

# OntoKin: An Ontology for Chemical Kinetic Reaction Mechanisms

Feroz Farazi,<sup>†</sup> Jethro Akroyd,<sup>†,‡,¶</sup> Sebastian Mosbach,<sup>†,‡,¶</sup> Philipp Buerger,<sup>†</sup> Daniel  
Nurkowski,<sup>¶</sup> Maurin Salamanca,<sup>†,‡</sup> and Markus Kraft<sup>\*,†,‡,§,¶</sup>

<sup>†</sup>*Department of Chemical Engineering and Biotechnology, University of Cambridge,  
Philippa Fawcett Drive, Cambridge CB3 0AS, United Kingdom*

<sup>‡</sup>*Cambridge Centre for Advanced Research and Education in Singapore (CARES),  
CREATE Tower, 1 Create Way, Singapore 138602*

<sup>¶</sup>*CMCL Innovations, Sheraton House, Castle Park, Castle Street, Cambridge CB3 0AX,  
United Kingdom*

<sup>§</sup>*School of Chemical and Biomedical Engineering, Nanyang Technological University, 62  
Nanyang Drive, Singapore 637459*

E-mail: mk306@cam.ac.uk

Phone: +44 (0)1223 762784

## Abstract

An ontology for capturing both data and the semantics of chemical kinetic reaction mechanisms has been developed. Such mechanisms can be applied to simulate and understand the behaviour of chemical processes, for example, the emission of pollutants from internal combustion engines. An ontology development methodology was used to produce the semantic model of the mechanisms, and a tool was developed to automate the assertion process. As part of the development methodology, the ontology is formally represented using OWL, assessed by domain experts and validated by applying a

reasoning tool. The resulting ontology, termed OntoKin, has been used to represent example mechanisms from the literature. OntoKin and its instantiations are integrated to create a Knowledge Base (KB), which is deployed using the RDF4J triple store. The use of the OntoKin ontology and the KB is demonstrated for three use cases - querying across mechanisms, modelling atmospheric pollution dispersion and a mechanism browser tool. As part of the query use-case, the OntoKin tools have been applied by a chemist to identify variations in the rate of a prompt  $\text{NO}_x$  formation reaction in the combustion of ammonia as represented by four mechanisms in the literature.

## Introduction

The Semantic Web<sup>1</sup> offers the ability to attach meaning to data and represent extracted knowledge from data. Ontologies<sup>2</sup> are the core means to represent data, its meaning and knowledge about data on the Semantic Web. Agents<sup>3</sup> can use the Semantic Web easily to find ontologies to assist (human and non-human) users. Having been influenced by the Semantic Web, many people have taken the initiative to build a Chemical Semantic Web<sup>4,5</sup> using chemical ontologies<sup>6,7</sup> to meet an increasing interest to generate knowledge from chemical data and to facilitate data sharing<sup>8</sup>.

A number of ontologies have been developed to capture and represent the semantics and knowledge of chemicals and chemical interactions with different levels of granularity. The Chemicals ontology<sup>7</sup> was developed to represent elements and chemical substances. Sankar and Aghila<sup>9</sup> applied the same methodology used to develop the Chemicals ontology to build an ontology for organic reactions, and developed the Chemical Ontological Support System (COSS) to retrieve instances of different types of organic reactions using a semantic search. Hastings et al.<sup>4</sup> created an ontology represented in OWL (Web Ontology Language) to classify chemical compounds based on their structure. Marquardt and co-workers developed OntoCAPE (Ontology for Computer Aided Process Modelling) as a formal ontology for modelling chemical process engineering<sup>10</sup>, including the concepts of elements, species,

molecular entities and reactions.

In addition, a number of cross-domain ontologies that cover aspects of chemical modelling have been developed. ChEBI<sup>11</sup> is an ontology created for representing concepts and relations (object properties) belonging to chemistry and biology. ChEBI supports the structural description of ‘small’ molecules. Hill et al.<sup>12</sup> combined ChEBI and the Gene Ontology to develop the BioChEBI ontology to integrate chemical and biological descriptions of gene products. CompChem<sup>13</sup> was developed as an RDF-based representation of the semantics of computational chemistry calculations, whilst the Gainesville Core ontology<sup>14</sup> allows the description of the geometric structure of molecules. PubChemRDF<sup>15</sup> represents structures and metadata of chemical substances and compounds. In addition to chemical semantic resources, there are initiatives that have led to well-established chemical databases (PubChem<sup>16</sup>, PrIme<sup>17</sup> and Reaxys (<https://www.reaxys.com>) - to name a few).

A chemical kinetic reaction mechanism (also termed mechanism or reaction mechanism) is a fundamental part of simulations to investigate the behaviour of chemical processes such as the emission of pollutants from combustion processes. The J-Park Simulator (JPS) (<http://www.theworldavatar.com>) is an example of an intelligent application toolset that, following ideas discussed by Kraft and Mosbach<sup>8</sup>, uses semantic representation to harnesses the reasoning and inferencing power of ontologies to perform cross-domain simulations. We can apply ideas related to the Chemical Semantic Web to address the needs of the combustion community, where there is an interest in understanding how the chemistry of a fuel affects its performance. Mechanisms for fuels can be very complex, containing thousands of species and reactions. One challenge facing the community is the inconsistency between different models of the same fuel, where models have been developed for different situations and tested against different data. Ideally, such models should be consistent and universally applicable.

The purpose of this paper is to take steps towards addressing these issues by developing an ontology to represent chemical kinetic reaction mechanisms. Whenever possible, we reuse

concepts and object properties from existing ontologies, in particular OntoCAPE. The paper makes the following contributions:

- The development of OntoKin, an ontology for representing reaction mechanisms.
- A Knowledge Base (KB), which includes OntoKin and representations of mechanisms using OWL. This KB consists of two components: *TBox* and *ABox*. The TBox provides general knowledge (also known as intensional knowledge) about domains of interest in the form of concepts, relations between them and properties that describe them. The concepts form the terminology of the domains involved, thus called terminological box or TBox. OntoKin is the TBox of the KB. On the other hand, the ABox provides knowledge specific to instances belonging to the domains. Such knowledge is termed existential knowledge or assertional knowledge, thus called assertional box or ABox. Reaction mechanisms form the ABox of the KB.
- An ABox Manager tool to convert mechanisms from CHEMKIN format to OWL, and vice versa. CHEMKIN-III<sup>18,19</sup>, and its predecessors and successors, defined a *de facto* standard file format for mechanisms in the combustion community.

The paper is structured as follows. The subsequent section gives an overview of the J-Park Simulator (JPS), which is the context of this work. Later, the Development of OntoKin section describes the methodological approach and creation of the ontology and the Population of the Knowledge Base section shows how the KB was filled in with mechanisms. Finally, we present some Use Cases to demonstrate the application of OntoKin and draw Conclusions.

## J-Park Simulator (JPS) and the World Avatar

JPS is an instance of the World Avatar project (<http://www.theworldavatar.com>) to develop a decentralized system that supports data-driven decision making via the use of data

and models that are represented and linked using a knowledge graph. The approach allows the use of automated intelligent software agents to generate, store and analyse data, and enables the interoperability of data and models across multiple domains.

JPS structures data using a knowledge graph built upon the principles of linked data using ontologies. This enables the encoding of data (both real-world observations and model output) to record the state of a system and models (both physics-based and data-based models) to describe the behavior of a system as a function of its state and other model parameters. JPS facilitates automation of tasks via the introduction of computational and representational agents to operate on the knowledge graph. Eibeck et al.<sup>20</sup> broadly classified the JPS agents into Types 0, 1, ..., 4 based on their activities. Agents are represented using the OntoAgent ontology<sup>21</sup>. Complete integration of newly developed agents with OntoAgent extends the agent eco-system of JPS.

JPS has been applied to many aspects of Industry 4.0<sup>22</sup>. One example concerns the Eco-Industrial Park (EIP) on Jurong Island in Singapore. An EIP is a group of product manufacturers and service providers working together based on exchange relations to address issues related to pollution and reuse (of waste materials) to achieve environmental and economic benefit<sup>23</sup>. In an EIP, the geographical vicinity of enterprises forming the group is important for cost efficiency and environmental footprint reduction<sup>24</sup>. EIPs may include material exchanges, energy systems and wastewater treatment networks, which can be modelled at different levels such as unit operations, processes, plants and networks as well as optimized for improved performance<sup>25,26</sup>.

In the context of JPS, ontologies including *OntoEIP*, the *EIP energy system ontology* and the *biodiesel plant ontology* have been developed by extending the relevant branches of OntoCAPE<sup>27</sup>. OntoEIP is an ontology designed for creating a knowledge base with information originating from resource and transportation networks, chemical process systems and plants for managing an EIP<sup>28</sup>. The EIP energy system ontology was developed for building a decision-making system by integrating data from heterogeneous sources<sup>29</sup>. The

biodiesel plant ontology was built for creating a knowledge base with information about the process simulation and optimization of biodiesel production<sup>30</sup>.

One application of interest is to couple simulations of emissions from internal combustion engines to simulations of the atmospheric dispersion of these emissions. This requires a chemical model (*i.e.* a reaction mechanism) for the fuel used by the engine to be represented as part of the knowledge graph. The work described in this paper is positioned within this context. It addresses the needs of JPS by developing an ontology to represent chemical mechanisms and a knowledge base to store corresponding mechanism data. This will support the automation of processes within JPS by enabling the introduction of intelligent agents to query the knowledge base, which becomes part of the JPS knowledge graph, to search for, find and extract mechanisms for a given task.

## The Development of OntoKin

An iterative approach was used to develop the ontology. This allowed learning from previous iterations to be incorporated into the development process to improve the quality of the ontology. At each iteration, three sequential top level activities were performed – preparation, creation and/or modification and quality assessment – to accomplish a macro task. For example, to extend the ontology to include a defined set of additional concepts and object properties.

- *Preparation:* Decide the tasks that will be executed in the current iteration and who should be involved for the successful accomplishment of the tasks.
- *Creation and/or Modification:* Create and/or modify the ontological TBox<sup>31</sup> consisting of concepts, relations between them and properties, possibly with assignment of their domains and ranges.
- *Quality assessment:* Assess the quality of the completed subtasks to identify opportunities to improve other areas of the TBox.

These activities follow the ontology development approaches proposed by Grüninger and Fox<sup>32</sup>, Uschold and King<sup>33</sup>, Noy and McGuinness<sup>34</sup>, Pinto and Martins<sup>35</sup>, Giunchiglia et al.<sup>36,37</sup> and Fernandez-Lopez et al.<sup>38</sup>. The following subsections describe the steps included in each phase in order to demonstrate how we developed the OntoKin ontology.

## Preparation

The *preparation* task is a sequential combination of understanding the purpose<sup>33,35</sup> and motivating scenario<sup>32</sup>, identifying competency questions<sup>32,34</sup> and defining the scope<sup>34,35</sup>.

### Identify Purpose

A purpose consists of three aspects - *what*, *who* and *how*. ‘*What*’ refers to the intended uses of the ontology being developed, ‘*who*’ specifies the intended users of the ontology, and ‘*how*’ refers to how it is going to be used.

