Supporting Information

for

OntoKin: An Ontology for Chemical Kinetic Reaction Mechanisms

Feroz Farazi[a)], Jethro Akroyd[a),b),c)], Sebastian Mosbach[a),b),c)], Philipp Buerger[a)], Daniel Nurkowski[c)], Maurin Salamanca[a),b)], Markus Kraft[a),b),c),d),*]

a) Department of Chemical Engineering and Biotechnology, University of Cambridge, Philippa Fawcett Drive, Cambridge CB3 0AS, United Kingdom

b) Cambridge Centre for Advanced Research and Education in Singapore (CARES), CREATE Tower, 1 Create Way, Singapore 138602

c) CMCL Innovations, Sheraton House, Castle Park, Castle Street, Cambridge CB3 0AX, United Kingdom

d) School of Chemical and Biomedical Engineering, Nanyang Technological University, 62 Nanyang Drive, Singapore 637459

*Corresponding author (MK): e-mail: mk306@cam.ac.uk

This document presents some queries to demonstrate how to retrieve information from the knowledge base using both the OntoKin system and SPARQL Endpoint and some formalizations.

Query 1. Show all of the mechanisms available in the knowledge base.

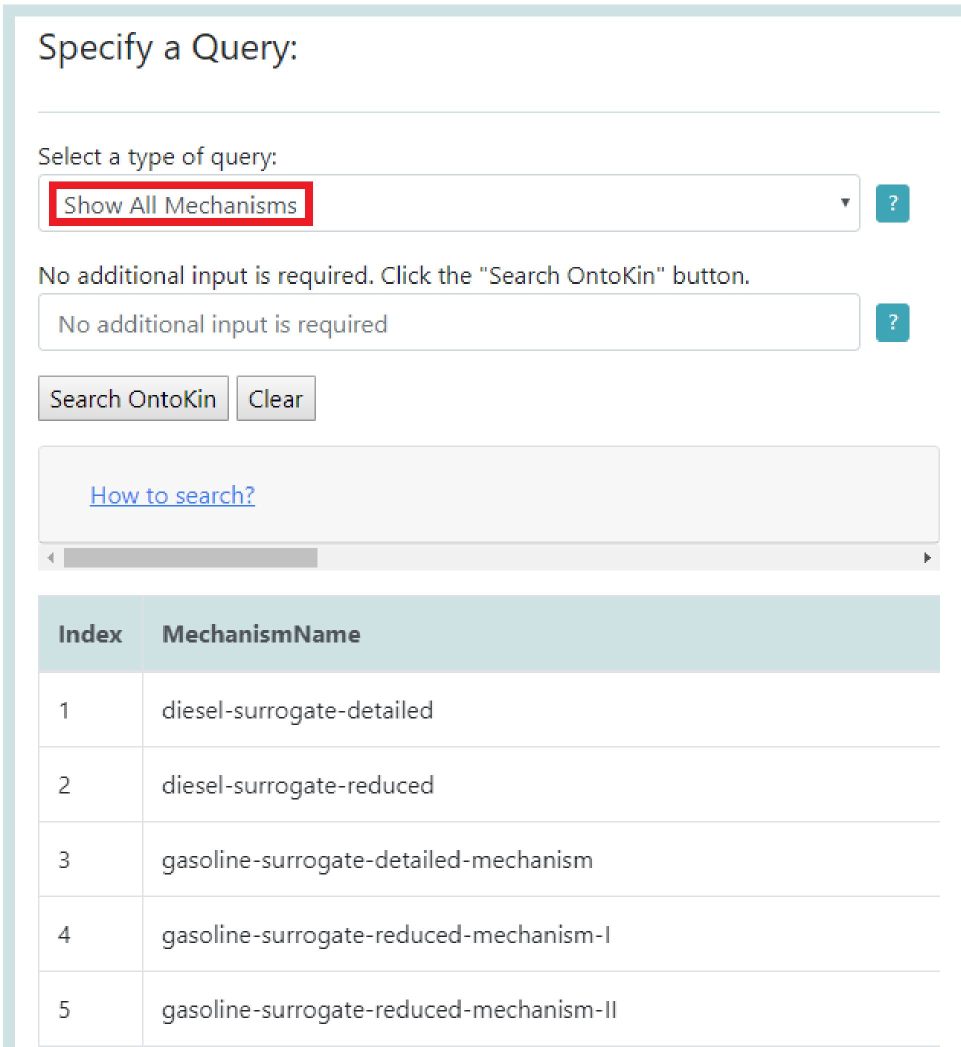

Figure **S1**: Screen-shot of the UI of the OntoKin system depicting how to query for viewing the list of all the mechanisms available in the knowledge base. The red-rectangle marked text in the drop-down menu is the query selected for this purpose and the search box below this menu is left empty. A click on the 'Search Ontokin' button shows the corresponding result. The first part of the result of Query 1 is shown here.

