Data Modeling and Translation

John Hewson, Bill Pitz, David Leahy

jchewso@ca.sandia.gov, pitz@llnl.gov,
djleahy@sandia.gov

SNL and LLNL

2nd CMCS Workshop

Argonne National Laboratory

June 10-14, 2002

Sample Application

Model Validation

- *General definition of a* **Model**: Set of rules for transforming inputs into outputs.
- **Models** reproduce observable phenomena.
- **Models** are validated using experimental data.
- **Complex models** may require extensive validation and verification.
  - *Many people may contribute* to such validation.
  - Creation of model inputs and comparison of output is *error prone.*
  - Chemical mechanisms are one model type where 100's or 1000's of validation tests are desired.

What might CMCS add to Model Validation?

- **Common data formats** provide a single point of translation.
  - Translate in and out of one format rather than many formats.
- CMCS facilitates automatic input file creation, execution of models and documentation of results. Enables **automation of validation** process.
- CMCS **Pedigree** links methods, models and data sets.
- Translation into **HTML** for discussion with remote colleagues is relatively easy.

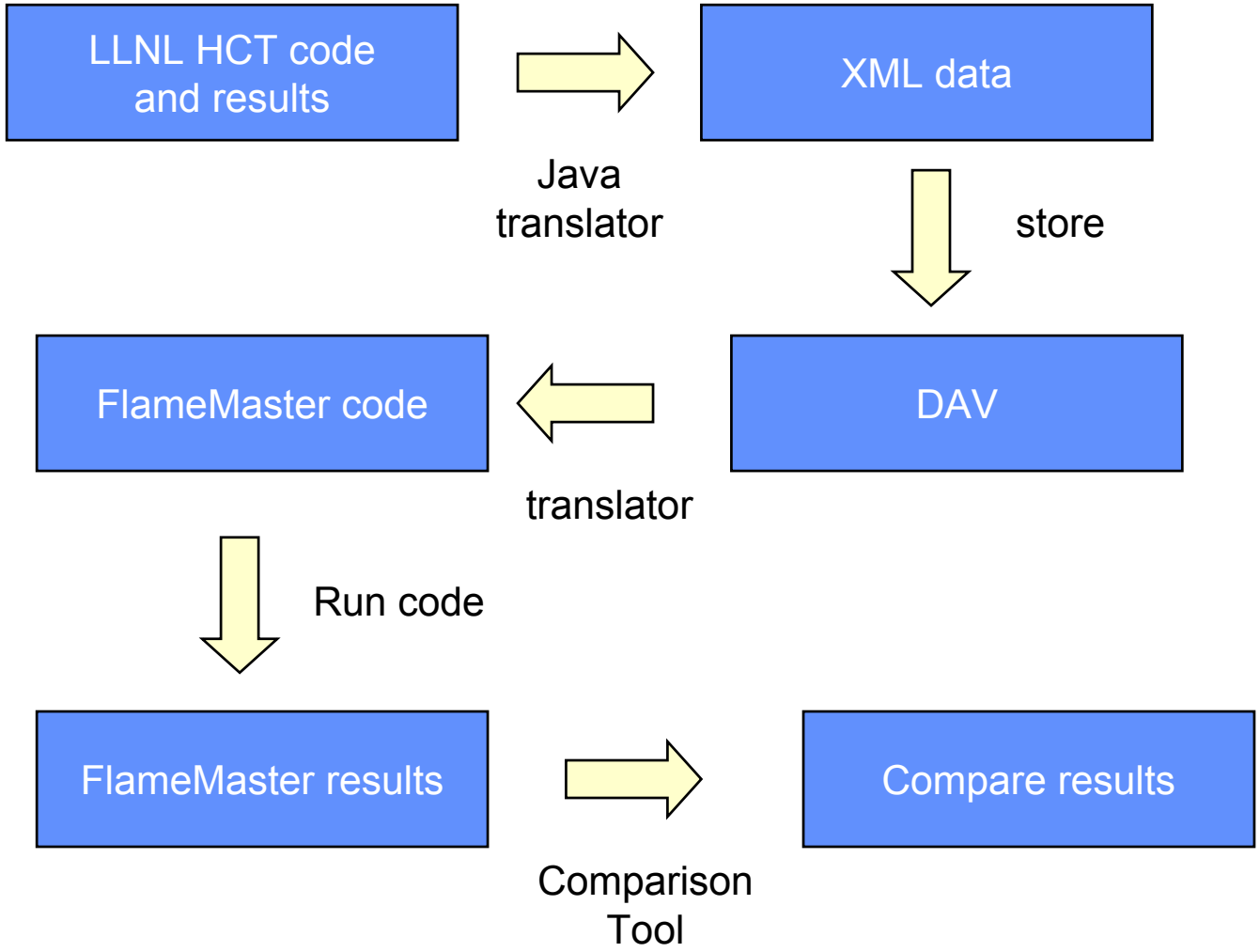
Introduction

