in Repository**

<table>
<thead>
<tr>
<th>Model</th>
<th>Scale</th>
<th>Functionality</th>
<th>Model Functionality</th>
<th>Ethanol</th>
<th>Complexity</th>
<th>Flowrate</th>
<th>Yield</th>
<th>Software</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>MODEL 1</td>
<td>Process</td>
<td>Conversion</td>
<td>C6 Fermentation</td>
<td>Yes</td>
<td>Shortcut</td>
<td>117,233</td>
<td>0.116</td>
<td>gProms</td>
<td>(Siougkrou et al. 2016)</td>
</tr>
<tr>
<td>MODEL 2</td>
<td>Process</td>
<td>Conversion</td>
<td>SSF*</td>
<td>Yes</td>
<td>Shortcut</td>
<td>449,353</td>
<td>0.055</td>
<td>AspenPlus</td>
<td>(Humbird et al. 2011)</td>
</tr>
<tr>
<td>MODEL 3</td>
<td>Process</td>
<td>Conversion</td>
<td>Transesterification</td>
<td>No</td>
<td>Detailed</td>
<td>1,004</td>
<td>0.000</td>
<td>AspenPlus</td>
<td>(Zhang et al. 2003)</td>
</tr>
<tr>
<td>MODEL 4</td>
<td>Process</td>
<td>Conversion</td>
<td>Gasification</td>
<td>Yes</td>
<td>Conceptual</td>
<td>3,967</td>
<td>0.066</td>
<td>Data</td>
<td>(Wei et al. 2009)</td>
</tr>
<tr>
<td>MODEL 5</td>
<td>Process</td>
<td>Conversion</td>
<td>C6 Fermentation</td>
<td>Yes</td>
<td>Detailed</td>
<td>74,256</td>
<td>0.121</td>
<td>AspenPlus</td>
<td>(AspenPlus 2007)</td>
</tr>
<tr>
<td>MODEL 6</td>
<td>Process</td>
<td>Conversion</td>
<td>Gasification</td>
<td>Yes</td>
<td>Conceptual</td>
<td>1,653</td>
<td>0.114</td>
<td>Data</td>
<td>(Wei et al. 2009)</td>
</tr>
<tr>
<td>MODEL 7</td>
<td>Process</td>
<td>Conversion</td>
<td>SSF*</td>
<td>Yes</td>
<td>Conceptual</td>
<td>10,722</td>
<td>0.016</td>
<td>Data</td>
<td>(Wei et al. 2009)</td>
</tr>
<tr>
<td>MODEL 8</td>
<td>Process</td>
<td>Conversion</td>
<td>C6 Fermentation</td>
<td>Yes</td>
<td>Conceptual</td>
<td>47,191</td>
<td>0.075</td>
<td>AspenPlus</td>
<td>(Siougkrou et al. 2016)</td>
</tr>
<tr>
<td>MODEL 9</td>
<td>Process</td>
<td>Conversion</td>
<td>Indirect Gasification</td>
<td>No</td>
<td>Detailed</td>
<td>6,507</td>
<td>0.000</td>
<td>AspenPlus</td>
<td>(Spath et al. 2005)</td>
</tr>
</tbody>
</table>

- **Elimination:**
  - No Ethanol & No Cost available

<table>
<thead>
<tr>
<th>Model</th>
<th>Scale</th>
<th>Functionality</th>
<th>Model Functionality</th>
<th>Ethanol</th>
<th>Complexity</th>
<th>Flowrate</th>
<th>Yield</th>
<th>Software</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>MODEL 1</td>
<td>Process</td>
<td>Conversion</td>
<td>C6 Fermentation</td>
<td>Yes</td>
<td>Shortcut</td>
<td>117,233</td>
<td>0.116</td>
<td>gProms</td>
<td>(Siougkrou et al. 2016)</td>
</tr>
<tr>
<td>MODEL 2</td>
<td>Process</td>
<td>Conversion</td>
<td>SSF*</td>
<td>Yes</td>
<td>Shortcut</td>
<td>449,353</td>
<td>0.055</td>
<td>AspenPlus</td>
<td>(Humbird et al. 2011)</td>
</tr>
<tr>
<td>MODEL 3</td>
<td>Process</td>
<td>Conversion</td>
<td>Transesterification</td>
<td>No</td>
<td>Detailed</td>
<td>1,004</td>
<td>0.000</td>
<td>AspenPlus</td>
<td>(Zhang et al. 2003)</td>
</tr>
<tr>
<td>MODEL 4</td>
<td>Process</td>
<td>Conversion</td>
<td>Gasification</td>
<td>Yes</td>
<td>Conceptual</td>
<td>3,967</td>
<td>0.066</td>
<td>Data</td>
<td>(Wei et al. 2009)</td>
</tr>
<tr>
<td>MODEL 5</td>
<td>Process</td>
<td>Conversion</td>
<td>C6 Fermentation</td>
<td>Yes</td>
<td>Detailed</td>
<td>74,256</td>
<td>0.121</td>
<td>AspenPlus</td>
<td>(AspenPlus 2007)</td>
</tr>
<tr>
<td>MODEL 6</td>
<td>Process</td>
<td>Conversion</td>
<td>Gasification</td>
<td>Yes</td>
<td>Conceptual</td>
<td>1,653</td>
<td>0.114</td>
<td>Data</td>
<td>(Wei et al. 2009)</td>
</tr>
<tr>
<td>MODEL 7</td>
<td>Process</td>
<td>Conversion</td>
<td>SSF*</td>
<td>Yes</td>
<td>Conceptual</td>
<td>10,722</td>
<td>0.016</td>
<td>Data</td>
<td>(Wei et al. 2009)</td>
</tr>
<tr>
<td>MODEL 8</td>
<td>Process</td>
<td>Conversion</td>
<td>C6 Fermentation</td>
<td>Yes</td>
<td>Conceptual</td>
<td>47,191</td>
<td>0.075</td>
<td>AspenPlus</td>
<td>(Siougkrou et al. 2016)</td>
</tr>
<tr>
<td>MODEL 9</td>
<td>Process</td>
<td>Conversion</td>
<td>Indirect Gasification</td>
<td>No</td>
<td>Detailed</td>
<td>6,507</td>
<td>0.000</td>
<td>AspenPlus</td>
<td>(Spath et al. 2005)</td>
</tr>
</tbody>
</table>

- **Matching & Performance Ranking Result**

<table>
<thead>
<tr>
<th>Model</th>
<th>Semantic Similarity</th>
<th>Cosine Similarity</th>
<th>Euclidean Similarity</th>
<th>Property Similarity</th>
<th>Aggregated Similarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>MODEL 1</td>
<td>0.833</td>
<td>1.000</td>
<td>0.832</td>
<td>0.916</td>
<td>0.875</td>
</tr>
<tr>
<td>MODEL 2</td>
<td>0.750</td>
<td>1.000</td>
<td>0.000</td>
<td>0.500</td>
<td>0.625</td>
</tr>
<tr>
<td>MODEL 3</td>
<td>0.667</td>
<td>1.000</td>
<td>0.839</td>
<td>0.870</td>
<td>0.818</td>
</tr>
<tr>
<td>MODEL 4</td>
<td>0.667</td>
<td>1.000</td>
<td>0.993</td>
<td>0.996</td>
<td>0.832</td>
</tr>
</tbody>
</table>
Acknowledgement