6	Methyl-decanoate-plus-methyl-9-decenoate-plus-n-heptane
7	Methyl-decanoate-plus-n-heptane-mechanism
8	N-Heptane
9	POLIMI_BIOGASOLINE_REDUCED171R_1410
10	POLIMI_BIO_1412
11	POLIMI_C1C3_HT_1412
12	POLIMI_C1C3_HT_NOX_1412
13	POLIMI_C1C3_LT_1412
14	POLIMI_DIESEL_BIODIESEL_REDUCED226_1410
15	POLIMI_DIESEL_REDUCED201C_1410
16	POLIMI_FAME_REDUCED177_1410
17	POLIMI_GASOLINE_REDUCED156R_1410
18	POLIMI_H2CO_1412
19	POLIMI_H2CO_NOX_1412
20	POLIMI_KEROSENE_LTHT_REDUCED231_1410

Figure **S2**: The second part of the result of Query 1 is shown here.

⋮

39	GRIMECH_30
40	GRIMECH_30_1
41	AramcoMech3_2018
42	AramcoMech2_2016
43	MP_Detail_Dooley_2013
44	MP_Reduced_Dooley_2013
45	MP_Felsmann_2017
46	MP_Huang_2014_3
47	MP_Reduced_Dooley_2013_1
48	NO_Konnov_2009_1
49	NO_Qi_Konnov_2019
50	GRI_Mech_2_11
51	GRI_Mech_1_2

Figure **S3**: The last part of the result of Query 1 is shown here.

Query 2. Show chemical mechanisms that contain O_2.

