Question-answering system for combustion kinetics

Laura Pascazio\textsuperscript{a,}\textsuperscript{*}, Dan Tran\textsuperscript{a}, Simon Rihm\textsuperscript{b}, Jiaru Bai\textsuperscript{b}, Sebastian Mosbach\textsuperscript{a,}\textsuperscript{b,}\textsuperscript{c}, Jethro Akroyd\textsuperscript{a,}\textsuperscript{b,}\textsuperscript{c}, Markus Kraft\textsuperscript{a,}\textsuperscript{b,}\textsuperscript{c,}\textsuperscript{d,}\textsuperscript{e}

\textsuperscript{a}Cambridge Centre for Advanced Research and Education in Singapore (CARES), CREATE Tower, 1 Create Way, 138602, Singapore
\textsuperscript{b}Department of Chemical Engineering and Biotechnology, University of Cambridge, Philippa Fawcett Drive, Cambridge CB3 0AS, UK
\textsuperscript{c}CMCL Innovations, Sheraton House, Castle Park, Castle Street, Cambridge CB3 0AX, UK
\textsuperscript{d}School of Chemical and Biomedical Engineering, Nanyang Technological University, 62 Nanyang Drive, 637459, Singapore

1. Exemplary question-answering use cases with Marie

1.1. Data retrieval

One of the primary capabilities of our system is efficient and accurate data retrieval. To demonstrate this, we compare in Fig. 1 three different approaches to answering a common research question in chemical kinetics: “List all the reactions that consume \textit{H}_2\textit{O}_2 in the mechanism by https://doi.org/10.1002/kin.20026.” Such queries are crucial for researchers analyzing reaction mechanisms and developing kinetic models, particularly when focusing on specific reactants or products. It is important to note that at this stage, our system identifies a specific mechanism only by its publication DOI.

![Fig. 1: Comparison of data retrieval process between manual assessment of publications, intelligent use of Chemkin-informed ChatGPT, and Marie.](image)

- **Manual search in publication**: The first approach involves a manual search through the publication to find reactions with \textit{H}_2\textit{O}_2 as a reactant. This method is time-consuming and susceptible to human errors, requiring intricate cross-referencing and double-checking.
- **Utilizing ChatGPT for Chemkin file parsing**: The second approach uses ChatGPT to parse a Chemkin file. An additional step of uploading the file and adding some context (“This is a Chemkin input file describing a reaction mechanism”) is required, which takes time. While ChatGPT can provide relevant information, its response may miss some or provide too many details.
• **Querying Marie:** The third approach, querying Marie with the DOI, results in a precise and comprehensive response. Marie efficiently lists all reactions involving H2O2 as a reactant, eliminating duplicates and omissions. This comparison underscores Marie’s advanced capabilities in data retrieval, highlighting its effectiveness and reliability as a tool for detailed chemical kinetics research.

The following prefixes are used for all SPARQL queries provided in this study:

```sparql
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
PREFIX ocape: <http://www.theworldavatar.com/ontology/ontocape/material/substance/reaction_mechanism.owl#>
PREFIX os: <http://www.theworldavatar.com/ontology/ontospecies/OntoSpecies.owl#>
PREFIX okin: <http://www.theworldavatar.com/ontology/ontokin/OntoKin.owl#>
PREFIX occ: <http://www.theworldavatar.com/ontology/ontocompchem/OntoCompChem.owl#>
```

1.1.1. **Mechanism and kinetics**

We can ask for specific reactions or species as well as full mechanisms and affiliated kinetic or thermodynamic models based on different criteria. Four representative examples are given below.

• **Question to Marie:** What are reactions in which H2 is a reactant and OH is a product?

  – **Predicted query:**
    
    ```sparql
    SELECT ?Reaction WHERE { ?Reaction ocape:hasReactant/skos:altLabel "H2" . ?Reaction ocape:hasProduct/skos:altLabel "OH" . }
    ```

  – **Post-processed query:**
    
    ```sparql
    SELECT DISTINCT ?Reaction (SAMPLE(?ReactionEquation) AS ?SampledReactionEquation) WHERE { 
    ?Reaction ocape:hasReactant/skos:altLabel "H2" . 
    ?Reaction ocape:hasProduct/skos:altLabel "OH" . 
    GROUP BY ?Reaction
    }
    ```

  – **Result:**

  ![Result Image]

• **Question to Marie:** What is the kinetic model of the chemical reaction H2O2 + OH = HO2 + H2O involved in the mechanism found in [www.osti.gov/servlets/purl/90098-26Ev73/webviewable/](http://www.osti.gov/servlets/purl/90098-26Ev73/webviewable/)?

