Chemical knowledge representation and access in Wolfram|Alpha and Mathematica

Eric Weisstein
Wolfram Research

ACS Symposium
Chemistry, Data, & the Semantic Web: An Important Triple to Advance Science
March 16, 2016
Outline

- Mathematica, the Wolfram Language, and Wolfram|Alpha
- built-in datasets in Mathematica: ChemicalData, ChemicalFunctionalGroup, ElementData, FoodData, IsotopeData, ProteinData, ThermodynamicData, ...
- Wolfram Language entity framework
- support for units and physical quantities
- compound search methods
- external database access and computation: "PubChem", "ChemSpider", "OpenPHACTS"
- chemistry functionality examples from Wolfram|Alpha (including step-by-step solution)
- future work
Mathematica, the Wolfram Language, and Wolfram|Alpha

- Mathematica is a full-featured computational system which includes:
  - advanced mathematics computation
  - a powerful programming language with 5500+ functions
  - visualization
  - image processing
  - built-in curated data and framework for entity-based representation and computation
  - external device and service connectivity
  - import/export to hundreds of common formats
  - much more
- the Wolfram Language is the knowledge-based programming language inside Mathematica
- Wolfram|Alpha is a free website that accepts natural language queries and returns computational results
- all are integrated together in the latest versions of Mathematica
ElementData

- curated collection of the chemical elements
- elements can be specified by full names, standard abbreviations, or atomic numbers
- available since 2007 (V6.0)

?ElementData

ElementData["name", "property"] gives the value of the specified property for the chemical element "name".
ElementData[n, "property"] gives the specified property for the nth chemical element. >>
ElementData entities

ElementData[]

{hydrogen, helium, lithium, beryllium, boron, carbon, nitrogen, oxygen, fluorine, neon, sodium, magnesium, aluminum, silicon, phosphorus, sulfur, chlorine, argon, potassium, calcium, scandium, titanium, vanadium, chromium, manganese, iron, cobalt, nickel, copper, zinc, gallium, germanium, arsenic, selenium, bromine, krypton, rubidium, strontium, yttrium, zirconium, niobium, molybdenum, technetium, ruthenium, rhodium, palladium, silver, cadmium, indium, tin, antimony, tellurium, iodine, xenon, cesium, barium, lanthanum, cerium, praseodymium, neodymium, promethium, samarium, europium, gadolinium, terbium, dysprosium, holmium, erbium, thulium, ytterbium, lutetium, hafnium, tantalum, tungsten, rhenium, osmium, iridium, platinum, gold, mercury, thallium, lead, bismuth, polonium, astatine, radon, francium, radium, actinium, thorium, protactinium, uranium, neptunium, plutonium, americium, curium, berkelium, californium, einsteinium, fermium, mendelevium, nobelium, lawrencium, rutherfordium, dubnium, seaborgium, bohrium, hassium, meitnerium, darmstadtium, roentgenium, copernicium, ununtrium, flerovium, ununpentium, livermorium, ununseptium, ununoctium}

Length [%]

118

ElementData["Carbon"]

carbon

ElementData["C"]

carbon

ElementData[6]

carbon
ElementData properties and classes

ElementData["Properties"]

{Abbreviation, AdiabaticIndex, AllotropeNames, AllotropicMultiplicities, 
AlternateNames, AlternateStandardNames, AtomicMass, AtomicNumber, 
AtomicRadius, Block, BoilingPoint, BrinellHardness, BulkModulus, 
CASNumber, Color, CommonCompoundNames, CovalentRadius, CriticalPressure, 
CriticalTemperature, CrustAbundance, CrystalStructure, CuriePoint, DecayMode, 
Density, DiscoveryCountries, DiscoveryYear, ElectricalConductivity, 
ElectricalType, ElectronAffinity, ElectronConfiguration, 
ElectronConfigurationString, Electronegativity, ElectronShellConfiguration, 
FusionHeat, GasAtomicMultiplicities, Group, HalfLife, HumanAbundance, 
IconColor, IonizationEnergies, IsotopeAbundances, KnownIsotopes, 
LatticeAngles, LatticeConstants, Lifetime, LiquidDensity, MagneticType, 
MassMagneticSusceptibility, MeltingPoint, Memberships, MeteoriteAbundance, 
MohsHardness, MolarMagneticSusceptibility, MolarVolume, Name, NeelPoint, 
NeutronCrossSection, NeutronMassAbsorption, OceanAbundance, Period, 
Phase, PoissonRatio, QuantumNumbers, Radioactive, RefractiveIndex, 
Resistivity, Series, ShearModulus, SolarAbundance, SoundSpeed, 
SpaceGroupName, SpaceGroupNumber, SpecificHeat, StableIsotopes, 
StandardName, SuperconductingPoint, ThermalConductivity, ThermalExpansion, 
UniverseAbundance, Valence, VanDerWaalsRadius, VaporizationHeat, 
VickersHardness, VolumeMagneticSusceptibility, YoungModulus}