The uses (*what*) of OntoKin are manifold. Firstly, to add semantics to mechanisms in order to allow agents to comprehend them automatically. Secondly, to support semantics-enabled complex query answering. Chemical engineers, combustion scientists, chemists and materials scientists are potential users (*who*) of the ontology. OntoKin is intended to be used (*how*) in the development of a KB.

### Motivating Scenario

Data interoperability across many chemical engineering systems which deal with kinetics was one of the issues encountered within the community. Use of OntoKin will have the potential to allow such systems to communicate data without human intervention. Detecting species thermodynamic data and transport data inconsistency across existing mechanisms will reduce the manual steps in creating a new mechanism from them.

## Competency Questions

The ontology developers or ontologists with the help of domain experts compile a list of questions that the TBox and ABox separately or combined must be able to answer. It is usually produced in a number of iterations in meetings between ontologists and domain experts. The goal is to collect all possible questions that the domain experts are interested in to query.

A list of example questions which needs the presence of both *TBox* and *ABox* statements of OntoKin is given as follows.

- How many *Three-body Reactions* are involved in a given *Reaction Mechanism*?
- List all *Rate Coefficients* used in a given *Reaction Mechanism*.
- Which *Reaction Mechanisms* are created by *Person P* and have a *Reference* to a *Journal* article?

On the other hand, a list of example questions which can be answered with statements provided in the OntoKin TBox is as follows.

- Is *Troe Reaction* a *Fall-off Reaction*, or *General Pressure Dependent Reaction*?
- Can the *Coverage Dependency* property be applied to a *Gas-phase Reaction*?
- What are the special properties of different types of *Rate Coefficients*?

The example question on *Troe Reaction* helped ontologists to understand that users are interested in the hierarchy of reactions, while the *Coverage Dependency* property and the special properties of *Rate Coefficient* questions emphasized the fact that the properties of these concepts should be modelled explicitly by setting their domains and ranges.

A list of example questions which requires statements generated through instantiations of classes and properties and encoded in OntoKin ABoxes is as follows.



- List all *Reaction Mechanisms* that contain *Species*  $S_1$ ,  $S_2$  and  $S_3$ .
- Show the high-temperature thermodynamic data of all *Species* in a given *Reaction Mechanism*.
- Compare the rate constant parameters of the reaction  $\text{O} + \text{C}_3\text{H}_4 \rightleftharpoons \text{CH}_3 + \text{HCCO}$  in all the mechanisms in which it appears.

## Identify Scope

With reference to the Motivating Scenario and Competency Questions described in the previous subsections, the domains Chemical Engineering and Publication were found pertinent. The Mechanism subdomain was found relevant for the former and the Reference subdomain for the latter. The following topics of the former domain were identified as relevant:

- *Phase*: A phase of a substance is a form of matter that is uniform throughout in chemical composition and state (i.e. a particular solid, liquid, or gas). Note that the ontological modelling of the liquid phase is beyond the scope of this paper.
- *Chemical Reaction*: A process in which one or more substances, the reactants, are converted to one or more new and entirely different substances, the products. In chemistry, substances are either chemical elements or compounds. A chemical reaction rearranges the constituent atoms of the reactants to create different substances as products.
- *Rate Coefficients*: The coefficients used to evaluate the rate constant that appears in a reaction rate expression.
- *Species*: An ensemble of chemically identical molecular entities.
- *Thermodynamic Model*: A model that specifies the relationship between certain intensive thermodynamic state variables.

- *Transport Model*: A model that specifies how to calculate the transport properties of species within a phase.

On the other hand, for the Reference subdomain, the topics which were found relevant are as follows:

- *Article*: A scientific document published in a publishing outlet, *e.g.* a journal or conference proceedings.
- *Author*: The creator of an article, *e.g.* a person or an organization.
- *Type of reference*: The type or category of a reference, *e.g.* a journal article or a proceedings article.

## Creation and/or Modification

The *creation and/or modification* task is a sequential combination of concept and property identification<sup>33,37</sup>, analysis<sup>37</sup>, synthesis<sup>37</sup>, standardization<sup>36,37</sup> and formalization<sup>35,37,38</sup>.

### Concept and Property Identification

A *concept* is the representative name of a group of things. A *property* is a means to characterize a concept. Property can be classified into two kinds: *data property* and *object property*. A *data property* is a means to characterize a concept with a data value, while an *object property* is a means to characterize a concept by relating it to another concept.

In order to develop the OntoKin ontology, concepts and properties were extracted from domain experts and published resources. Notable resources include research papers, technical reports for example,<sup>18,19,39</sup> and chemical process knowledge. This involved significant contributions and input from domain experts. As recommended by the domain experts, 20 mechanisms were studied thoroughly. The mechanisms were chosen to have good coverage

of different types of species and reactions. The technical reports provided detailed mathematical descriptions of the thermodynamic model (thermo model from now on), transport model and rate coefficient models.

For the topics of the Reference subdomain (provided in the Identify Scope section), the Bibliographic Ontology (<http://bibliontology.com/>), Dublin Core (<http://dublincore.org/>) and FOAF (<http://xmlns.com/foaf>) were consulted to extract concepts related to Author, Article and Type of Reference. In the end, 57 atomic concepts were identified. The number of identified object properties and data properties is 35 and 100, respectively.

## Analysis

The goal of the analysis phase is twofold: to find clusters of homogenous concepts and to understand properties of concepts. We call each cluster a *haplotype*, as opposed to an *isolate*, which is a concept that does not belong to any cluster. In a haplotype, concepts are connected with *is-a* relations. The amount of isolates is an indication of how hierarchical an ontology is. The bigger the amount of isolates belonging to an ontology, the less hierarchical the ontology is. A haplotype consists of a root concept and one or more child concepts. Each child concept in a haplotype is defined with respect to its direct parent(s). The definition also includes properties of a concept to distinguish from its siblings.

In the analysis phase, a definition is added to each concept of the OntoKin ontology with respect to a parent concept and a distinguishable object property whenever possible. Some examples are provided below:

- *Gas-Phase Reaction: A Chemical Reaction* that occurs in the *Gas Phase*.
- *Arrhenius Reaction: A Gas-Phase Reaction* that is described by an Arrhenius law.
- *Surface Reaction: A Chemical Reaction* that involves an adsorbed species or desorption of products into the gas phase, or adsorption of reactants from the gas phase onto the surface.

- *Sticking Coefficient Reaction: A Surface Reaction* that is described by a sticking coefficient rate model.

## Synthesis

This phase combines the analytical understanding of concepts and properties to create concept hierarchies and/or object property-based relationships between concepts for developing the ontology. A concept hierarchy is created using *is-a* (more specific) or *part-of* (meronymy) relations. The creation of hierarchy and object relationships is exemplified as follows. A *Gas-Phase Reaction* is more specific than a *Chemical Reaction*. This is due to the fact that the former holds all the properties of the latter. In addition, the former also holds an object property which expresses that it belongs to the *Gas Phase*. Similarly, the *Surface Reaction* is more specific than a *Chemical Reaction*. Here the former holds a specialized object property which represents that it belongs to a material. By following this approach, we have developed the OntoKin ontology. The key concepts and object properties of OntoKin are depicted in Figure 1.

To enable semantic interoperability across the chemical space, OntoKin has been mapped to OntoCAPE, which is one of the major ontologies for representing chemical processes. We also looked carefully at other existing chemical ontologies including PubChemRDF and ChEBI. However, their ontological infrastructures were not designed to support the representation of chemical processes, hence, they are not suitable for our purpose. The mapped concepts and properties are denoted by an asterisk (\*) in Figure 1. As shown in this figure, the OntoKin ontology is conceptually divided into 5 modules, each of which is described below.

- *Reaction Mechanism*: A reaction mechanism refers to a set of elementary reactions with specific rate laws, for example to model the combustion of hydrogen. The reaction mechanism module is linked to the phase module to model the fact that a reaction mechanism contains a gas phase and may contain any number of (solid) materials.

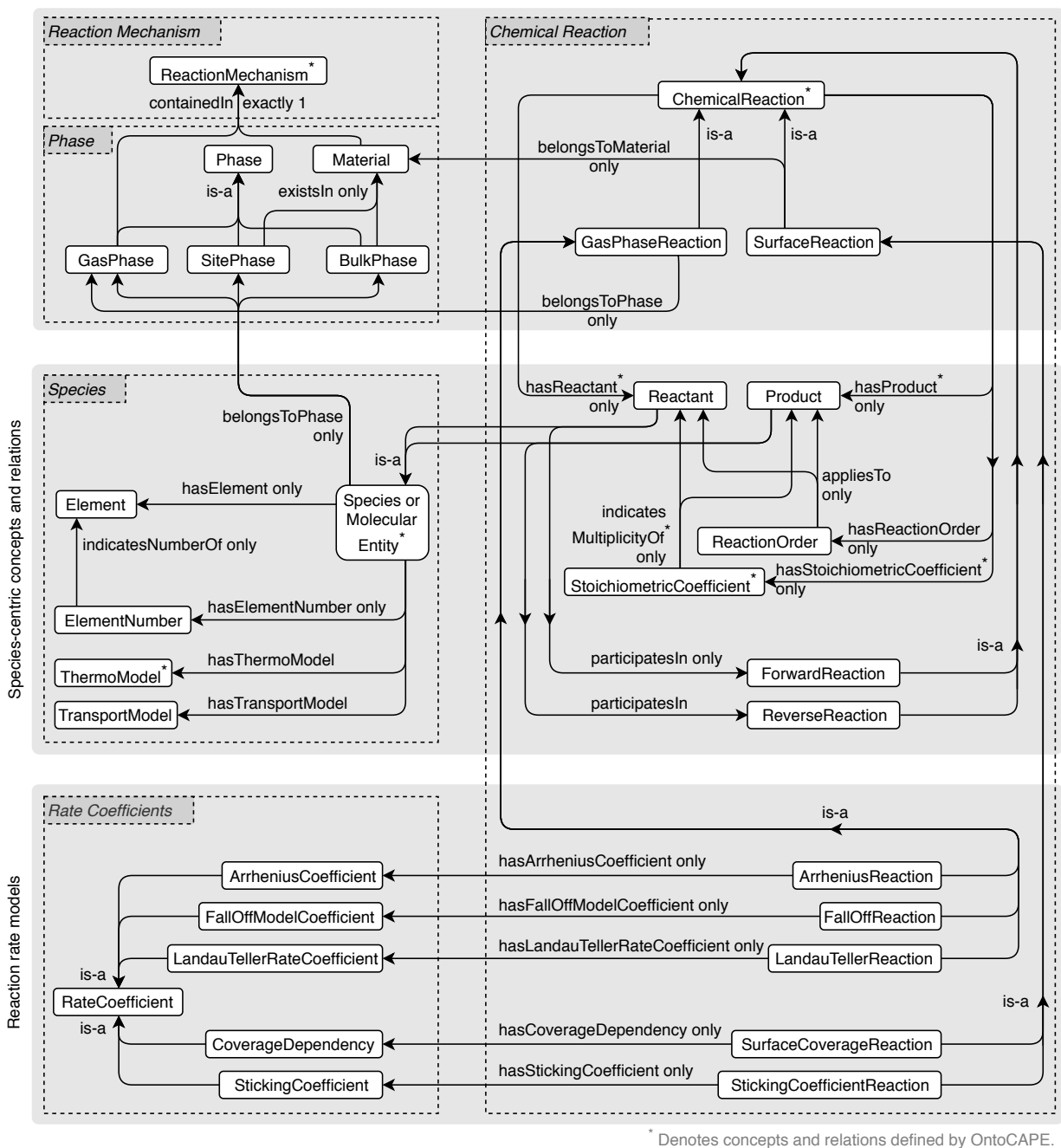


Figure 1: The core concepts and properties of the OntoKin ontology.