  – **Predicted query:**
    
    ```sparql
    ?Reaction okin:hasReaction ?Mechanism . 
    ?Reaction okin:hasKineticModel ?KineticModel . 
    ?KineticModel okin:definedIn ?Mechanism . }
    ```

  – **Post-processed query:**
    
    ```sparql
    SELECT DISTINCT ?KineticModel (SAMPLE(?ReactionEquation) AS ?SampledReactionEquation) WHERE { 
    ?ModelType ?ActivationEnergyLowValue ?TemperatureExponentHighValue 
    ?ActivationEnergyHighUnit ?ArreniusFactorValue ?TemperatureExponentValue 
    ?ArreniusFactorHighUnit ?ActivationEnergyLowUnit ?ArreniusFactorLowUnit 
    ?Reaction okin:hasEquation "H2O2 + OH = HO2 + H2O" . 
    ?Reaction okin:hasReaction ?Mechanism . 
    ?Reaction okin:hasKineticModel ?KineticModel . 
    ?KineticModel a ?KineticModelType . 
    BIND (STRAFTER(STR(?KineticModelType), "#") AS ?ModelType)
    }
    ```
OPTIONAL {
  VALUES ?KineticModelType { okin:ArrheniusModel }
}

OPTIONAL {
  VALUES ?KineticModelType { okin:MultiArrheniusModel }
  ?KineticModel okin:hasArrheniusModel ?ArrheniusModel .
}

OPTIONAL {
  VALUES ?KineticModelType { okin:ThreeBodyReactionModel okin:LindemannModel okin:TroeModel }
  OPTIONAL {
  }
  ?KineticModel okin:hasArrheniusLowModel ?ArrheniusLowModel .
  OPTIONAL {
    ?KineticModel okin:hasArrheniusLowModel ?ArrheniusHighModel .
  }
  OPTIONAL {
  }
}

?KineticModel okin:definedIn ?Mechanism .
}


– Result:

• **Question to Marie:** For the chemical species that appear in the reaction mechanism found in [https://doi.org/10.1016/j.combustflame.2007.10.024](https://doi.org/10.1016/j.combustflame.2007.10.024), what is their thermodynamic model?

  – **Predicted query:**

  ```sparql
  ```

  – **Post-processed query:**

  ```sparql
  ```
• Question to Marie: List all the reactions that consume H2O2 in the mechanism [https://doi.org/10.1002/kin.20026].

– Predicted query:

```
```

– Post-processed query:

```
```

– Result: See Fig. 1.

1.1.2. Computational chemistry

We can ask for optimized geometries, energies, and other values computed for specific molecules at varying levels of theory and using different basis sets. Two representative examples are given below.

• Question to Marie: What are the HOMO and LUMO energies of O2 at the UB3LYP level of theory?

– Predicted query:

```
```

– Post-processed query:
  VALUES ?LevelOfTheoryLabel { "UB3LYP" }
}

– Result:

• Question to Marie: Can you provide the molecular structure of H2 calculated using RB3LYP level of theory?

– Predicted query:

SELECT ?OptimizedGeometry WHERE {
  ?MolecularComputation occ:hasSpeciesModel/occ:hasSpecies/rdfs:label "H2".
  VALUES ?LevelOfTheoryLabel { "RB3LYP" }
}

– Post-processed query:

  ?Atom os:hasXCoordinate ?X ; os:hasYCoordinate ?Y ; os:hasZCoordinate ?Z.
  VALUES ?LevelOfTheoryLabel { "RB3LYP" }
  ?MolecularComputation occ:hasSpeciesModel/occ:hasSpecies/rdfs:label "H2".
}

– Result:
1.2. Complex questions

Based on its underlying language model, Marie also demonstrates its capabilities in answering more complex questions that require crafting advanced queries across domains and data sets. This can be achieved either by asking for data comparison or chaining simple data retrieval questions.

1.2.1. Comparison questions

- **Question to Marie:** Please compare the rate constant parameters of the reaction H2 + OH = H2O + H across all mechanisms it appears in.