ElementData["Classes"] // Shallow

{ actinides, alkali metals, alkaline earth metals,
  antiferromagnetic elements, chalcogens, conductors, diamagnetic elements,
  f-block elements, ferromagnetic elements, gaseous elements, }
ElementData examples

Find the atomic weight of carbon:

ElementData["Carbon", "AtomicWeight"]
12.0107 u

Isotopes of carbon:

ElementData["Carbon", "KnownIsotopes"]
{8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22}

Find the thermal conductivity of tungsten (in SI units):

ElementData["Tungsten", "ThermalConductivity"]
170. W/(mK)

ElementData["Tungsten", "ThermalConductivity", "Units"]
WattsPerMeterKelvin

Plot melting points of elements as a function of atomic number:

ListLinePlot[Table[ElementData[z, "MeltingPoint"], {z, 118}]]
ElementData in Wolfram|Alpha

WolframAlpha["periodic table"]

Assuming "periodic table" is referring to elements | Use as an invention or a word instead

Input interpretation:
- elements
- periodic table

Results:
ChemicalData

- curated collection of 44,089 chemicals
- chemicals can be specified by their common names such as "Water" or "AceticAcid", registry numbers such as "CID962", IUPAC-like names such as "2Methylpropane", or structure strings
- available since 2007 (V6.0)

? ChemicalData

ChemicalData["name", "property"] gives the value of the specified property for the chemical "name".  
ChemicalData["name"] gives a structure diagram for the chemical with the specified name.  
ChemicalData["class"] gives a list of available chemicals in the specified class.  >>
### ChemicalData entities

```plaintext
ChemicalData[] // Shallow

{ liquid hydrogen, hydrogen, deuterium hydride, liquid helium, helium, deuterium,  
  tritium deuteride, lithium, lithium hydride, lithium deuteride, <<44 079>> }

Length[%] 44 089

ChemicalData["Water"]

\( \text{H}_2\text{O} \)

ChemicalData["H2O"]

\( \text{H}_2\text{O} \)

ChemicalData["CID962"]

\( \text{H}_2\text{O} \)

ChemicalData["CAS7732-18-5"]

\( \text{H}_2\text{O} \)

ChemicalData["Water", "AlternateStandardNames"]

{ DihydrogenMonoxide, DihydrogenOxide, DistilledWater, HydrogenHydroxide }
ChemicalData properties and classes

ChemicalData["Properties"]

\{AcidityConstants, AdjacencyMatrix, AlternateNames, AtomPositions,
AutoignitionPoint, BeilsteinNumber, BlackStructureDiagram, BoilingPoint,
BondTally, CASNumber, CHBlackStructureDiagram, CHColorStructureDiagram,
CIDNumber, Codons, ColorStructureDiagram, CombustionHeat, CompoundFormulaDisplay,
CompoundFormulaString, CriticalPressure, CriticalTemperature, Density,
DielectricConstant, DOTHazardClass, DOTNumbers, EdgeRules, EdgeTypes,
EGECNumber, ElementMassFraction, ElementTally, ElementTypes, EUNumber,
FlashPoint, FormalCharges, FormattedName, GmelinNumber, HBondAcceptorCount,
HBondDonorCount, HenryLawConstant, HildebrandSolubility, InChI, IonEquivalents,
Ions, IonTally, IsoelectricPoint, IsomericSMILES, IUPACName, LogAcidityConstant,
LowerExplosiveLimit, MDLNumber, MeltingPoint, Memberships, MolarMass,
MolarVolume, MolecularFormulaDisplay, MolecularFormulaString, MolecularMass,
MoleculePlot, Name, NFPAFireRating, NFPAHazards, NFPAHealthRating,
NFPALabel, NFPAReactivityRating, NonHydrogenCount, NonStandardIsotopeCount,
NonStandardIsotopeNumbers, NonStandardIsotopeTally, NSCNumber, OdorThreshold,
OdorType, PartitionCoefficient, pH, Phase, RefractiveIndex, Resistivity,
RotatableBondCount, RTECSClasses, RTECSNumber, SideChainAcidityConstant,
SMILES, Solubility, SpaceFillingMoleculePlot, StandardName, StructureGraph,
SurfaceTension, TautomerCount, ThermalConductivity, TopologicalPolarSurfaceArea,
UpperExplosiveLimit, VanDerWaalsConstants, VaporDensity, VaporizationHeat,
VaporPressure, VertexCoordinates, VertexTypes, Viscosity\}

ChemicalData["Classes"] // Shallow

\{\[\[\text{acid anhydrides}\], \[\text{acid halides}\], \[\text{acids}\], \[\text{alcohols}\], \[\text{aldehydes}\],
\[\text{aldoses}\], \[\text{aliphatic ketones}\], \[\text{alkanes}\], \[\text{alkenes}\], \[\text{alkyl halides}\], \text{\textless 128\textgreater }\}\}
ChemicalData examples