A relation is established between each of these concepts and the reaction mechanism using the *containedIn* object property.

- *Phase*: The phase module includes *gas phase*, *site phase* and *bulk phase* as subclasses of *phase*, where *site* and *bulk* phases exist as part of a *material*. Note that the ontological

modelling (or representation) of liquid phases is beyond the scope of the current work.

- *Chemical Reaction*: In the ontology, both the reactants and products are represented as *species* or *molecular entities*. The ontology models gas-phase reactions, where the reactants and products exclusively belong to the *gas phase*, and *surface reactions*, where the reactants and products belong to a combination of the *gas phase* and a *material*. The ontology models each reaction in terms of a *forward reaction* and a *reverse reaction* to allow for the possibility that the reaction is reversible, *i.e.* that it may convert both reactants to products and products to reactants.

The ratio in which the reactants combine and the corresponding proportion of products participating in a chemical reaction are modelled using a *stoichiometric coefficient*. The order of reaction with respect to each reactant participating in the forward reaction and each product participating in the reverse reaction (if one exists) is modelled using a *reaction order*. In other words, the *order of reaction* with respect to an entity is the exponent to which the term for the entity is raised in the reaction rate equation.

The ontology includes a number of common reaction rate models, including *Arrhenius*, *Landau-Teller* and *fall-off* models for gas-phase reactions, and coverage-dependent and sticking coefficient models for surface reactions.

- *Rate Coefficients*: The details of the coefficients that are required depend on the choice of rate model.
- *Species*: The elemental composition of a *species* or *molecular entity* is modelled using an *element number* to describe the multiplicity of each chemical element appearing in the species.

Each *species* or *molecular entity* may be associated with a *thermo model* describing its thermodynamic properties (for example, enthalpy, heat capacity and entropy) and a *transport model* describing its transport properties (for example, viscosity, thermal conductivity and diffusion coefficients).

A subset of data properties of the OntoKin ontology is shown in Table 1. The OntoKin ontology can represent *mechanisms* consisting of a set of *chemical reactions* occurring between *species*. A species is composed of *elements* (for example, water is composed of hydrogen and oxygen) and may exist in different *phases* (for example, water may exist in a *gas phase* or may be adsorbed on some *surface phase*). With respect to a given reaction, the *reactants* are the set of species consumed by the reaction and the *products* are the set of species produced by the reaction. The ontology can represent the *relations* between these concepts as well as *data* and *metadata* about instances of them.

Table 1: Description of some example data properties.

Property Name	Description
hasPreExponentialFactor	The pre-exponential factor in an Arrhenius rate constant
hasTemperatureExponent	The temperature exponent in a modified Arrhenius rate constant
hasActivationEnergy	The activation energy in an Arrhenius rate constant
hasEquation	A string that shows the reaction equation
hasBulkSpeciesDensity	Density of a bulk phase species
hasSiteDensity	The number of sites per unit area on a surface (i.e. a site phase)
hasOccupancy	Number of surface sites occupied by a surface species
hasAtomicMass	Atomic mass of an element

An example *reaction* is  $2\text{H}_2 + 1\text{O}_2 \longrightarrow 2\text{H}_2\text{O}$ .  $\text{H}_2$ ,  $\text{O}_2$  and  $\text{H}_2\text{O}$  are species, of which  $\text{H}_2$  and  $\text{O}_2$  are *reactants* and  $\text{H}_2\text{O}$  is a *product*. The *elements* contained within this set of species are H and O. The *relation* between a chemical entity and a reaction can be expressed as: a reaction having a reactant ( $\text{O}_2$ ), which is composed of a species ( $\text{O}_2$ ) and a stoichiometric coefficient (1), expressing the amount of the species consumed by the reaction. Following are some examples of *data*: the reaction rate (to describe the speed of a reaction) and thermodynamic data (to describe the enthalpy, heat capacity and entropy of a species as a function of temperature and pressure). *Metadata* can be exemplified as the creator of the mechanism and bibliographic data about where it was published.

## Standardization

This is performed in collaboration with domain experts, who based on their experience decide the preferred terms for the concepts which are denoted by multiple terms. For example, in the development of OntoKin, Rate Coefficient is preferred over Reaction Rate Model.

## Formalization

Concepts and properties are formally defined using a logic language as the formalization provides their unambiguous definitions, and allows users to check the consistency of the ontology. *Description Logic (DL)* was selected as the logic language as it is simple and yet powerful enough to fulfill our purpose. Some example formalizations using DL are given in Table 2. The ontology is codified using an ontology representation language. OWL is

Table 2: Formalization of some example concepts using DL.

DL Axiom	English sentence
$\text{GasPhaseReaction} \equiv \text{ChemicalReaction} \sqcap \forall \text{belongsToPhase}.\text{GasPhase}$	A chemical reaction that (always) belongs to (or occurs in) the gas phase.
$\text{ArrheniusReaction} \equiv \text{GasPhaseReaction} \sqcap \forall \text{hasArrheniusCoefficient}.\text{ArrheniusCoefficient}$	A gas-phase reaction that (always) has an Arrhenius reaction rate model.
$\text{SurfaceReaction} \equiv \text{ChemicalReaction} \sqcap \forall \text{belongsToMaterial}.\text{Material}$	A chemical reaction that (always) belongs to a material (or occurs on a surface).
$\text{StickingCoefficientReaction} \equiv \text{SurfaceReaction} \sqcap \forall \text{hasStickingCoefficient}.\text{StickingCoefficient}$	A surface reaction that (always) has a sticking coefficient reaction rate model.

selected as it is sufficient for our purpose. Concepts and *is-a* relations between concepts are represented using the *owl:Class* and *rdfs:subClassOf* constructs, respectively. Object properties are represented using the *owl:ObjectProperty*. The equivalence ( $\equiv$ ), intersection of ( $\sqcap$ ) and union of ( $\sqcup$ ) relations, are represented using *owl:equivalentClass*, *owl:intersectionOf* and *owl:unionOf*, respectively. The universal quantification ( $\forall$ ) and existential quantification ( $\exists$ ) are represented using *owl:allValuesFrom* and *owl:someValuesFrom*, respectively. In terms of expressivity and reasoning, the ontological representation of OntoKin is valid in OWL 1



and OWL 2, such that reasoners compatible with DL SROIQ<sup>40</sup> can respond to reasoning and inference-based queries.

## Quality Assessment

Both the domain experts and ontology experts actively participate to assess the quality of the ontology. The assessment proceeds with the execution of automatic and manual procedures.

## Evaluation

This step includes the measurement of the *ease of query writing*, *types of queries supported* and the *query performance*. The measurement aspects are described below:

- *Ease of query writing*: The amount of minimum number of triples needs to be included in a query to retrieve the intended result. For example, knowing the Internationalized Resource Identifier (IRI) of a chemical reaction, how many triples are required to be written to extract its equation and the phase to which it belongs? Note that some instances are created for the evaluation. Writing only one triple was enough to extract the equation. The triple was the following: *?chemical\_reaction\_iri ontokin:hasEquation ?equation*. Similarly to the reaction equation, the phase extraction was successfully performed with only one triple, which is as follows: *?chemical\_reaction\_iri ontokin:belongsToPhase ?phase*. However, three triples were needed to query the name of the mechanism to which a reaction belongs. Queries which involve more than one triple were investigated further to see if the number of triples could be reduced.
- *Types of queries supported*: For a given set of information modelled in an ontology, how many different types of queries can be performed? For example, following the finding of a chemical reaction in a mechanism, what are the other mechanisms containing the same reaction? Also, what would be the rate coefficients of the same chemical reaction

in all the mechanisms in which it appears?

- *Query performance*: The indicator for measuring query performance is the time duration each query takes between the moment of submission and returning the result. Queries which return results in less than a second are marked as satisfactory. However, queries which do not fit into one second upper limit are investigated further in order to improve their performance by applying the following two modification techniques: the *query modification* and *ontological model modification*. The *query modification* deals with the alternative way of writing the query to achieve the same result, while the *ontological model modification* deals with the change of the already created model.

## Validation

In the validation step, the ontology is assessed from the accuracy and correctness perspectives by applying a combination of a set of manual and automatic processes. Firstly, the logical consistency of the ontological TBox and ABox is verified with Hermit (<http://www.hermit-reasoner.com/>), a reasoner developed for OWL ontologies<sup>41</sup>. Secondly, the domain experts and ontology experts validate the ontological modelling. The experts also validate the correctness of object properties and data properties in terms of their domain (concept) and range (concept or data type).

An archived version of the OntoKin ontology is available via the University of Cambridge data repository at <https://doi.org/10.17863/CAM.35439> and the living version of the ontology is available via the JPS website at <http://www.theworldavatar.com/kb/ontokin/ontokin.owl>.

## Population of the Knowledge Base

A preliminary knowledge base (KB) by integrating OntoKin with ontological representation of 34 publicly available mechanisms obtained from research organizations including universit-

ies and scientific laboratories. The largest mechanism contained approximately 3,000 species and 19,000 chemical reactions. The smallest mechanism, on the other hand, contained as few as 14 species and 33 reactions. The mechanisms can be analysed from the perspective of the level of granularity and the parametric conditions. The level of granularity relates to whether it is a detailed mechanism that attempts to describe the full chemistry of a system, or a reduced mechanism that attempts to mimic the behavior of the full chemistry under a restricted set of conditions. The parametric conditions relate to whether different mechanisms describe things that purport to describe the same physics in the same way.

Figure 2 shows the toolset used to populate the KB. The toolset includes an *ABox Manager* that uses the OntoKin TBox and OWL API (<https://github.com/owlcs/owlapi>) to convert CHEMKIN<sup>18,19</sup> mechanism files to OntoKin ABoxes, and a *KB Generator* that imports the OntoKin TBox and ABox(es) into a KB. The KB was deployed using the RDF4J (<http://rdf4j.org>) triple store. The toolset supports running the conversion and import independently as batch processes.

The ABox Manager was developed using CHEMKIN mechanisms to provide examples of the mapping to the ontology. CHEMKIN mechanisms can consist of up to four files. The main file specifies the chemical elements, the species and reactions that belong to the gas phase, and optionally, thermodynamic data for species in this file. A second file specifies any materials, any site and bulk phases that belong to a material, any species that belong to these phases, any thermodynamic data for these species, and any reactions that belong to a material. These files may be supplemented by thermodynamic and transport data files. These provide data for any species for which thermodynamic and transport data have not been defined. In the current work, the thermodynamic data is in the form of 7-coefficient NASA polynomials<sup>39</sup>. Bidirectional conversion was used to prove that the ABox Manager faithfully preserved the source data, and facilitates backward compatibility.