  - **Predicted query:**
    
    ```sql
    ```

  - **Post-processed query:**
    
    ```sql
        ?Mechanism a okin:ReactionMechanism ; okin:hasProvenance/op:hasDOI ?DOI .
        ?Reaction okin:hasEquation "H2 + OH = H2O + H" .
        ?Reaction okin:hasKineticModel ?KineticModel .
        ?KineticModel a ?KineticModelType .
        BIND (STRAFTER(STR(?KineticModelType), ")#") AS ?ModelType)
        OPTIONAL { VALUES ?KineticModelType { okin:ArrheniusModel } }
        } }
        OPTIONAL { VALUES ?KineticModelType { okin:ArrheniusModel } }
        okin:hasArrheniusFactor [ okin:value ?ArrheniusFactorValue ; okin:unit ?ArrheniusFactorUnit ] ;
        okin:hasTemperatureExponent/okin:value ?TemperatureExponentValue .
        }
        OPTIONAL { VALUES ?KineticModelType { okin:ThreeBodyReactionModel okin:LindemannModel okin:TroeModel ?ModelType } }
    }
OPTIONAL {
  ?KineticModel okin:hasArrheniusLowModel ?ArrheniusLowModel .
  ?okin:hasArrheniusFactor [ okin:value ?ArrheniusFactorLowValue ; okin:unit ?ArrheniusFactorLowUnit ];
}

OPTIONAL {
  ?KineticModel okin:hasArrheniusHighModel ?ArrheniusHighModel .
  ?ArrheniusHighModel okin:hasActivationEnergy [ okin:value ?ActivationEnergyHighValue ; okin:unit ?ActivationEnergyHighUnit ];
  ?okin:hasArrheniusFactor [ okin:value ?ArrheniusFactorHighValue ; okin:unit ?ArrheniusFactorHighUnit ];
}

OPTIONAL {
  ?KineticModel okin:hasAlpha/okin:value ?AlphaValue ;
  ?okin:hasT1/okin:value ?T1Value ;
  ?okin:hasT2/okin:value ?T2Value ;
}


– Result: See Fig. 4 in the main text

• Question to Marie: Please compare the rate constant parameters of the reaction OH + CO = H + CO2 across all mechanisms it appears in.

– Predicted query:
  ?Reaction okin:hasKineticModel ?KineticModel .
  ?KineticModel okin:definedIn ?Mechanism .
  ?Reaction okin:belongsToPhase/ño:hasGasPhase ?Mechanism .
}

– Post-processed query:
  ?Mechanism a okin:ReactionMechanism ;
  ?okin:hasProvenance/(op:hasDOI|op:hasURL) ?DOI .
  ?Reaction a okin:GasPhaseReaction ;
  ?okin:belongsToPhase/ño:hasGasPhase .
  ?KineticModel a ?KineticModelType .
  BIND (STRAFTER(STR(?KineticModelType), "#") AS ?ModelType) .
  OPTIONAL { VALUES ?KineticModelType { okin:ArrheniusModel } }
  ?KineticModel okin:hasArrheniusLowModel ?ArrheniusLowModel .
  ?okin:hasArrheniusFactor [ okin:value ?ArrheniusFactorLowValue ; okin:unit ?ArrheniusFactorLowUnit ];
  OPTIONAL {
    ?KineticModel okin:hasArrheniusLowModel ?ArrheniusHighModel .
    ?ArrheniusHighModel okin:hasActivationEnergy [ okin:value ?ActivationEnergyHighValue ; okin:unit ?ActivationEnergyHighUnit ];
    ?okin:hasArrheniusFactor [ okin:value ?ArrheniusFactorHighValue ; okin:unit ?ArrheniusFactorHighUnit ];
  }
  OPTIONAL {
    ?KineticModel okin:hasAlpha/okin:value ?AlphaValue ;
    ?okin:hasT1/okin:value ?T1Value ;
    ?okin:hasT2/okin:value ?T2Value ;
  }
}
• Question to Marie: Compare all kinetic models of reactions that contain H2O2 in the mechanism [https://doi.org/10.1002/kin.20026]