```
ChemicalData["Caffeine"]

ChemicalData["Caffeine", "MolecularMass"]
194.191 u

ChemicalData["Caffeine", "ElementTally"]
{{C, 8}, {H, 10}, {N, 4}, {O, 2}}

Total[#2 ElementData[#1, "AtomicMass"] &@@%]
194.191 u

UnitConvert[%, "Kilograms"]
3.22461 \times 10^{-25} \text{ kg}
```
ChemicalData examples

ChemicalData["Caffeine", "MoleculePlot"]

![MoleculePlot of Caffeine]

ChemicalData["Caffeine", "SpaceFillingMoleculePlot"]

![SpaceFillingMoleculePlot of Caffeine]
ChemicalData in Wolfram|Alpha

WolframAlpha["caffeine"]
ProteinData

- curated collection of all 27,479 reference human proteins
- protein sequences are represented as strings of standard single-letter amino acid codes
- available since 2008 (V7.0)

? ProteinData

ProteinData[entity] gives the reference amino acid sequence for the protein entity.
ProteinData[entity, property] gives the value of the specified property for the protein entity.
ProteinData[entity, property, annotation] gives the specified annotation associated with the given property.  ➞
ProteinData entities

ProteinData[] // Shallow

{ alpha 1B-glycoprotein precursor, alpha-2-macroglobulin precursor, N-acetyltransferase 1, 
arylamide acetylase 2, serpin peptidase inhibitor, clade A, member 3 precursor, 
aryacetamide deacetylase, angio-associated, migratory cell protein, 
arylalkylamine N-acetyltransferase, alanyl-tRNA synthetase, 
4-aminobutyrate aminotransferase precursor, <<27469>> }

Length[%]
27479
ProteinData properties and classes

ProteinData["Properties"]
{AdditionalAtomPositions, AdditionalAtomTypes, AtomPositions, AtomRoles, AtomTypes,
BiologicalProcesses, CellularComponents, ChainLabels, ChainSequences,
DihedralAngles, DNACodingSequence, DNACodingSequenceLength, DomainIDs,
DomainPositions, Domains, Gene, GeneID, GyrationRadius, Memberships,
MolecularFunctions, MolecularWeight, MoleculePlot, Name, NCBIAccessions, PDBIDList,
PrimaryPDBID, SecondaryStructureRules, Sequence, SequenceLength, StandardName}

ProteinData["Classes"] // Shallow
{10 formyltetrahydrofolate catabolic process,
11 beta hydroxysteroid dehydrogenase activity, 11 cis retinal binding,