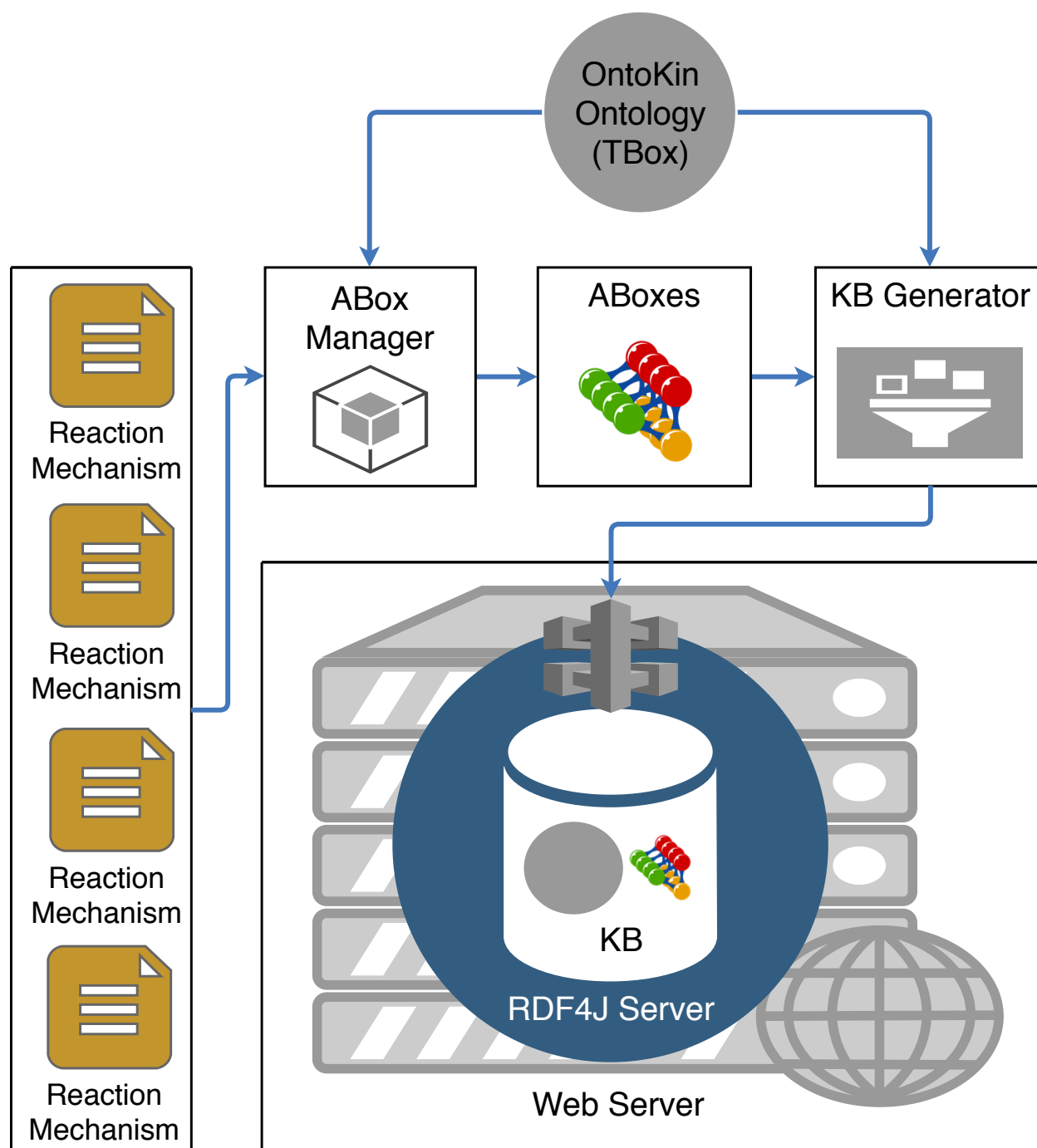


Figure 2: The toolset for generating the OntoKin KB.

## Use Cases

This section introduces some use cases to show how Semantic Web technologies such as the OntoKin ontology and knowledge base can add value. The use cases show examples of how

mechanisms can be uploaded, compared, queried and employed for modelling.

## OntoKin System Development: Mechanism Upload and Querying

OntoKin has been developed to allow any user to upload chemical mechanisms to the OntoKin KB, and to query the KB to retrieve and compare species and reaction data. The following are examples of the types of queries that are of interest to different practitioners:

- Show all of the mechanisms available in the knowledge base.
- Show mechanisms that contain a specific species, e.g. O<sub>2</sub>.
- Show thermodynamic data of a specific species in all of the mechanisms in which the species appears.
- Compare thermodynamic data for a specific species across a set of mechanisms.
- Show mechanisms that contain a specific reaction, e.g. O<sub>2</sub> + N → O + NO.
- Show the rate parameters for a specific reaction in all of the mechanisms in which the reaction appears.
- Compare the rate parameters for a specific reaction in all of the mechanisms in which the reaction appears.

The OntoKin system consists of three main components - a User Interface (UI), a business logic layer and the underlying knowledge base. A web-based UI to demonstrate the OntoKin system is available at the following link: <http://theworldavatar.com/ontokin>. The UI is also shown in Figure 3. The left panel of the UI was designed to enable users to upload mechanisms to the knowledge base and the right panel was designed to support users to query the knowledge base, including new mechanisms added by user.

The business logic layer includes a CHEMKIN to OWL conversion agent, an OWL file consistency checking agent, an OWL file uploading component and a query component.

**OntoKin**

### Select CHEMKIN files to upload:

---

Select a mechanism file to upload:

No file chosen

Select a thermo chemistry file to upload:

No file chosen

Select a surface chemistry file to upload:

No file chosen

Select a transport file to upload:

No file chosen

Provide a name for the mechanism:

### Specify a Query:

---

Select a type of query:

?

Specify the name of a species, e.g. O2. Click the "Search OntoKin" button.

?

[How to search?](#)

**How to search for a species?**

In the 'OntoKin Search' box above, provide a species (e.g. O2) by typing in or pasting, select either 'Show Mechanism(s) Containing Species' or 'Show Thermodynamic Data' or 'Compare Thermodynamic Data' in the 'Select query type' drop-down menu and click on the 'Search OntoKin' button.

**How to search for a reaction?**

In the 'OntoKin Search' box above, provide a reaction (e.g. O2 + N =>O + NO) by typing in or pasting, select either 'Show Mechanism(s) Containing Reaction' or 'Show Arrhenius Rate Constant Parameters' or 'Compare Arrhenius Parameters and Rate Constants' in the 'Select query type' drop-down menu and click on the 'Search OntoKin' button.

Figure 3: A screen-shot depicting a web-based UI developed for the OntoKin System to allow users to upload mechanisms published in CHEMKIN format to the OntoKin KB and to query the KB for species and reaction data.

The conversion agent can assess the validity of a mechanism in CHEMKIN format, which usually contains a mechanism file and a thermodynamic data file and can optionally contain a surface chemistry file and a transport data file. If user-provided files represent a complete mechanism, the converter proceeds with the conversion and reports the success or failure. Following a successful conversion, the consistency checking agent uses the HermiT Reasoner to determine the consistency of logical axioms provided in the OWL file. If the OWL file passes the consistency check, it is uploaded to the OntoKin knowledge base.

The UI allows the user to select from a list of predefined queries. The user may additionally constrain the query by specifying a species or reaction. The UI combines the user input to create a SPARQL (SPARQL Protocol and RDF Query Language) query that is used to search for relevant results in the knowledge base. The results are returned to the UI. Queries for the comparison of thermodynamic data of a species and rate coefficients of a reaction across mechanisms are visualized as line charts. The results of the rest of the queries are shown in tabular form. Figure 4 depicts the results of comparison of thermodynamic data (the heat capacity, enthalpy and entropy) of  $O_2$  across a selected set of mechanisms in the knowledge base.

The system was developed iteratively in consultation with researchers, in particular chemists. As part of the process, the researchers uploaded several mechanisms for testing the mechanism upload feature and comparing thermodynamic data of species and reaction data with already existing mechanisms in the knowledge base. The summary of the evaluation is as follows:

Uploading a mechanism with experimental data represented in CHEMKIN format is self explanatory and supported by the OntoKin system. If a user wants to know the availability of any mechanism for a specific species, the user can check this with the system. Once the user looks up the species of interest, a comparison between thermodata reported in each mechanism can be performed quickly and easily. In this case, the user can identify whether there are inconsistencies between the thermodata used in different mechanisms.

The system allows different mechanisms to be searched for specific reactions and enables easy comparison of the Arrhenius parameters for the reactions. This is important because different methods can be used to estimate the Arrhenius parameters for a given reaction and the evaluation of the parameters can improve over time. With this feature included in the system, changes in these parameters can be easily identified.

Users who are developing new mechanisms can upload their mechanisms to the OntoKin system to crosscheck data for species and reactions of interest. In the future, users will be

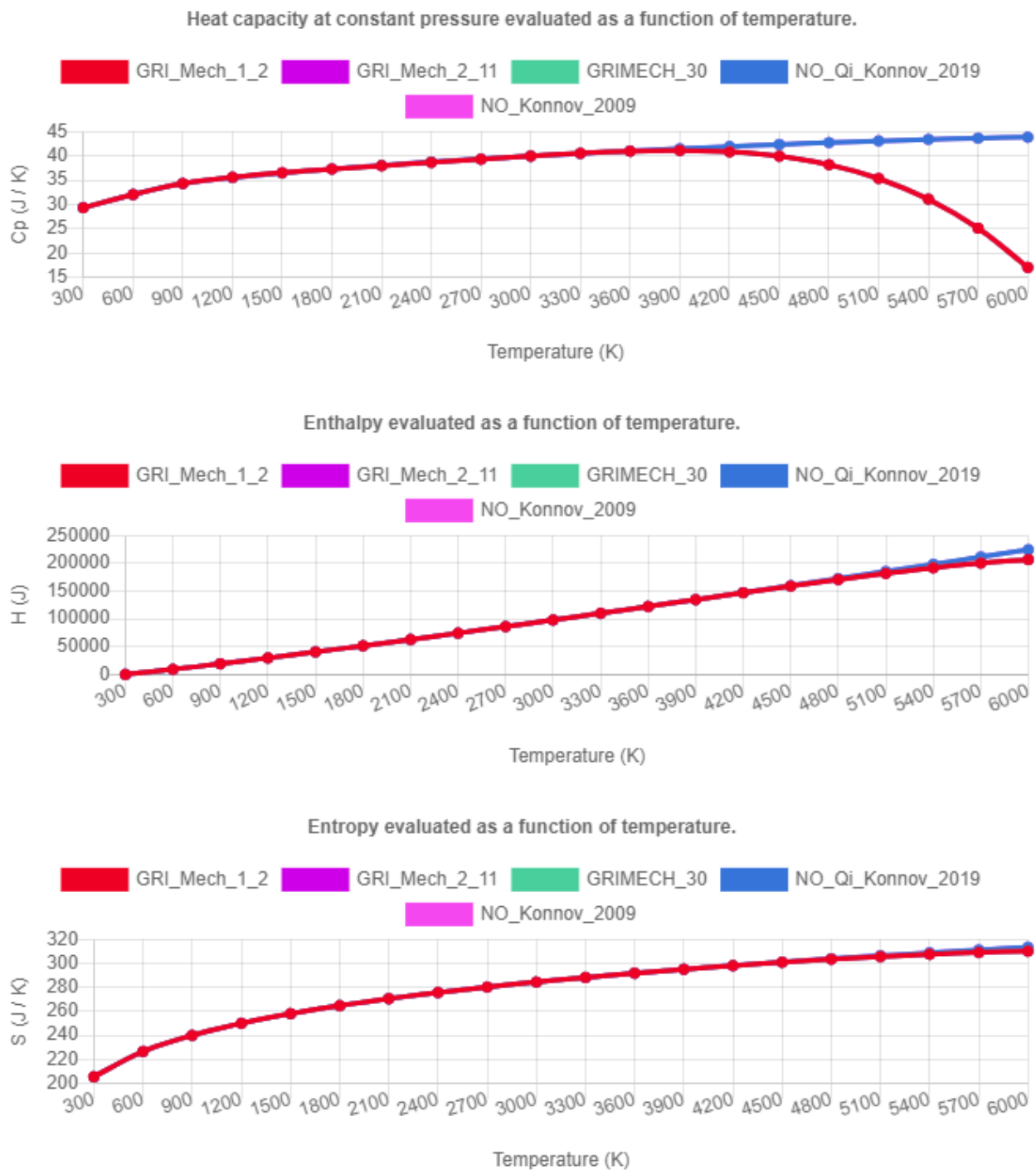


Figure 4: A comparison of thermodynamic data for the species O<sub>2</sub>. Two different trends of heat capacity, enthalpy and entropy are observed across mechanisms at high temperatures.

able to use the knowledge base to retrieve the data that is most suitable for a given purpose and include it in their mechanism.