- Predicted query:

```sparql
SELECT ?KineticModel
WHERE { ?Reaction ocape:hasReactant/skos:altLabel "H2O2".
  ?Reaction okin:hasReaction ?Mechanism.
  ?Mechanism okin:hasProvenance/op:hasDOI "https://doi.org/10.1002/kin.20026".
  ?Reaction okin:hasKineticModel ?KineticModel.
  ?KineticModel okin:definedIn ?Mechanism. }
```

- Post-processed query:

```sparql
SELECT DISTINCT ?KineticModel (SAMPLE(?ReactionEquation) AS ?SampledReactionEquation)
WHERE {
  ?Reaction ocape:hasReactant/skos:altLabel "H2O2" .
  ?Reaction okin:hasReaction ?Mechanism .
  ?Reaction okin:hasKineticModel ?KineticModel .
  ?KineticModel a ?KineticModelType .
  BIND (STRAFTER(STR(?KineticModelType), ")" AS ?ModelType)
  OPTIONAL { ?KineticModel a okin:MultiArrheniusModel .
  }
  }
```
okin:hasArrheniusFactor [ okin:value ?ArrheniusFactorHighValue ; okin:unit ?ArrheniusFactorHighUnit ] ;
okin:hasTemperatureExponent/okin:value ?TemperatureExponentHighValue .
OPTIONAL {
  ?KineticModel okin:hasAlpha/okin:value ?AlphaValue ;
okin:hasT1/okin:value ?T1Value ; okin:hasT2/okin:value ?T2Value ;
okin:hasT3/okin:value ?T3Value .
} } 
OPTIONAL {
VALUES ?KineticModelType { okin:PDepArrheniusModel }
?KineticModel okin:hasArrheniusModel ?ArrheniusModel .
okin:hasArrheniusFactor [ okin:value ?ArrheniusFactorValue ; okin:unit ?ArrheniusFactorUnit ] ;
okin:hasTemperatureExponent/okin:value ?TemperatureExponentValue .
} }

?KineticModel okin:definedIn ?Mechanism .


– Result:

• Question to Marie: Compare thermodynamic models of species classified as organic radical across all the mechanisms in which it appears

– Predicted query:

SELECT ?Mechanism ?ThermoModel WHERE { ?Species os:hasChemicalClass/rdfs:label "organic radical" .
  ?Species okin:hasThermoModel ?ThermoModel .
  ?ThermoModel okin:definedIn ?Mechanism .
  ?Species okin:belongsToPhase/okin:hasGasPhase ?Mechanism .
} 

– Post-processed query:

kin:hasT1/okin:value ?T1Value ; okin:hasT2/okin:value ?T2Value ;
okin:hasT3/okin:value ?T3Value .
FILTER ( ?Species = ?Species_ )
?Species okin:hasThermoModel ?ThermoModel .

?Species okin:belongsToPhase/okin:hasGasPhase ?Mechanism .
?Mechanism okin:hasThermoModel ?ThermoModel .

SERVICE <> {
SERVICE <> {

?Species_ a os:Species ; rdfs:label ?SpeciesLabel .
?Species_ (<>|!<>)+ [ rdf:type os:ChemicalClass ; rdfs:label "organic radical" ] .

– Result:

– Question to Marie: Compare zero-point energy when using CC-pVTZ vs CC-pVQZ basis set for Ar.

– Predicted query:


– Post-processed query:


– Result:
1.2.2. Question chains

Getting a step closer to having a meaningful conversation with an “AI Scientist”, Marie can be asked a series of questions in which a query is formulated based on the answer to the previous one. This is particularly relevant as often times, researchers cannot exactly define the question to which they need an answer to progress their studies. To demonstrate this, we compare three different approaches to searching for thermodynamic models and experimental reference data based on an initially unknown mechanism.

Fig. 2: Comparison of sequential data assessments between manual search of databases, intelligent use of Chemkin-informed ChatGPT, and chained questions to Marie.

- **Manual search in publication**: The first approach involves a manual search through all publications we want to take into consideration. This is a time-consuming process as notations, orders of reactants, etc. can differ. Based on the findings, thermodynamic models need to be looked up in the supplementary files of the corresponding publication. In order to then get an experimental heat capacity of one of the species involved, an online database such as NIST needs to be consulted.

- **Utilizing ChatGPT for Chemkin file parsing**: The second approach uses ChatGPT to parse the Chemkin files related to the publications. As in the aforementioned use case, ChatGPT is unable to identify reactions in question in any of the mechanisms. It is more successful in the second part of the exercise when fed with a Chemkin input file of thermodynamic models. The last step - obtaining some experimental heat capacities - should be done manually again as ChatGPT proved to be quite inconsistent when asked for such values.