- The problem: Collaborating researchers each have access to his/her own preferred set of analysis tools. Difficulties arise because researchers typically expend a great deal of effort in making data and models compatible with their unique tools.
- The solution: CMCS will provide an environment to ease translation by taking advantage of certain technologies described in this poster.
- Issues:
  - Must understand data models in order to translate between formats.
  - Means of translation must be much easier than current practice

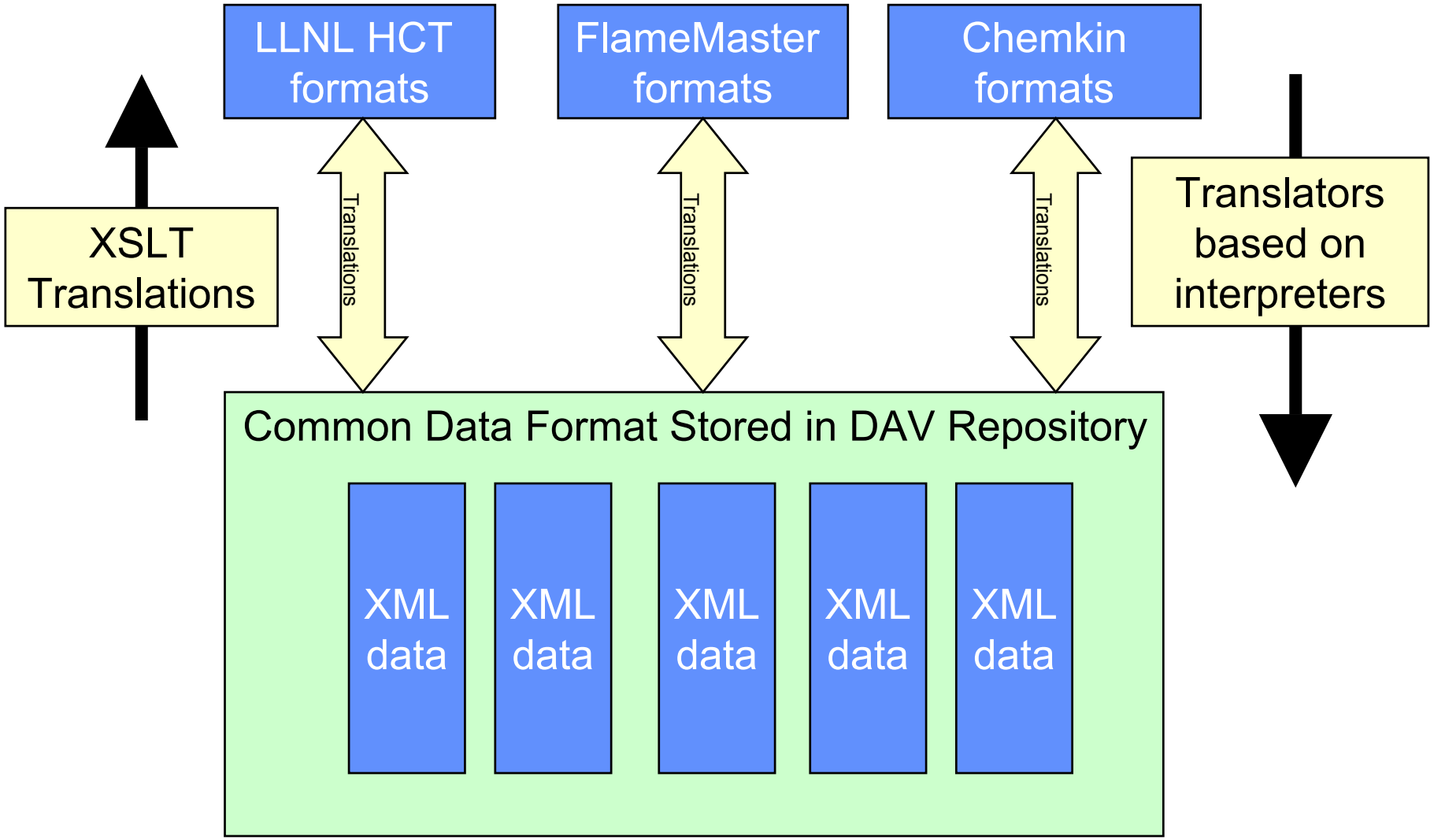
Technologies

- The eXtensible Markup Language (XML) is a widely supported standard for data description.
- XML provides a framework for generating a sufficient descriptive format for models and data.
  - Core elements allow ease of exchange for basic information.
  - XML descriptions are readily extensible allowing change as necessary. Extensions build on core elements.
  - XML elements can be given definitive meaning by associating them with a definition following XML/RDF (Resource Description Framework).
  - Pedigree associated with models will be described primarily with existing frameworks: Dublin Core (DC) and RDF. These indicate relationships between models/data and their creation.
- XSLT (eXtensible Stylesheet Language for Translations) provides a means of translating data from XML to arbitrary formats.