The combination of experimental and theoretical results provides a sound basis for com-



bustion mechanisms, in which thousands of elementary reactions are involved. These mechanisms attempt to represent the complexity of reactions in engines and combustors at high temperatures and elevated pressures. For example, if a user is interested in modelling ammonia combustion to predict  $\text{NO}_x$  formation, the user will need to choose a mechanism that contains appropriate reactions and kinetic parameters. To compare different mechanisms, the user can employ the UI developed for the OntoKin System. In this case, four mechanisms were uploaded to the OntoKin KB: GRI\_Mech2\_11<sup>42</sup>, GRIMECH\_30<sup>43</sup>, NO\_Konnov\_2009<sup>44</sup> and NO\_Qi\_Konnov\_2019<sup>45</sup>.

A well-known formation pathway is thermal NO, which includes the following key reaction:



The Arrhenius parameters for the reaction in Eq. 1 are shown in Figure 5. It can be observed that the pre-exponential factor has different values between mechanisms. The value reported in GRI\_Mech2\_11 is three orders of magnitude larger than the other mechanisms. It is important to remember that GRI\_Mech2\_11 is the oldest mechanism. The opposite behavior is observed for the temperature exponent, which has a value of zero in GRI\_Mech2\_11. Three different values are obtained for the activation energy. Since the release of GRI\_Mech2\_11, a significant amount of experimental data has allowed an improvement in the estimates of the Arrhenius parameters for the reaction in Eq. 1. As can be observed in Figure 6, different rate constant values are obtained with the different Arrhenius parameters. The estimated uncertainty in the rate constant, for this reaction, is a factor of two<sup>46</sup>.

Another interesting reaction in  $\text{NO}_x$  formation is shown in Eq. 2. This plays a significant role in the NO formation through the prompt route:



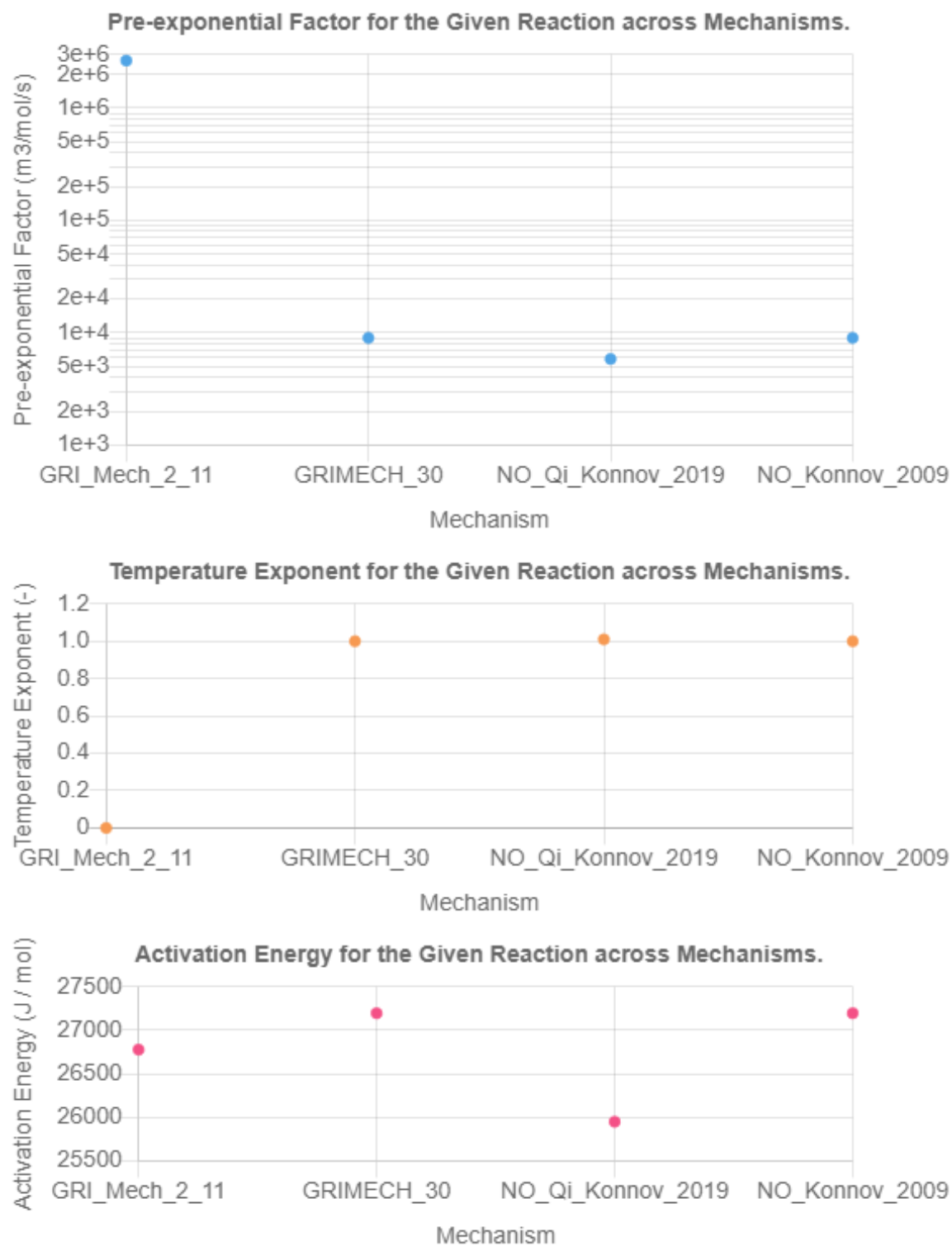


Figure 5: A comparison of the Arrhenius rate constant parameters activation energy, temperature exponent and pre-exponential factor of the reaction  $O_2 + N \Rightarrow O + NO$  across mechanisms.

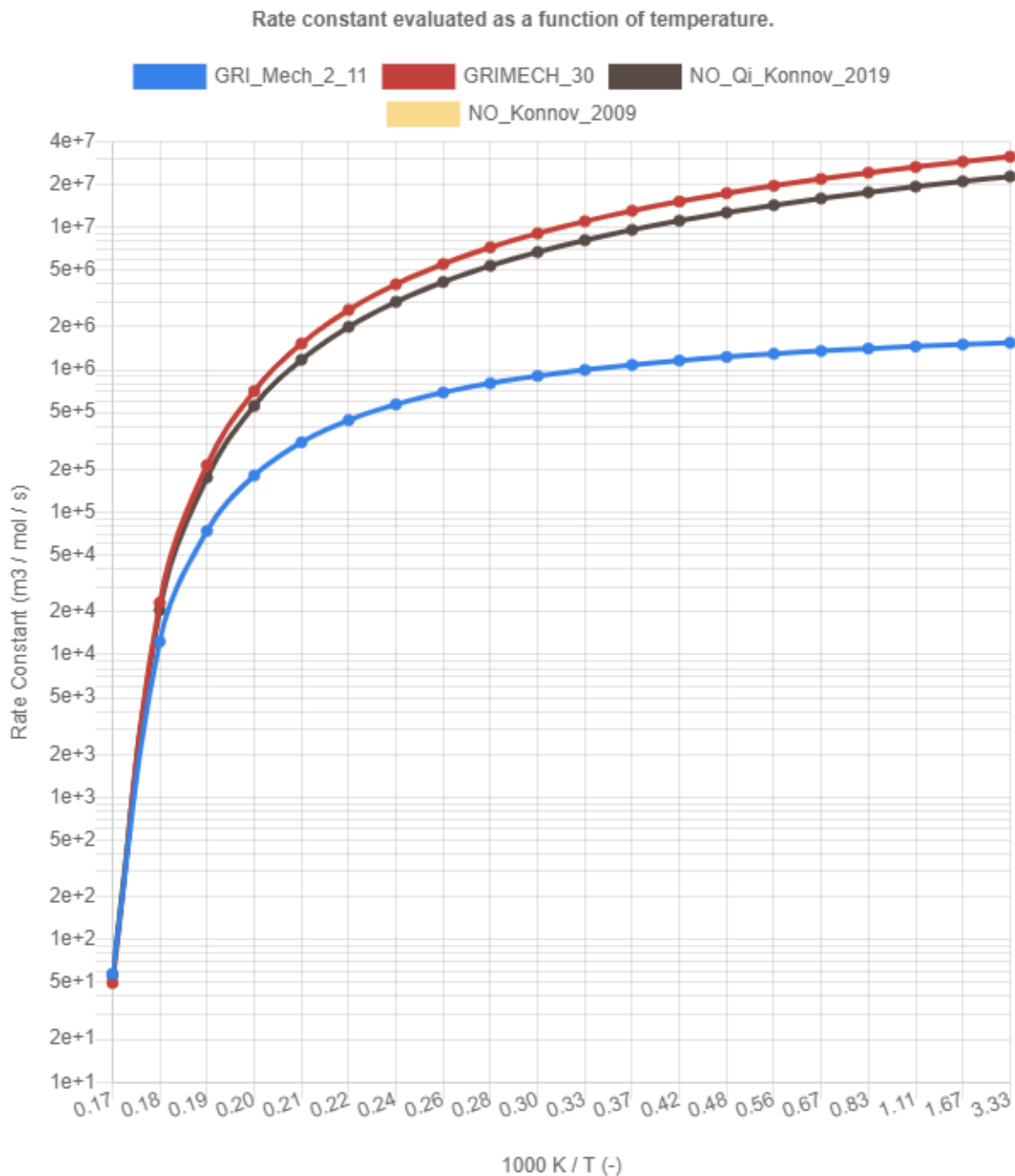


Figure 6: A comparison of rate constants of the reaction  $O_2 + N \Rightarrow O + NO$  across mechanisms.