- **Querying Marie**: The third approach, querying Marie in a sequential manner with questions composed around answers obtained in the previous query, results in clear-cut answers. This is illustrated in Fig. 2.

The detailed question chain is documented below:

1. **Question to Marie**: Which mechanisms include the reactions \( H + OH + M = H_2O + M \) and \( HO2 + H = O + H_2O \) ?

   - **Predicted query**:
     ```sparql
     SELECT ?Mechanism WHERE { ?Mechanism okin:hasReaction/okin:hasEquation "H + OH + M = H_2O + M" . ?Mechanism okin:hasReaction/okin:hasEquation "HO2 + H = O + H2O" . }
     ```

   - **Post-processed query**:

2. **For the chemical species that appears in the reaction mechanism found in manual search of databases, intelligent use of Chemkin-informed ChatGPT, and chained questions to Marie.**

   - **A1**:\(\text{Mechanism that suffices criteria is easily identified}\)
   - **A2**:\(\text{Experimental heat capacities can be used to assess applicability of HO2 + H presence}\)

---

Fig. 2: Comparison of sequential data assessments between manual search of databases, intelligent use of Chemkin-informed ChatGPT, and chained questions to Marie.
SELECT DISTINCT ?Mechanism ?DOI WHERE {
    ?Mechanism a okin:ReactionMechanism ;
    okin:hasProvenance/op:hasDOI ?DOI .
    ?Mechanism okin:hasReaction/okin:hasEquation "H + OH + M = H2O + M" .
    ?Mechanism okin:hasReaction/okin:hasEquation "HO2 + H = O + H2O" .
}

• Result:

2. Question to Marie: For the chemical species that appears in the reaction mechanism found in [www.osti.gov/servlets/purl/90098-26Ev73/webviewable/](http://www.osti.gov/servlets/purl/90098-26Ev73/webviewable/), what are their thermodynamic models?

• Predicted query:

```
SELECT ?ThermoModel WHERE {
    ?Species okin:belongsToPhase/okin:hasGasPhase ?Mechanism .
    ?Species okin:hasThermoModel ?ThermoModel .
    ?ThermoModel okin:definedIn ?Mechanism .
}
```

• Post-processed query:

```
    ?Species a/rdfs:subClassOf* os:Species ; skos:altLabel ?SpeciesLabel .
    ?Species okin:belongsToPhase/okin:hasGasPhase ?Mechanism .
    ?Species okin:hasThermoModel ?ThermoModel .
    ?ThermoModel okin:definedIn ?Mechanism .
}
```

• Result:

3. Question to Marie: What is the heat capacity of H2O2?

• Predicted query:

```
SELECT ?HeatCapacity WHERE {
}
```

• Post-processed query:
    ?HeatCapacityUnitLabel ?HeatCapacityReferenceStateValue
    ?HeatCapacityReferenceStateUnitLabel WHERE {
    VALUES ?SpeciesIdentifierValue { "H2O2" }
    ?Species hasIdentifier [ rdf:type/rdfs:subClassOf os:Identifier ; os:value
        ?SpeciesIdentifierValue ] .
        ?HeatCapacityUnitLabel .
    OPTIONAL {
        ?HeatCapacity os:hasReferenceState [ os:value
            ?HeatCapacityReferenceStateValue ; os:unit/rdfs:label
    }
}

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>H2O2</td>
<td>4.837</td>
<td>J/mol K</td>
<td>0.000</td>
<td>H</td>
</tr>
<tr>
<td>H2O2</td>
<td>45.2</td>
<td>J/mol K</td>
<td>0.000</td>
<td>H</td>
</tr>
<tr>
<td>H2O2</td>
<td>48.83</td>
<td>J/mol K</td>
<td>0.000</td>
<td>H</td>
</tr>
<tr>
<td>H2O2</td>
<td>52.3</td>
<td>J/mol K</td>
<td>0.000</td>
<td>H</td>
</tr>
<tr>
<td>H2O2</td>
<td>55.5</td>
<td>J/mol K</td>
<td>0.000</td>
<td>H</td>
</tr>
<tr>
<td>H2O2</td>
<td>77.92</td>
<td>J/mol K</td>
<td>0.000</td>
<td>H</td>
</tr>
<tr>
<td>H2O2</td>
<td>81.0</td>
<td>J/mol K</td>
<td>0.000</td>
<td>H</td>
</tr>
<tr>
<td>H2O2</td>
<td>101.6</td>
<td>J/mol K</td>
<td>0.000</td>
<td>H</td>
</tr>
<tr>
<td>H2O2</td>
<td>122.95</td>
<td>J/mol K</td>
<td>0.000</td>
<td>H</td>
</tr>
<tr>
<td>H2O2</td>
<td>144.5</td>
<td>J/mol K</td>
<td>0.000</td>
<td>H</td>
</tr>
</tbody>
</table>