Goal: CMCS infrastructure allows translation of data and comparison of results from different combustion codes



Goal: CMCS infrastructure allows translation of data to and from different combustion codes



Translate LLNL Chemistry Databases to CMCS

LLNL Chemistry Database CDAT

	reaction			frc	frp	frx	rrc	rrp	rrx
1	ch4	=	ch3+h	1.400E+17	0.00	8.840E+04	8.000E+26	-3.00	0.000E+00
2	ch4+h	=	ch3+h2	1.250E+14	0.00	1.190E+04	4.800E+12	0.00	1.143E+04
3	ch4+oh	=	ch3+h2o	3.320E+12	0.00	3.772E+03	5.500E+11	0.00	1.845E+04
4	ch4+o	=	ch3+oh	2.130E+06	2.21	6.480E+03	3.557E+04	2.21	3.920E+03
5	ch3+o2	=	hco+h2o	2.000E+10	0.00	0.000E+00	7.270E+09	0.00	8.304E+04
6	hco+oh	=	co+h2o	1.000E+14	0.00	0.000E+00	2.800E+15	0.00	1.051E+05
7	oh+co	=	h+co2	1.500E+07	1.30	-7.650E+02	1.680E+09	1.30	2.158E+04
8	h+o2	=	o+oh	2.200E+14	0.00	1.679E+04	1.740E+13	0.00	6.770E+02
9	h2+o	=	h+oh	1.800E+10	1.00	8.900E+03	8.300E+09	1.00	6.950E+03
10	o+h2o	=	oh+oh	6.800E+13	0.00	1.835E+04	6.300E+12	0.00	1.100E+03

```
<nomenclature>
<equation>CH4 = CH3+H</equation> </nomenclature>
<dc:relation.references id="1">
hartig, troe, and wagner, 13th comb. symp., 1971;
and warnatz, chem. of comb. reactions, 1984.
</dc:relation.references>
<species>
<reactant id="16">CH4</reactant>
<product id="15">CH3</product>
<product id="1">H</product> </species>
<rate direction="forward" hctType="A+M=B+C+M"
type="modifiedArrhenius">
<A> 1.4E17 <units unit="cm^3 s^-1 mole^-1" />  </A>
<n> 0.0 <units unit="dimensionless" /> </n>
<E> 88400.0 <units unit="cal/mole" /> </E>
```

to XML

(entire database has been converted)