A comparison of rate constants of this reaction is shown in Figure 7. As can be observed, only the Konnov<sup>44</sup> and Li et al.<sup>45</sup> mechanisms contain the reaction  $CH + N_2 \longrightarrow NCN + H$  (Eq. 2).

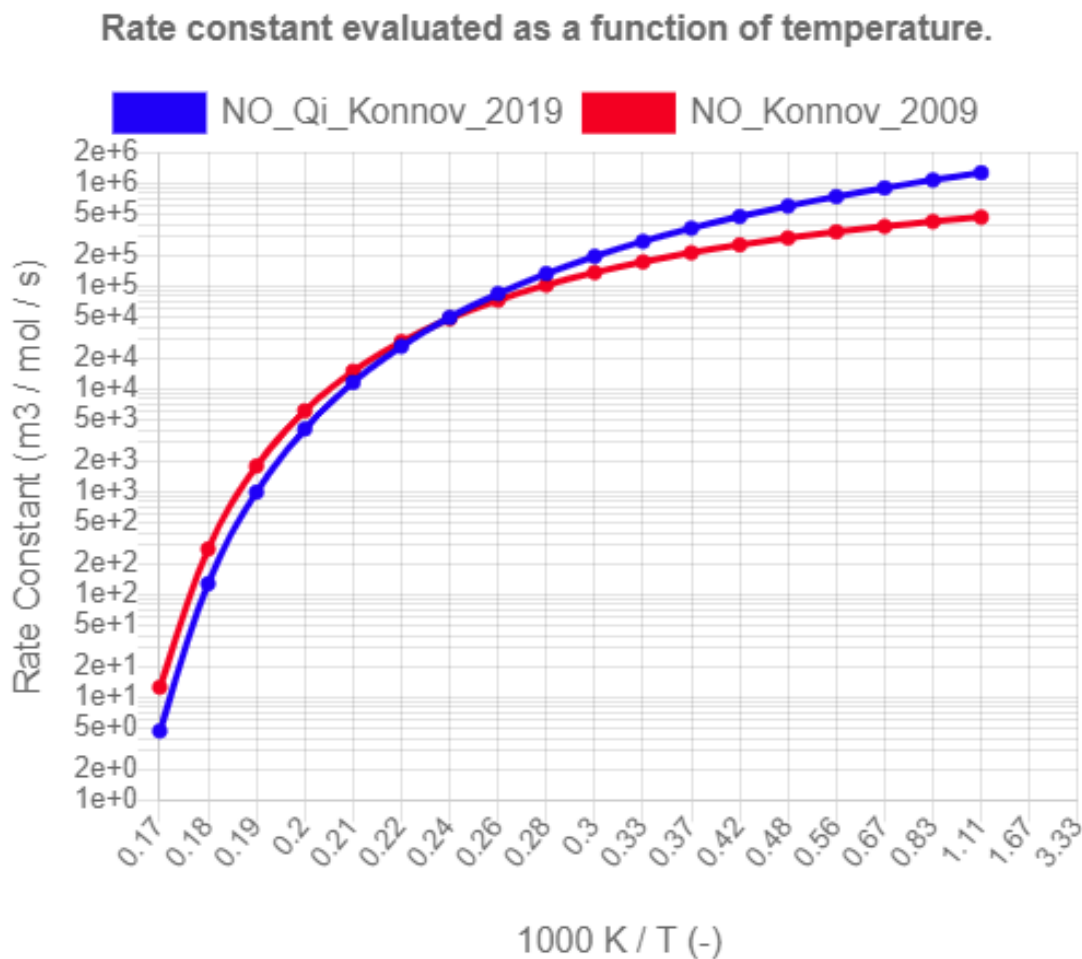


Figure 7: A comparison of rate constants of the reaction  $\text{CH} + \text{N}_2 \Rightarrow \text{NCN} + \text{H}$  across mechanisms.

This means that GRI\_Mech2\_11 and GRIMECH\_30 do not consider the formation of NO through the prompt route and should not be used to model systems in which a significant concentration of CH radicals is expected.

These two reactions (Eq. 1 and 2) are simple examples in which the UI developed for the OntoKin system can be used. See Supporting Information for a list of UI-based queries and their corresponding SPARQL queries as well as how to use the OntoKin system and query the knowledge base. The Ontokin system will allow the comparison of mechanisms containing species and reactions of interest. This will simplify the identification of mechanisms containing different species and the thermodynamic data associated with each species.

In addition, it will allow the comparison of different reactions and their kinetic parameters. This comparison will be an important path to decide which mechanism will suit the researchers better for their applications of interest.

## SRM-ADMS for Simulating Atmospheric Dispersion of Pollutants

Figure 8 shows a cross-domain use case from JPS (<http://www.theworldavatar.com/JPS/?lat=52.076&lon=4.31&zoom=14.5&tilt=0.0&rotation=0.6>). In this example, JPS uses the SRM Engine Suite (<https://cmclinnovations.com/products/srm>), a toolset developed to model the performance of and emissions from internal combustion engines, to estimate the exhaust emissions from fixed diesel generators.

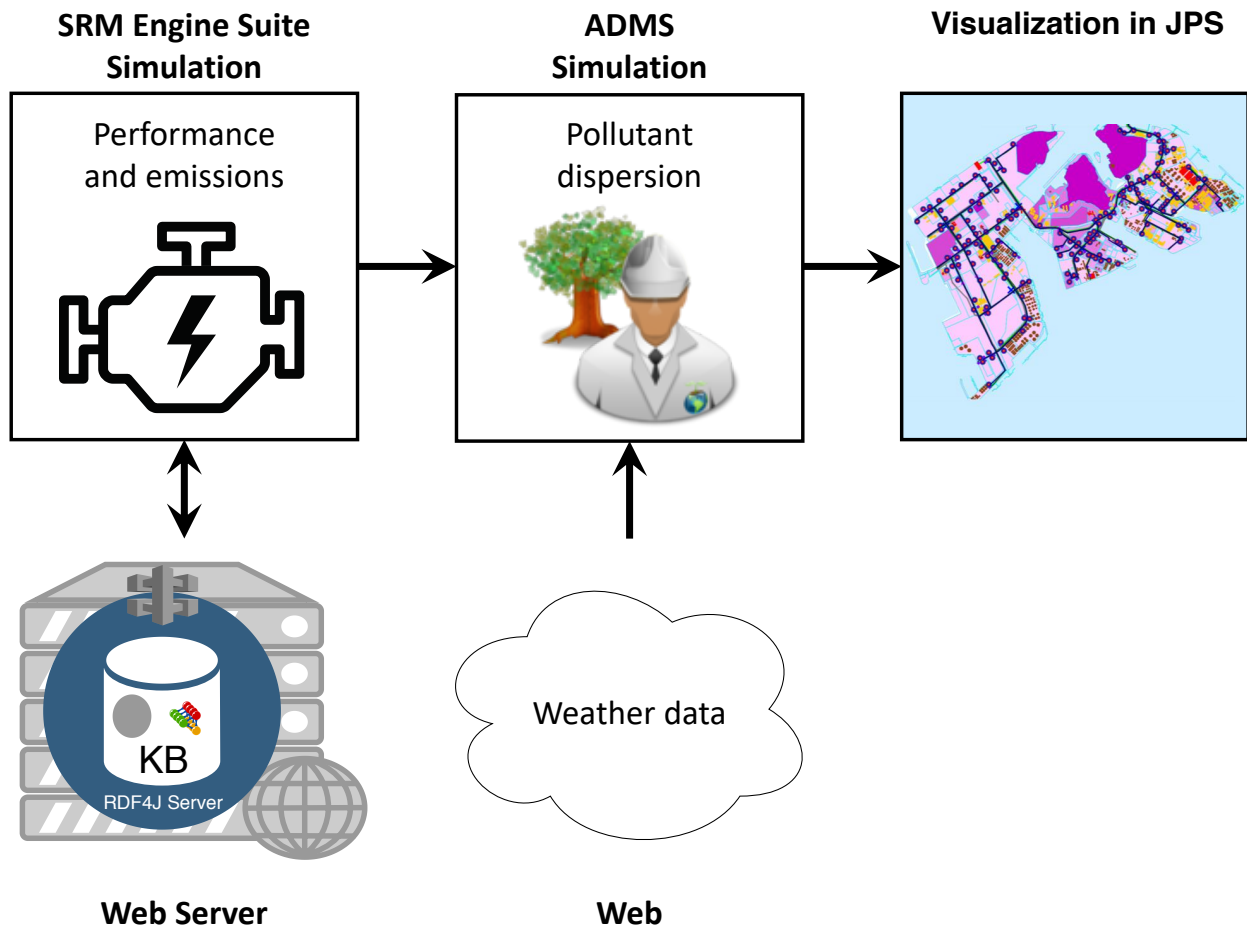


Figure 8: The SRM-ADMS use case.

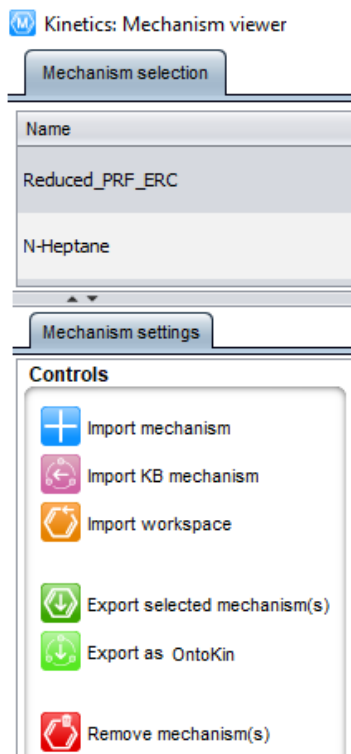
JPS then uses ADMS (<https://cerc.co.uk/environmental-software.html>), the Atmospheric Dispersion Modelling System, to simulate the dispersion of these emissions in the vicinity of each generator. The SRM engine simulations require as an input amongst many others a chemical mechanism which contains species and reactions describing combustion of the fuel used to operate the generators. These mechanisms are retrieved by an agent from the OntoKin KB by querying the JPS knowledge-graph. The ADMS simulations use weather data queried from the World Wide Web. The distribution of the emissions is visualized in JPS as an overlay on a Google Map.

An API has been developed using the RDF4J Model API (<http://docs.rdf4j.org/programming>). The API takes advantage of the choice to represent mechanisms using ontologies (in any of the Semantic Web formats including OWL and RDF). The API enables the SRM to programmatically import mechanisms by sending a request with the IRI of the mechanism to the KB. The KB responds by returning the RDF representation of the mechanism, which can be converted to a form that is readable by the SRM. Here we have exploited the ontologies including OntoKin and OntoCAPE and the KB to provide interoperability between software packages from different domains, enabling them to be used in both autonomous and interactive environments in JPS.