14 alpha glucan branching enzyme activity, 15 hydroxyprostaglandin dehydrogenase nad+activity,
15 hydroxyprostaglandin dehydrogenase nadp+activity,
15 oxoprostaglandin 13 oxidase activity, 1 acylglycerol 3 phosphate o acyltransferase activity,
1 acylglycerophosphocholine o acyltransferase activity,
1 alkyl 2 acetylhydrophosphocholine esterase activity, <<6996>> }

ACS2016-small.nb
ProteinData examples

Find the English name of a protein:

ProteinData["A2M", "Name"]
alpha-2-macroglobulin precursor

Amino acid sequence:

ProteinData["A2M"]
MGKNKLLHPSLVLLLLLVTASVSGLPKQVMVLPSLLHTTTEKGCVLLSYLNETVTSASLESVRGRRSLFTDLEAEN DVLHCVAFAVPKSSSNEEVMFLTVQGPTQEKFARKRTMVKNEDSLVVFQTDKISKGGQPQTVKFRVSMMDNFHPLN EIPLVYIQDPKNRIAQQWQSFQLEGKLQFSFSLSEPFQGYSYKVVQKSGGRTHEPFTVEEFVLPKFQTVTPK IIITLEEPMVSCGLTYGKLPVGPHVTSICRYSDADOCHGEDSFQAFCIEKSQGQLNSHGFQYQQVTKVFLKRKE YENKLHTEAQIEEETVVELSTGRSSEITITIKLSFVKDSHRQGIPFQQVRLVDGKVPGIPKINVFIRGNEANY YSNATTDEHGLVQFSINTTNMGTSLTVRVNYKDRSPCQYQWSEEHEEAAHTAYLFVSPKSVHLLEPMHSLPCG HTQTVAQHYILGNNLKLKQYLMMEGGIVRTGTHGGLVQEDMKGHFSISIQVKSIPDKVAPILLAYAVLPTG TIVIGSAKDYVMNCLANKVDSFSQSPSLPAHHRVTAAPQVSCALRRAVDQVSLMMKPAELSASSYVNNLPEKDL TGFPGLNQDDDECISHRNVYINGITYTPVSSTNEDMISLEDMLKAFNSKIRPKPKMCQPLQQYEMHSPGGTLPV GFYESDMGRGVRHARVHVIEEPHTETVRKYPETWIDLVNMVASQAEGVTPDIETWEKAGAFCLSEDAGLGSST ASLRAFPQFFVELTMYPSVIRGEAFLKATVNLNPKCRVQLASPAFLAVPEKEQAPHICANGRQTVSWAVTA PKSLGNVFNTSAEAESQELVESQPEVPEHGRKDKTVKPKLVEPELEKTFETSFMLPQSCGGEVSEEESLKLPPNV VEESARASVSVLGDILGSAQNTQNLQYMPGCEGQMNMLFAPMNYLNETQQLTPEISKAILGYNLTQYQQLN YKHYDGSSYGTERYGRQGNTWLTAFLKTAQARAYIFIDEAHIQALIWLSQRKQDKNCFRSNGSLNNAAAQGGV EDEVTLSAYTIALIEILVTIVHVRNALSFWTSAWKTQAOYQHGSHYVTKLALAYAFAALGNQDOKKVRKLKSLNEE AVKKDNSVHWERPQKPKAVGHFYEPQQAPSAEEMTYSVLLAYLTAQAPMTSEDLTSATIVKWKFFQNMAYXGGFSSTQ QDTVVALHALSKYGAAFTTRTGGQAQQVTIQSSGQSFFQVDNMNRRLQVQLPPEEGVSMTGCYQLTSLK YNIIPEKEFPFALQGVTLPQTCDERPRAHTSFQIQSLVSSTYSQRSNASMVDFQVGFPEPDKPTVKMLERSNHVSR TEVSSNHLIYDLKQSVQTLSSLFFTVQLDVPVRDLKPAIVKVYDYETDEFAIAYNACPSKDLGNA

StringLength[%]
1474

ProteinData["A2M", "SequenceLength"]
1474 amino acids
ProteinData examples

Ribbon diagram:

```
ribbon = ProteinData["A2M", "MoleculePlot"]
```

Atom positions:

```
(pos = ProteinData["A2M", "AtomPositions"] // Take[#, 5] &

{{ 418.3 pm, -2427.4 pm, -1083.8 pm },
 { 436.5 pm, -2370.9 pm, -1007.8 pm },
 { 418.4 pm, -2393.4 pm, -987.3 pm },
 { 473.2 pm, -2477.6 pm, -1144.3 pm },
 { 479.2 pm, -2550.5 pm, -1197.5 pm }}
```
Plot using atom positions and van der Waals radii:

\[
\text{With[\{p = "A2M", cr = Dispatch@ColorData["Atoms", "ColorRules"]\},}
\]
\[
\text{\text{Graphics3D[\{Specularity[White, 30], MapThread[\{\#2 / cr,
\text{ Sphere[\#1, QuantityMagnitude@ElementData[\#2, "VanDerWaalsRadius"]\} \&,
\text{ \{QuantityMagnitude/@pos, Flatten[ProteinData[p, "AtomTypes"]]\}]\}],
\text{Boxed \to False, Lighting \to "Neutral", ViewPoint \to \{-0.2, -0.1, 3.4\}\}]\}]}}
\]
ProteinData in Wolfram\|Alpha

WolframAlpha["A2M protein"]
ThermodynamicData

- curated collection of 85 substances
- available since 2014 (V10.0)

? ThermodynamicData

```
ThermodynamicData["name", "property"] gives the value of the specific property for the substance "name".
ThermodynamicData["name", "property", {"Temperature" -> quantity1, "Pressure" -> quantity2}] gives the
value of the specific property for the substance "name" at the specified temperature and pressure. >>
```
ThermodynamicData entities

ThermodynamicData[]

Length[%]
85
ThermodynamicData examples

Simple example:

\texttt{ThermodynamicData["Air", "Density", \{"Pressure" \rightarrow \texttt{Quantity}[0.2, "MegaPascals"],
   "Temperature" \rightarrow \texttt{Quantity}[300, "Kelvins"]\}]}

2.32335 kg/m³
**ThermodynamicData examples**

Phase diagram of water:

```plaintext
pressures = Quantity[Table[10^x, {x, -3, 9, 0.1}], "Pascals"];

vaporliquidLine = ThermodynamicData["Water", "LiquidVaporPhaseBoundary", {"Pressure" -> pressures}];
solidliquidLine = ThermodynamicData["Water", "SolidLiquidPhaseBoundary", {"Pressure" -> pressures}];
solidvaporLine = ThermodynamicData["Water", "SolidVaporPhaseBoundary", {"Pressure" -> pressures}];

triplePoint = {ThermodynamicData["Water", "TriplePointTemperature"], ThermodynamicData["Water", "TriplePointPressure"]