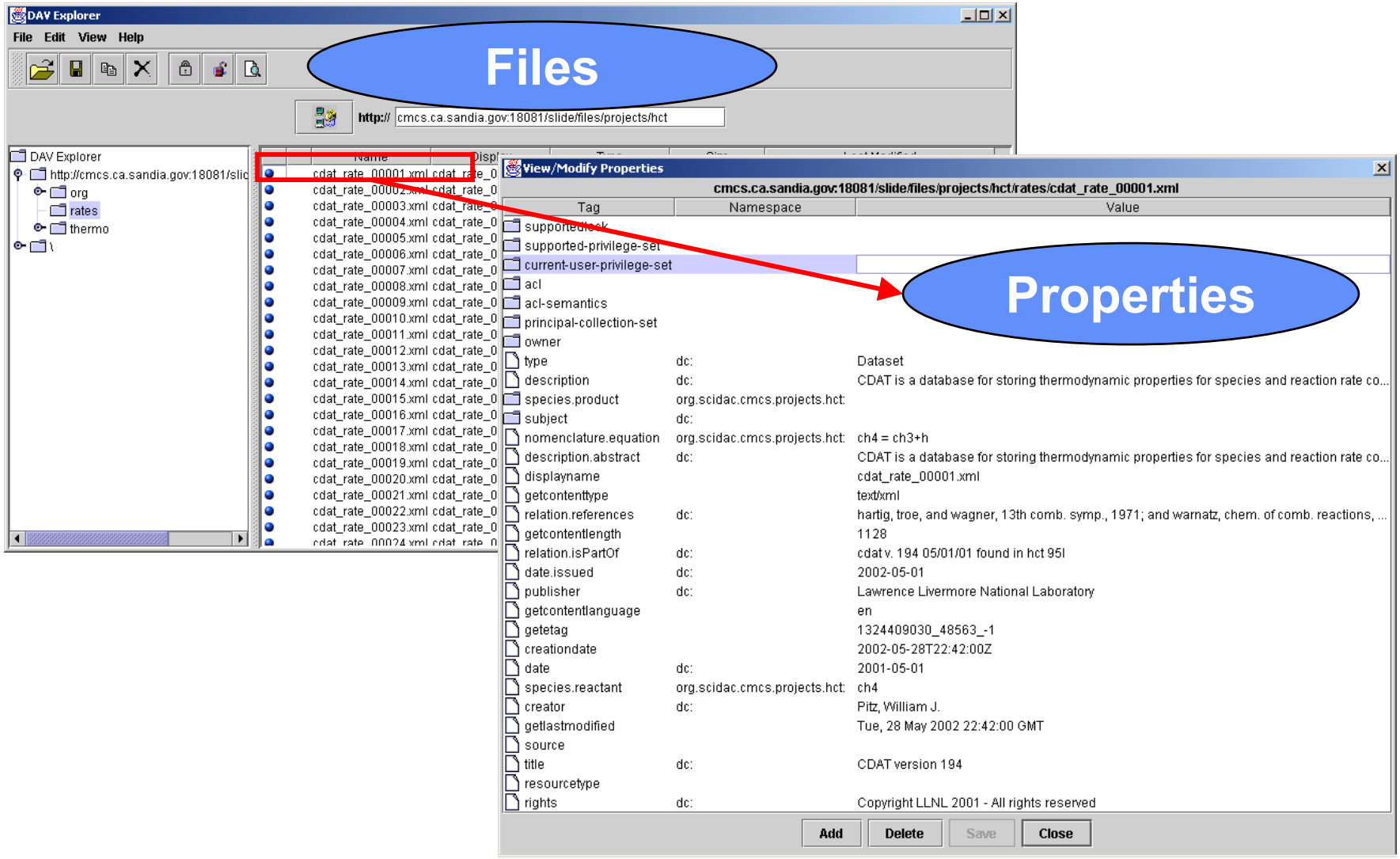
LLNL Species XML

```
<?xml version="1.0" encoding="UTF-8" ?>
<molecule id="1" xmlns:dc="dc:">
<nomenclature>
  <name>H</name>
 </nomenclature>
 <dc:relation.isPartOf>cdat v. 194 05/01/01 found in hct 95l</dc:relation.isPartOf>
<thermo id="1" type="polyfit">
  <coef TempRange="500 3000" type="hct">4.968 0.0 0.0 0.0 0.0 0.0</coef>
</thermo>
<thermo id="1" type="hf298"> 52.1  <units unit="cal/mole" />  </thermo>
<thermo id="1" type="EquilibriumConstantOfFormation">
<coef TempRange="500 3000" type="modifiedArrhenius">
  <A> 1244.0  <units unit="atm units" />  </A>
  <n> 0.0 <units unit="dimensionless" />  </n>
  <E>53720.0 <units unit="cal/mole" />  </E>
</coef>
</thermo>
<molecularWeight>1.0</molecularWeight>
</molecule>
```