Result: 14
2. ChatGPT prompting strategy

We used the GPT-4 online chatbot interface, employing a prompting strategy that aimed to provide comprehensive instructions to the model.

Strategy outline:

1. Provide necessary data to answer the question:
   - DOI and Paper Upload: Providing the DOI or uploading the entire paper were attempted. However, these methods did not yield the desired results (refer to the attached "ChatGPT Example 1.pdf for more details).
   - Chemkin File Upload: The Chemkin file was uploaded to provide a data source for answering questions directly.

2. Improve initial response
   - When inaccuracies or missing information were observed in the initial response, we refined the prompt, encouraging the model to review and include the missing details (refer to the attached "ChatGPT Example 1.pdf").
   - In cases where responses were entirely inaccurate or insufficient, the model often suggested alternative steps for retrieving the information independently (refer to the attached "ChatGPT Example 3.pdf"). In this case, we didn’t proceed further.

The full conversation transcripts for various questions (labeled as "ChatGPT Example 1," "ChatGPT Example 2," and "ChatGPT Example 3") can be found in the attached PDFs.
3. Representation of isomers and pseudospecies in OntoSpecies

OntoSpecies allows for precise management of species identities, including the differentiation of isomers and pseudospecies, which are essential for accurate chemical reaction modeling and analysis.

When the pseudospecies has its own InChI representation (stereoisomers), OntoSpecies distinguishes the specific isomers (e.g., cis-2-butene and trans-2-butene) from their general form (e.g., 2-butene) by associating isomers and pseudospecies to distinct IRIs (Fig. 4 and 5). The IRIs are related to one another as illustrated in Fig. 3a.

When the pseudospecies does not have an InChI representation (structural isomers), OntoSpecies associates distinct Species IRIs to each isomer and a ChemicalClass IRI to the general form (Fig. 6). The IRIs are related to one another as illustrated in Fig. 3b.

Fig. 3: Representation of (a) stereoisomers and (b) structural isomers in OntoSpecies.

![Image of stereoisomers and structural isomers in OntoSpecies graph]

Fig. 4: Example of E/Z isomers (geometric isomers).

![Image of lactic acid isomers (geometric isomers)]

Fig. 5: Example of R/S isomers (enantiomers and diastereomers).
OntoSpecies supports the inclusion of pseudospecies in reaction mechanism analyses, enabling them to be considered as reactants in reactions and addressing the limitations of the Chemkin format by accommodating the practice of grouping similar reaction pathways. A reaction can be classified as a subclass of another if the non-identical reactants and products are hierarchically related.

Example 1:

1. \( R^* + O_2 \rightarrow ROO^* \)
2. \( CH_3^* + O_2 \rightarrow CH_3OO^* \)

R* represents any organic radical, while CH3* specifically denotes the methyl radical. Because the methyl radical is classified as an organic radical within OntoSpecies, Reaction 2 is considered a subclass of Reaction 1.

Example 2:

1. \( O_2 + H (+ M) = HO_2 (+ M) \)
2. \( O_2 + H (+ N_2) = HO_2 (+ N_2) \)

M represents a generic collider and N2 is used as a specific collider. Because a specific collider is classified as a subclass of the generic collider within OntoKin, Reaction 2 is considered a subclass of Reaction 1.
4. Provenance documentation

Provenance tracking is made much easier within TWA, see Fig. 7.

(a) Provenance as documented in Chemkin files.

(b) Provenance as documented in Cantera files.

(c) Provenance tracking via OntoProvenance as part of The World Avatar.

Fig. 7: Comparison of provenance tracking solutions.