{ 273.16 K, 612.48 Pa }

ListLogPlot[Join[Transpose[#, pressures]] &/@ {vaporliquidLine, solidliquidLine, solidvaporLine}, {{{triplePoint}}}, PlotRange -> All, Joined -> True, Frame -> True, FrameLabel -> Automatic, PlotLegends -> {"liquid-vapor boundary", "solid-liquid boundary", "solid-vapor boundary", "triple point"}, PlotStyle -> {Green, Blue, Orange, {Thickness[.03], Red}}]
```

![Phase diagram of water](image-url)


ThermodynamicData examples

Sketch the region in the T-V phase diagram where the liquid and gas forms of nitrogen coexist:

\[
\{T_t, T_c\} = \text{ThermodynamicData["Nitrogen", 
\{"TriplePointTemperature", "CriticalTemperature"\]};
\]

Calculate the T-V curves along the phase boundaries:

\[
\text{TVCurve}[\{p_{\text{min}}, p_{\text{max}}, \Delta p\}, \text{boundary}_-, f_] := \\
\text{With}\left[\{\text{ps} = \text{Quantity}\left[10^\text{\text{Table}}[p, \{p, p_{\text{min}}, p_{\text{max}}, \Delta p\}], \text{"Pascals"}\}\right], \\
\text{Cases}[\#, \{\_\text{Quantity} ..\}] \& \& \\
\{\{\text{ThermodynamicData["Nitrogen", "SpecificVolume"}, \\
\{"Pressure" \to (1 + f \times 0.001) \#\text{1}, "Temperature" \to \#\text{2}\], \#\text{2}\} \& \& \& \\
\text{Transpose}[\{\text{ps}, \text{ThermodynamicData["Nitrogen", boundary, \{"Pressure" \to \text{ps}\}]\}]\}\}
\]
\]

\text{TVCurveVapor} = \text{TVCurve}[\{4.1, 6.8, 0.025\}, \text{"LiquidVaporPhaseBoundary"}, -1];
\text{TVCurveLiquid} = \text{TVCurve}[\{4.1, 6.6, 0.025\}, \text{"LiquidVaporPhaseBoundary"}, +1];
\text{curveSV} = \text{TVCurve}[\{1, 4.099, 0.025\}, \text{"SolidVaporPhaseBoundary"}, -1];
\text{curveLV} = \text{TVCurve}[\{4.1, 9, 0.025\}, \text{"SolidLiquidPhaseBoundary"}, -1];

\text{TVPointc} = \{\text{ThermodynamicData["Nitrogen", "SpecificVolume"}, \\
\{"Pressure" \to 1.001 \text{ThermodynamicData["Nitrogen", "CriticalPressure"}], \\
\{"Temperature" \to T_c\}, T_c\};
\]

Show the vapor/liquid coexistence region in gray, the triple point line in blue and the critical point in red:
With[{g = Log10[QuantityMagnitude[#]] &},
Graphics[{Lighter[Gray, 0.6],
  FilledCurve[
    {Line[Join[g[TVCurveLiquid], g[{TVPointc}], Reverse[g[TVCurveVapor]]]]},
    Black, Line[g[Append[curveSV, First[TVCurveVapor]]]], Line[g[curveLV]],
    Text[
      Column[{
        Subscript["N", 2], "vapor / liquid"}, Alignment -> Center], {-2., 1.9}],
    Blue, Thickness[0.01], Line[g[{{TVCurveLiquid[[1, 1]], Tt},
          {TVCurveVapor[[1, 1]], Tt}}]],
    Red, PointSize[0.02], Point[g[TVPointc]], Frame -> True, AspectRatio -> 1, FrameLabel ->
    {HoldForm[Log10["specific volume"/Quantity[None, "Meters"^3/"Kilograms"]]],
     HoldForm[Log10["temperature"/Quantity[None, "Kelvins"]]]}]]
Wolfram Language entity framework: Entity and EntityProperty

torus = Entity["Solid", "SolidTorus"]

InputForm[torus]
Entity["Solid", "SolidTorus"]

CanonicalName[torus]
SolidTorus

EntityTypeName[torus]
Solid

torus["Properties"]

{ alternate names , associated people , boundary , boundary mesh representation ,
  boundary surface , bounding box corners , centroid , circumcenter ,
  circumradius , circumsphere , classes , concave , constant width , convex ,
  convexity coefficient , cross sections , generalized diameter , 3-D graphics ,
  height , hexahedron , image , implicit region , incenter , defining inequalities ,
  moment of inertia tensor , inradius , insphere , lateral surface area , CDF of lengths ,
  mean line segment length , PDF of lengths , mean cylindrical radius , mean spherical radius ,
  mean square cylindrical radius , mean square spherical radius , mesh representation ,
  name , parallelepiped , parametric region , polyhedron , region , regular polyhedron ,