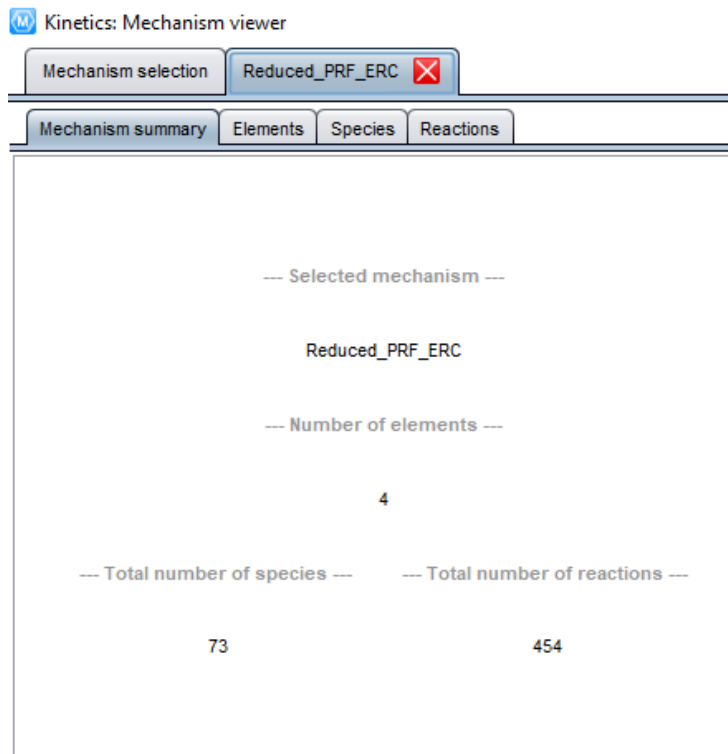
## **Importing and Exporting OWL/RDF Data Using a Mechanism Viewer**

The Mechanism Viewer, originally developed as part of the SRM Engine Suite, is a tool to visualise chemical mechanisms, including the elements, species, thermodynamic and transport data, reactions and rate parameters belonging to the mechanism. The viewer has been extended using the API developed for the SRM-ADMS use case to import and export mechanisms in OWL and RDF formats, and to browse, query and import mechanisms from the KB. The Mechanism Viewer is part of the SRM Engine Suite - a commercial tool<sup>47</sup>. The screen-shots in the subsequent figures are reproduced here with permission.

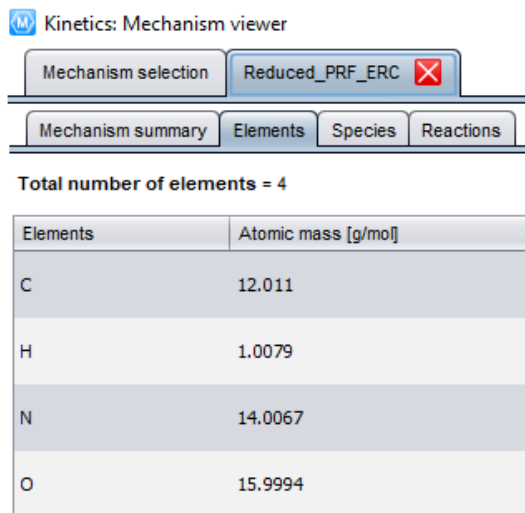
Figure 9 shows the Graphical User Interface (GUI) of the Mechanism Viewer. Figure 9a



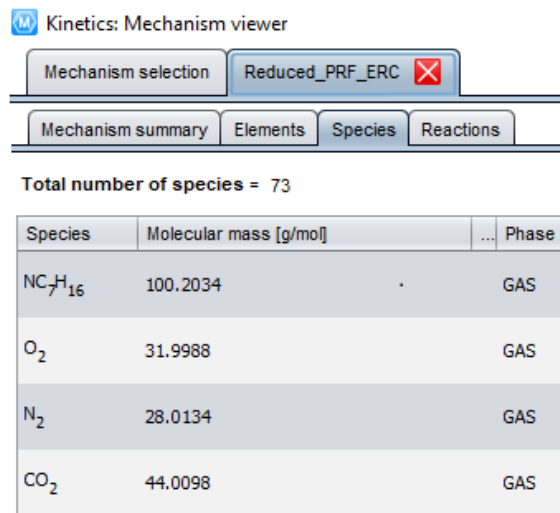
(a) The first page of the Mechanism Viewer user interface.



(b) The summary of a mechanism.



(c) The elements of a mechanism.



(d) A partial list of species of a mechanism.

Figure 9: The visualization of the first page of the Mechanism Viewer. It includes a summary, elements and a partial list of species of a mechanism.

depicts the first page of the GUI showing two mechanisms called Reduced PRF ERC<sup>48</sup> and N-Heptane. The Reduced PRF (Primary Reference Fuel) mechanism was developed by ERC

(Engine Research Centre) at the University of Wisconsin, Madison.

**Thermodynamic data**

	$T_{\min}$ [K]	$T_{\max}$ [K]	$a_1$	$a_2$	$a_3$	$a_4$	$a_5$	$a_6$	$a_7$
1	300.0	1391.0	-1.26836187	0.085435582	-5.2534678...	1.62945721...	-2.0239492...	-25658.6565	35.3732912
2	1391.0	5000.0	22.2148969	0.034767575	-1.1840712...	1.83298478...	-1.0613026...	-34276.0081	-92.3040196

(a) Thermodynamic data of the species  $\text{NC}_7\text{H}_{16}$ .

**Transport data**

Shape: nonlinear

Parameters

Parameter	Value 1	Unit
Lennard-Jones well depth	459.6	K
Lennard-Jones diameter	6.253	Å
Dipole moment	0.0	Debye
Polarizability	0.0	Å <sup>3</sup>
Rotation relaxation number	1.0	-

(b) Transport data of the species  $\text{NC}_7\text{H}_{16}$ .

Kinetics: Mechanism viewer

Mechanism selection: Reduced\_PRF\_ERC

Mechanism summary | Elements | Species | Reactions

Total number of reactions = 454

ID	Equation
(1)	$\text{NC}_7\text{H}_{16} \rightarrow \text{H} + \text{C}_7\text{H}_{15.2}$
(2)	$\text{H} + \text{C}_7\text{H}_{15.2} \rightarrow \text{NC}_7\text{H}_{16}$
(3)	$\text{NC}_7\text{H}_{16} + \text{H} \rightarrow \text{H}_2 + \text{C}_7\text{H}_{15.2}$
(4)	$\text{H}_2 + \text{C}_7\text{H}_{15.2} \rightarrow \text{NC}_7\text{H}_{16} + \text{H}$
(5)	$\text{NC}_7\text{H}_{16} + \text{O} \rightarrow \text{OH} + \text{C}_7\text{H}_{15.2}$
(6)	$\text{OH} + \text{C}_7\text{H}_{15.2} \rightarrow \text{NC}_7\text{H}_{16} + \text{O}$
(7)	$\text{NC}_7\text{H}_{16} + \text{OH} \rightarrow \text{H}_2\text{O} + \text{C}_7\text{H}_{15.2}$
(8)	$\text{H}_2\text{O} + \text{C}_7\text{H}_{15.2} \rightarrow \text{NC}_7\text{H}_{16} + \text{OH}$

(c) Reactions (partial list) of a mechanism.

**Arrhenius parameters**

Parameter	Value 1	Unit
A	1.3E88	K <sup>21.01</sup> /s
n	-21.01	-
$E_A$	583668.0	J/mol

(d) The Arrhenius rate constant parameters of the reaction (1) in Figure 10c.

Figure 10: Visualization of thermodynamic data and transport data of a species, a partial list of reactions and rate constant parameters of a reaction.

In the *Mechanism Settings* area of the figure, it shows the *Import KB mechanism* and *Export as OntoKin* options to allow users to import mechanisms from the KB to the Mechanism Viewer and to convert mechanisms into ontologies, respectively. Figure 9b illustrates the summary of the imported *Reduced PRF ERC* mechanism showing the fact that it has 4 elements, 73 species and 454 reactions. All the elements and their atomic masses are shown in Figure 9c. A partial list of species and their molecular masses are shown in Figure 9d.

Figure 10, on the other hand, demonstrates thermodynamic data, transport data and reaction data. Thermodynamic data and transport data of the species  $\text{NC}_7\text{H}_{16}$  are illustrated



in Figure 10a and Figure 10b, respectively. A partial list of reactions of the PRF ERC mechanism is shown in Figure 10c. The rate constant parameters of the first reaction in the same mechanism is shown in Figure 10d.

In this manner, the Mechanism Viewer is one tool that can be used to facilitate community contribution to the growth of the Semantic Web by providing an easy way to query, retrieve and distribute mechanisms. In the future it is proposed to develop tools to support more advanced interactions with the KB. For example, to compare data between mechanisms, to link to other data sources, to assess the quality of the data, to find and resolve inconsistencies and inaccuracies and to retrieve consistent mechanisms using the best quality data for a given system.

## Conclusions

This paper has presented an ontological model to capture the semantics of chemical kinetic reaction mechanisms, with a particular focus on combustion chemistry. The ontology was developed using an iterative method that divided the development into smaller macro tasks, with each task covering one aspect of the ontology in a given iteration. Learning was carried forward between iterations, contributing to the overall quality of the ontology.

An ABox Manager tool was developed to convert mechanisms from the CHEMKIN format to OWL, and vice versa. Bidirectional conversion was used to ensure full preservation of the source information and facilitates backward compatibility. A KB containing mechanisms from the literature has been developed and deployed using an RDF4J triple store and included in the JPS knowledge graph. The use of the knowledge base was demonstrated for example use cases, ranging from responding to queries to simulation of the atmospheric dispersion of pollutants from combustion processes.

It is hoped that the tools developed in this paper will provide a first step to providing an easy way to query, compare, and retrieve mechanisms via the Semantic Web. In the future

it is proposed to develop tools to support more advanced community involvement, including formulating an ontological model for the unique representation of chemical species, linking to other data sources and identifying the best quality data for a given system.

## Supporting Information

Supporting Information Available: [Queries to retrieve information from the knowledge base via the OntoKin system and SPARQL Endpoint, and some formalizations.]

## Acknowledgements

This work was partly funded by the National Research Foundation (NRF), Prime Minister's Office, Singapore under its Campus for Research Excellence and Technological Enterprise (CREATE) programme, and by the European Union Horizon 2020 Research and Innovation Programme under grant agreement 646121. Markus Kraft gratefully acknowledges the support of the Alexander von Humboldt foundation.

## References

- (1) Berners-Lee, T.; Hendler, J.; Lassila, O. The Semantic Web. *Sci. Am.* **2001**, 28–37.
- (2) Gruber, T. A translation approach to portable ontology specifications. *Knowledge Acquisition* **1993**, 5, 199 – 220.
- (3) Hendler, J. Agents and the Semantic Web. *IEEE Intell. Syst.* **2001**, 16, 30–37.
- (4) Hastings, J.; Magka, D.; Batchelor, C. R.; Duan, L.; Stevens, R.; Ennis, M.; Steinbeck, C. Structure-based classification and ontology in chemistry. *J. Cheminf.* **2012**, 4, 8.