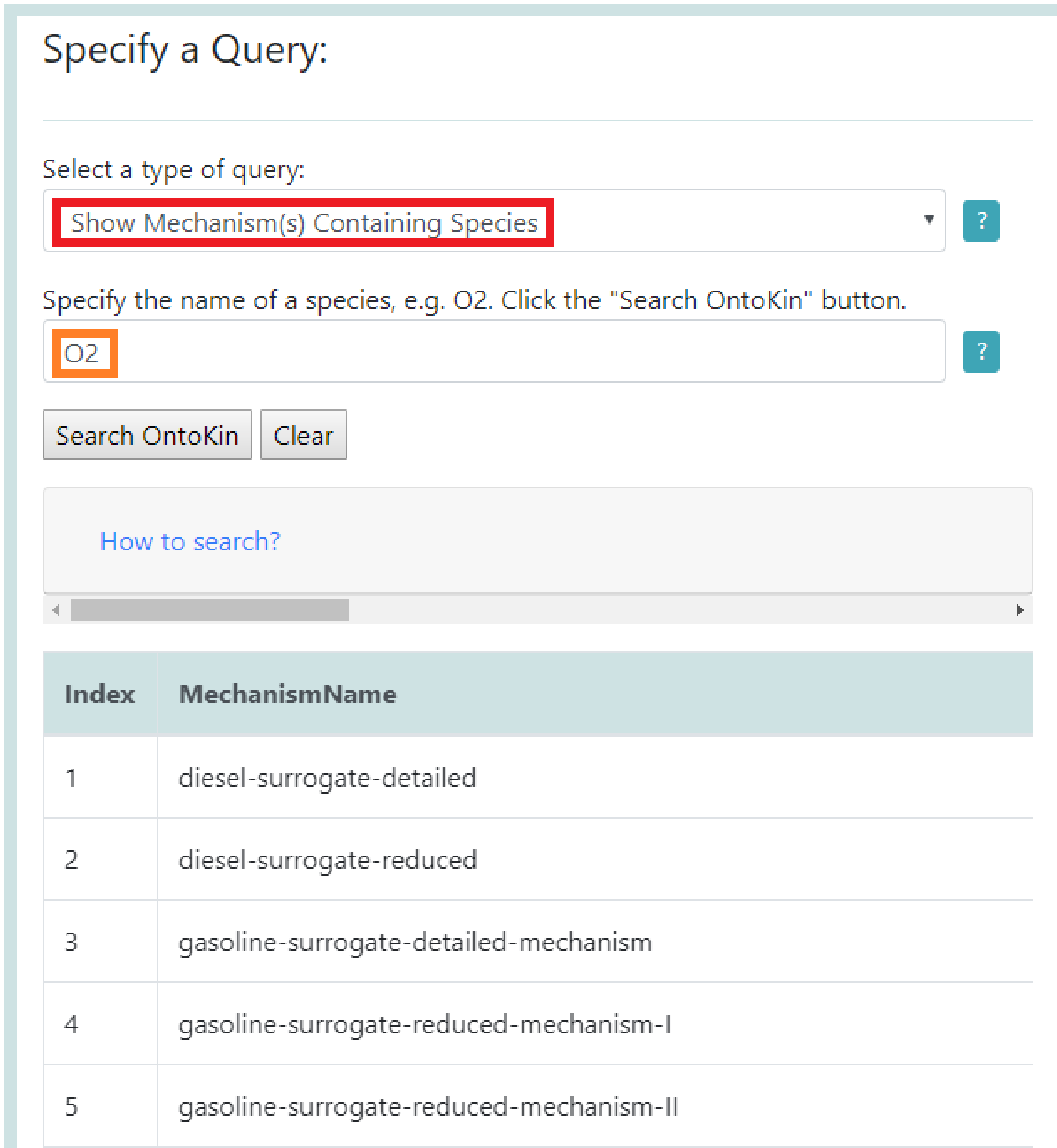

Figure **S4**: Screen-shot depicting how to query for viewing the list of mechanisms containing O_2. The orange-rectangle marked text is the user supplied (typed in) query and the red-rectangle marked text in the drop-down menu is the type of query selected for this purpose. A click on the 'Search OntoKin' button shows the corresponding result and a partial list of retrieved mechanisms is shown here.

Query 3. A SPARQL query to retrieve all the mechanisms available in the knowledge base is provided in the following text box.

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX ontokin: <http://www.theworldavatar.com/kb/ontokin/ontokin.owl#>
SELECT ?MechanismName
WHERE
  {
    ?MechanismIRI rdf:type ontokin:ReactionMechanism .
    ?MechanismIRI rdfs:label ?MechanismName .
  }
```

Click on `http://www.theworldavatar.com/OntoKinGUI` to open the SPARQL query interface of the knowledge base, of which a screenshot is shown in Figure **S5**. Press the 'Clear' button to empty the 'Query' box (white marked area), copy the query from the above text box, paste it in the emptied box and click on the 'Execute' button to see the query result.

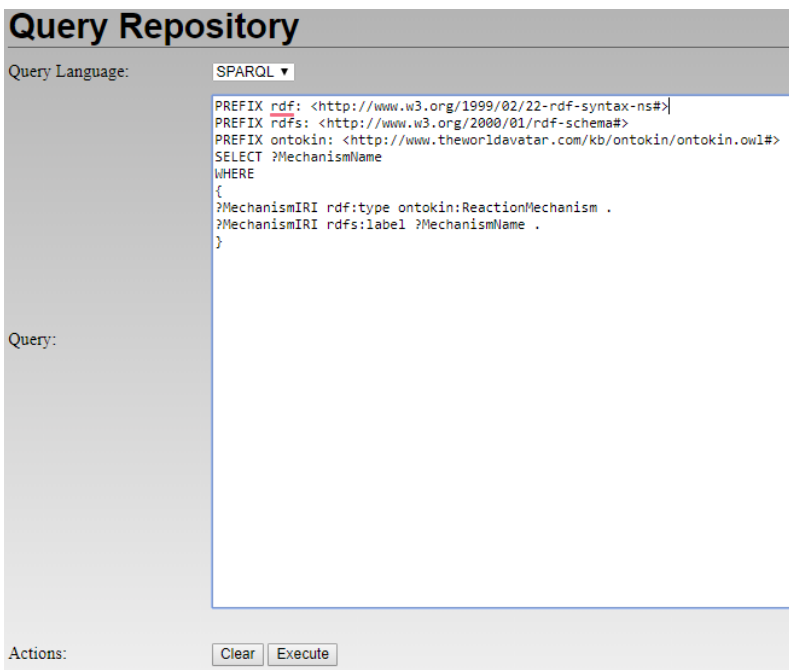

Figure **S5**: The user interface to query the OntoKin knowledge base using SPARQL. A SPARQL query that corresponds to Query 3 is shown in the 'Query' box.

Query 4. A SPARQL query to retrieve chemical mechanisms that contain a specific species, in this case O_2, is provided in the following text box.

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX ontokin: <http://www.theworldavatar.com/kb/ontokin/ontokin.owl#>
SELECT ?MechanismName
WHERE
  {
   ?species rdfs:label "O2" .
   ?species ontokin:belongsToPhase ?phase .
   ?phase ontokin:containedIn ?mechanism .
   ?mechanism rdf:type ontokin:ReactionMechanism .
   ?mechanism rdfs:label ?MechanismName .
  }
```

Query 5. A SPARQL query to compare thermodynamic data of the species O_2 across all the mechanisms in which it appears.
The query is given in the following text box.