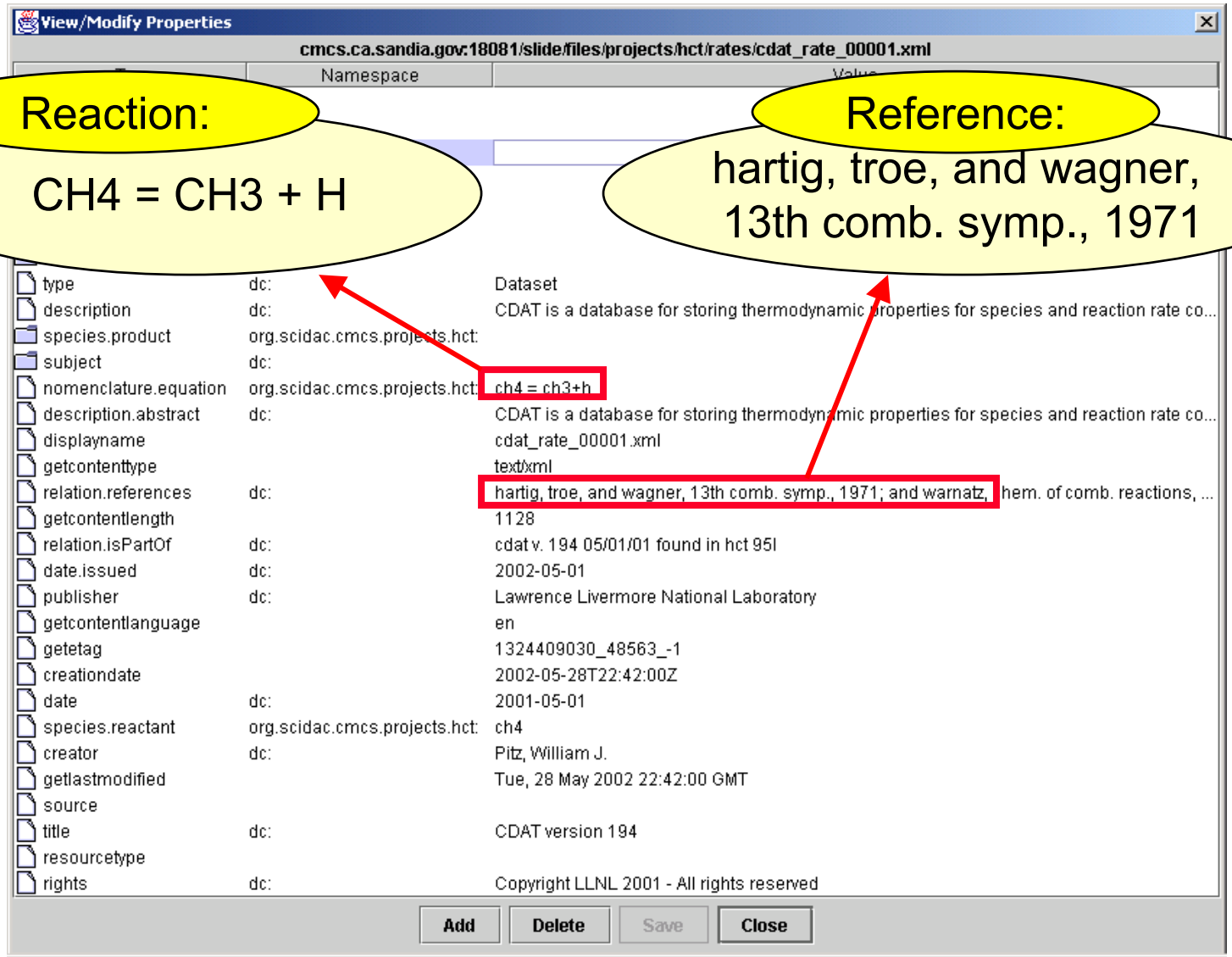
LLNL Reaction XML

```
<?xml version="1.0" encoding="UTF-8" ?>
<reaction id="2" xmlns:dc="dc:">
<nomenclature> <equation>    ch4+h = ch3+h2     </equation>   </nomenclature>
 <dc:relation.isPartOf>cdat v. 194 05/01/01 found in hct 95l</dc:relation.isPartOf>
 <dc:relation.references id="2">baldwin, r. r., hopkins, d. e., norris, a. c., and walker, r. w.,
                                combustion and flame 15, 33 (1970).</dc:relation.references>
<species>
   <reactant id="16">ch4  </reactant>
   <reactant id="1">   h     </reactant>
   <product id="15"> ch3  </product>
   <product id="2">    h2   </product>
 </species>
<rate direction="forward" hctType="A+B=C+D" type="modifiedArrhenius">
  <A> 1.24999998E14 <units unit="cm^3 s^-1 mole^-1" />  </A>
  <n>  0.0                  <units unit="dimensionless" />  </n>
  <E> 11900.0             <units unit="cal/mole" /> </E>
</rate>
<rate direction="reverse" hctType="A+B=C+D" type="modifiedArrhenius">
  <A> 4.7999998E12 <units unit="cm^3 s^-1 mole^-1" /> </A>
  <n>  0.0                 <units unit="dimensionless" />  </n>
  <E> 11430.0             <units unit="cal/mole" />  </E>
 </rate>
 </reaction
```

DAV allows each file to have properties



Each data file has properties that can be searched



Accomplishments

- Draft XML format describing Species and Reactions and Reacting Systems designed.
- Draft schema and object models developed for Reacting System descriptions.
- Translators available from Chemkin, FlameMaster and LLNL data formats into XML formats. Next step is the translation out of XML format into native file formats.
- Entire LLNL Chemistry database has been converted to XML (Thermodynamic parameters for 4,000 species and rate constants for 16,000 reactions)
- Utility program written to produce property data for DAV (David Leahy)
- Representative data put into DAV.