- (5) Hastings, J.; Chepelev, L.; Willighagen, E.; Adams, N.; Steinbeck, C.; Dumontier, M. The Chemical Information Ontology: Provenance and Disambiguation for Chemical Data on the Biological Semantic Web. *PLoS One* **2011**, *6*, 1–13.
- (6) Taylor, K. R.; Gledhill, R. J.; Essex, J. W.; Frey, J. G.; Harris, S. W.; De Roure, D. C. Bringing Chemical Data onto the Semantic Web. *J. Chem. Inf. Model.* **2006**, *46*, 939–952.
- (7) Fernandez-Lopez, M.; Gomez-Perez, A.; Sierra, J. P.; Sierra, A. P. Building a chemical ontology using Methontology and the Ontology Design Environment. *IEEE Intell. Syst.* **1999**, *14*, 37–46.
- (8) Kraft, M.; Mosbach, S. The future of computational modelling in reaction engineering. *Philos. Trans. R. Soc., A* **2010**, *368*, 3633–3644.
- (9) Sankar, P.; Aghila, G. Design and Development of Chemical Ontologies for Reaction Representation. *J. Chem. Inf. Model.* **2006**, *46*, 2355–2368.
- (10) Morbach, J.; Yang, A.; Marquardt, W. OntoCAPE - A large-scale ontology for chemical process engineering. *Eng. Appl. Artif. Intell.* **2007**, *20*, 147–161.
- (11) Degtyarenko, K.; Matos, P.; Ennis, M.; Hastings, J.; Zbinden, M.; McNaught, A.; Alcántara, R.; Darsow, M.; Guedj, M.; Ashburner, M. ChEBI: a database and ontology for chemical entities of biological interest. *Nucleic Acids Res.* **2008**, *36*, D344–D350.
- (12) Hill, D.; Adams, N.; Bada, M.; Batchelor, C.; Berardini, T.; Dietze, H.; Drabkin, H.; Ennis, M.; Foulger, R.; Harris, M.; Hastings, J.; Kale, N.; Matos, P.; Mungall, C.; Owen, G.; Roncaglia, P.; Steinbeck, C.; Turner, S.; Lomax, J. Dovetailing biology and chemistry: integrating the Gene Ontology with the ChEBI chemical ontology. *BMC Genomics* **2013**, *14*, 513.

- (13) Phadungsukanan, W.; Kraft, M.; Townsend, J. A.; Murray-Rust, P. The semantics of Chemical Markup Language (CML) for computational chemistry : CompChem. *J. Cheminf.* **2012**, *4*, 1–16.
- (14) Wang, B.; Dobosh, P.; Chalk, S.; Sopek, M.; Ostlund, N. Computational Chemistry Data Management Platform Based on the Semantic Web. *J. Phys. Chem. A* **2017**, *121*, 298–307.
- (15) Fu, G.; Batchelor, C.; Dumontier, M.; Hastings, J.; Willighagen, E.; Bolton, E. PubChemRDF: towards the semantic annotation of PubChem compound and substance databases. *J. Cheminf.* **2015**, *7*, 34.
- (16) Kim, S.; Thiessen, P.; Bolton, E.; Chen, J.; Fu, G.; Gindulyte, A.; Han, L.; He, J.; He, S.; Shoemaker, B.; Wang, J.; Yu, B.; Zhang, J.; Bryant, S. PubChem Substance and Compound databases. *Nucleic Acids Res.* **2016**, *44*, D1202–D1213.
- (17) Frenklach, M. Transforming data into knowledge – Process Informatics for combustion chemistry. *Proc. Combust. Inst.* **2007**, *31*, 125–140.
- (18) Kee, R. J.; Rupley, F. M.; ; Meeks, E.; Miller, J. A. *CHEMKIN-III: A FORTRAN Chemical Kinetics Package for the Analysis of Gas-Phase Chemical and Plasma Kinetics*; Sandia Natl. Lab. [Tech. Rep.] SAND 96-8216, 1996.
- (19) Coltrin, M. E.; Kee, R. J.; Rupley, F. M.; ; Meeks, E. *SURFACE CHEMKIN-III: A FORTRAN Package for the Analyzing Heterogeneous Chemical Kinetics at a Solid-surface–Gas-phase Interface*; Sandia Natl. Lab. [Tech. Rep.] SAND 96-8216, 1996.
- (20) Eibeck, A.; Lim, M. Q.; Kraft, M. J-Park Simulator: An ontology-based platform for cross-domain scenarios in process industry. *Comput. Chem. Eng.* **2019**, 106586.
- (21) Zhou, X.; Eibeck, A.; Lim, M. Q.; Krdzavac, N. B.; Kraft, M. An agent composition

- framework for the J-Park Simulator - A knowledge graph for the process industry. *Comput. Chem. Eng.* **2019**, *130*, 106577.
- (22) Pan, M.; Sikorski, J.; Kastner, C. A.; Akroyd, J.; Mosbach, S.; Lau, R.; Kraft, M. Applying Industry 4.0 to the Jurong Island Eco-industrial Park. *Energy Procedia* **2015**, *75*, 1536–1541.
- (23) Martin, S.; Weitz, K.; Cushman, R.; Sharma, A.; Lindrooth, R. *Eco-industrial parks: a case study and analysis of economic, environmental, technical and regulatory issues*; Research Triangle Institute: Research Triangle Park Report 6050 FR, 1996.
- (24) Heeres, R.; Vermeulen, W.; Walle, F. Eco-industrial park initiatives in the USA and the Netherlands: first lessons. *J. Cleaner Prod.* **2004**, *12*, 985 – 995.
- (25) Pan, M.; Sikorski, J.; Akroyd, J.; Mosbach, S.; Lau, R.; Kraft, M. Design Technologies for Eco-industrial Parks: From Unit Operations to Processes, Plants and Industrial Networks. *Appl. Energy* **2016**, *175*, 305–323.
- (26) Kastner, C.; Lau, R.; Kraft, M. Quantitative Tools for Cultivating Symbiosis in Industrial Parks; a Literature Review. *Appl. Energy* **2015**, *155*, 599–612.
- (27) Marquardt, W.; Morbach, J.; Wiesner, A.; Yang, A. *OntoCAPE - A Re-Usable Ontology for Chemical Process Engineering*, 1st ed.; Springer-Verlag Berlin Heidelberg, 2010.
- (28) Zhou, L.; Zhang, C.; Karimi, I. A.; Kraft, M. An Ontology Framework towards Decentralized Information Management for Eco-industrial Parks. *Comput. Chem. Eng.* **2018**, *118*, 49–63.
- (29) Zhang, C.; Romagnoli, A.; Zhou, L.; Kraft, M. Knowledge Management of Eco-industrial Park for Efficient Energy Utilization through Ontology-based Approach. *Appl. Energy* **2017**, *204*, 1412–1421.

- (30) Zhou, L.; Pan, M.; Sikorski, J. J.; Garud, S.; Aditya, L. K.; Kleinlanghorst, M. J.; Karimi, I. A.; Kraft, M. Towards an Ontological Infrastructure for Chemical Process Simulation and Optimization in the Context of Eco-industrial Parks. *Appl. Energy* **2017**, *204*, 1284–1298.
- (31) Liu, H.; Lutz, C.; Miličić, M.; Wolter, F. Reasoning About Actions Using Description Logics with General TBoxes. *Logics in Artificial Intelligence*. Berlin, Heidelberg, 2006; pp 266–279.
- (32) Grüninger, M.; Fox, M. Methodology for the Design and Evaluation of Ontologies. IJCAI'95, Workshop on Basic Ontological Issues in Knowledge Sharing, April 13. 1995.
- (33) Uschold, M.; King, M. Towards a methodology for building ontologies. In Workshop on Basic Ontological Issues in Knowledge Sharing, held in conjunction with IJCAI'95. 1995.
- (34) Noy, N. F.; McGuinness, D. L. *Ontology Development 101: A Guide to Creating Your First Ontology*; Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880, 2001.
- (35) Pinto, H.; Martins, J. Ontologies: How can They be Built? *Knowl. Inf. Syst.* **2004**, *6*, 441–464.
- (36) Giunchiglia, F.; Dutta, B.; Maltese, V.; Farazi, F. A Facet-Based Methodology for the Construction of a Large-Scale Geospatial Ontology. *Journal on Data Semantics* **2012**, *1*, 57–73.
- (37) Giunchiglia, F.; Dutta, B.; Maltese, V. From Knowledge Organization to Knowledge representation. *Knowl. Organ.* **2014**, *41*, 44–56.
- (38) Fernandez-Lopez, M.; Gomez-Perez, A.; Juristo, N. METHONTOLOGY: from Onto-

- logical Art towards Ontological Engineering. Proceedings of the AAAI'97 Spring Symposium. Stanford, USA, 1997; pp 33–40.
- (39) McBride, B.; Heibel, S.; Ehlers, J.; Gordon, S. *Thermodynamic Properties to 6000 K for 210 Chemical Substances Involving the First 18 Elements*; NASA Report.
- (40) Horrocks, I.; Kutz, O.; Sattler, U. The Even More Irresistible SROIQ. Proc. of the 10th Int. Conf. on Principles of Knowledge Representation and Reasoning. 2006; pp 57–67.
- (41) Glimm, B.; Horrocks, I.; Motik, B.; Stoilos, G.; Wang, Z. HerMiT: An OWL 2 Reasoner. *J. Autom. Reasoning* **2014**, *53*, 245–269.
- (42) Gardiner, W.; Lissianski, V.; Qin, Z.; Smith, G.; Golden, D.; FRENKLACH, M.; Eiteneer, B.; Goldenberg, M.; Moriarty, C., N. Bowman; Hanson, R.; Song, S.; Schmidt, C.; Serauskas, R. The GRI-Mech<sup>TM</sup> Model for Natural Gas Combustion and NO Formation and Removal Chemistry. 5th International Conference on Combustion Technologies for a Clean Environment. 1999.
- (43) Smith, G.; Golden, D.; Frenklach, M.; Moriarty, N.; Eiteneer, B.; Goldenberg, M.; Bowman, C.; Hanson, R.; Song, S.; Gardiner, W.; Lissianski, V.; Qin, Z. GRI-MECH 3.0. <http://combustion.berkeley.edu/gri-mech/version30/text30.html>.
- (44) Konnov, A. Implementation of the NCN pathway of prompt-NO formation in the detailed reaction mechanism. *Combust. Flame* **2009**, *156*, 2093 – 2105.
- (45) Li, R.; Konnov, A.; He, G.; Qin, F.; Zhang, D. Chemical mechanism development and reduction for combustion of NH<sub>3</sub>/H<sub>2</sub>/CH<sub>4</sub> mixtures. *Fuel* **2019**, *257*, 116059.
- (46) Baulch, D. L.; Bowman, C. T.; Cobos, C. J.; Cox, R. A.; Just, T.; Kerr, J. A.; Pilling, M. J.; Stocker, D.; Troe, J.; Tsang, W.; Walker, R. W.; Warnatz, J. Evaluated Kinetic Data for Combustion Modeling: Supplement II. *J. Phys. Chem. Ref. Data* **2005**, *34*, 757–1397.

- (47) Kraft, M.; Maigaard, P.; Mauss, F.; Christensen, M.; Johansson, B. Investigation of combustion emissions in an HCCI engine – Measurements and a new computational model. *Proc. Combust. Inst.* **2002**, *28*, 1195–1201.
- (48) Wang, H.; Yao, M.; Reitz, R. Development of a Reduced Primary Reference Fuel Mechanism for Internal Combustion Engine Combustion Simulations. *Energy Fuels* **2013**, *27*, 7843–7853.



# Graphical TOC Entry