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX ontokin: <http://www.theworldavatar.com/kb/ontokin/ontokin.owl#>
SELECT ?MechanismName ?MechanismIRI ?SpeciesIRI ?ThermoModelIRI ?Coeffi-
cientValues ?MinTemp ?MaxTemp ?Pressure
WHERE
  {
   ?SpeciesIRI rdfs:label "O2" .
   ?SpeciesIRI ontokin:belongsToPhase ?Phase .
   ?Phase ontokin:containedIn ?MechanismIRI .
   ?MechanismIRI rdf:type ontokin:ReactionMechanism .
   ?MechanismIRI rdfs:label ?MechanismName .
   ?SpeciesIRI ontokin:hasThermoModel ?ThermoModelIRI .
   ?ThermoModelIRI ontokin:hasCoefficientValues ?CoefficientValues .
   ?ThermoModelIRI ontokin:hasMinimumTemperature ?MinTemp .
   ?ThermoModelIRI ontokin:hasMaximumTemperature ?MaxTemp .
   ?ThermoModelIRI ontokin:hasPressure ?Pressure .
  }
```

Query 6. A SPARQL query to compare Arrhenius parameters of the reaction $O_2 + N \longrightarrow O + NO$ across all the mechanisms in which it appears.
The query is given in the following text box.

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX ontokin: <http://www.theworldavatar.com/kb/ontokin/ontokin.owl#>
SELECT ?MechanismName ?MechanismIRI ?ReactionIRI ?ActivationEnergy ?ActivationEnergyUnits ?PreExpFactor ?PreExpFactorUnits ?TempExponent ?TempExpUnits
WHERE
  {
    ?ReactionIRI ontokin:hasEquation "O2 + N => O + NO" .
    ?ReactionIRI ontokin:belongsToPhase ?Phase .
    ?Phase ontokin:containedIn ?MechanismIRI .
    ?MechanismIRI rdfs:label ?MechanismName .
    ?ReactionIRI ontokin:hasArrheniusCoefficient ?ArrheniusRateCoefficients .
    ?ArrheniusRateCoefficients ontokin:hasActivationEnergy ?ActivationEnergy .
    ?ArrheniusRateCoefficients ontokin:hasActivationEnergyUnits ?ActivationEnergyUnits .
    ?ArrheniusRateCoefficients ontokin:hasPreExponentialFactor ?PreExpFactor .
    ?ArrheniusRateCoefficients ontokin:hasPreExponentialFactorUnits ?PreExpFactorUnits .
    ?ArrheniusRateCoefficients ontokin:hasTemperatureExponent ?TempExponent .
    ?ArrheniusRateCoefficients ontokin:hasTemperatureExponentUnits ?TempExpUnits .
  }
```

This section presents formalizations of the methodological steps Motivating Scenario, Competency Questions and Scope.

Formalization 1. Motivating scenarios are a pair $\langle P_{ca},\ P_{sa}\rangle$, where P_{ca} refers to the problems which when resolved may lead to complete automation, and P_{sa} refers to the problems which when resolved may lead to semi-automation with less manual steps than before.

Formalization 2. Competency questions are a 3-tuple $\langle Q, Q_{TBox}, Q_{ABox}\rangle$, where Q means the list of questions that requires the inclusion of both the TBox and ABox to produce responses against them, while Q_{TBox} is the list of questions that the TBox alone should be able to answer and Q_{ABox} refers to the list of questions that ABox alone should be able to answer.

Formalization 3. A scope is a 5-tuple $\langle M_s, C_Q, D_i, d_i^j, T_k^{ij}\rangle$, where M_s is a set of motivating scenarios, C_Q is a set of competency questions, D_i is a set of all possible domains and i=1,. . . ,n, d_i^j is a set of subdomains belonging to the domain D_i and j=1,. . . ,m and T_k^{ij} is a set of topics belonging to the subdomain d_i^j and k=1,. . . ,p.