  related Wolfram Language symbols , slant height , solid of revolution , sport objects ,
  surface area , tetrahedron , mean tetrahedron volume , variable constraints ,
  variable descriptions , variables , number of vertices , vertices , volume }
Wolfram Language entity framework: EntityValue

EntityValue[Entity["Solid", "SolidTorus"], "SurfaceArea"]
Function[{\(\hat{a}, \hat{c}\)}, 4 \(\pi\)^2 \(\hat{a}\) \(\hat{c}\)]

\textbf{torus (solid)} ["Volume"]
Function[{\(\hat{a}, \hat{c}\)}, 2 \(\pi\)^2 \(\hat{a}\)^2 \(\hat{c}\)]

\textbf{ellipse (lamina)} ["area"]
Function[{\(\hat{a}, \hat{b}\)}, \(\pi\) \(\hat{a}\) \(\hat{b}\)]

\textbf{ellipse (plane curve)} ["notation table"]
Function[{a, b}, \begin{array}{l}
\text{a semimajor axis length} \\
\text{b semiminor axis length}
\end{array}]
Function[{\(\hat{a}, \hat{b}\)}, \(\pi\) \(\hat{a}\) \(\hat{b}\)][r, r]
\(\pi\) \(r\)^2
Wolfram Language entity framework: Implicitly defined entities

Chemicals with small molar masses:

Entity["Chemical", {"MolarMass" -> TakeSmallest[5]}]["Entities"]

\{(liquid hydrogen, hydrogen, deuterium hydride, liquid helium, helium)\}

Grid[{{#, ChemicalData[#, "MolarMass"]} & /@ %, Dividers -> All}]

<table>
<thead>
<tr>
<th>Chemical</th>
<th>Molar Mass</th>
</tr>
</thead>
<tbody>
<tr>
<td>liquid hydrogen</td>
<td>2.01588 g/mol</td>
</tr>
<tr>
<td>hydrogen</td>
<td>2.01588 g/mol</td>
</tr>
<tr>
<td>deuterium hydride</td>
<td>3.02204 g/mol</td>
</tr>
<tr>
<td>liquid helium</td>
<td>4.002602 g/mol</td>
</tr>
<tr>
<td>helium</td>
<td>4.002602 g/mol</td>
</tr>
</tbody>
</table>

Chemicals with high boiling points:

Entity["Chemical", {"BoilingPoint" -> TakeLargest[5]}]["Entities"]

\{(tungsten carbide, tungsten, rhenium, tantalum, hafnium oxide)\}

Grid[{{#, ChemicalData[#, "BoilingPoint"]} & /@ %, Dividers -> All}]

<table>
<thead>
<tr>
<th>Chemical</th>
<th>Boiling Point</th>
</tr>
</thead>
<tbody>
<tr>
<td>tungsten carbide</td>
<td>6000.°C</td>
</tr>
<tr>
<td>tungsten</td>
<td>5660.°C</td>
</tr>
<tr>
<td>rhenium</td>
<td>5596.°C</td>
</tr>
<tr>
<td>tantalum</td>
<td>5425.°C</td>
</tr>
<tr>
<td>hafnium oxide</td>
<td>5400.°C</td>
</tr>
</tbody>
</table>

Chemical with high molar masses and boiling points:

Entity["Chemical",

\{("MolarMass" -> TakeLargest[500], "BoilingPoint" -> TakeLargest[500])\}]["Entities"]

\{(pentabromophenyl ether, glycerol tristearate)\}
Grid[{{#, ChemicalData[#, "MolarMass"], ChemicalData[#, "BoilingPoint"]} & /@ %, 
Dividers → All}

<table>
<thead>
<tr>
<th>Compound</th>
<th>Molar Mass</th>
<th>Boiling Point</th>
</tr>
</thead>
<tbody>
<tr>
<td>pentabromophenyl ether</td>
<td>959.17 g/mol</td>
<td>573 °C</td>
</tr>
<tr>
<td>glycerol tristearate</td>
<td>891.480 g/mol</td>
<td>812 °C</td>
</tr>
</tbody>
</table>
Wolfram Language entity framework: Entity copies, groups, and instances

Compute with multiple copies of an entity:

EntityValue[EntityCopies[Entity[Earth, planet], 2], "Mass"]

$1.1944397 \times 10^{25}$ kg

Convert mass of a substance to volume:

EntityValue[EntityInstance[Entity[Oxygen, chemical], 1 kg], "AbsoluteVolume"]

699.8 L

Obtain total mass of a combination of chemicals specified in different ways:

EntityValue[EntityGroup[EntityInstance[Entity[Carbon dioxide, chemical], 3 mol], EntityInstance[Entity[Oxygen, chemical], 800 L]], "AbsoluteMass"]

1275. g
Wolfram Language entity framework: Entity types

EntityValue[]
Length[%]
235
FoodData

- > 8600 foods, ~150 nutrients
- covers a wide range of food, mostly ingredients/whole foods
- uses USDA SR27 was released late 2014, minor update in May 2015; will be updating to the recently released SR28
- > 20,000 foods (and growing) from ItemMaster (packaged foods)
- under development
FoodData entities and properties

Table[EntityValue["Food", "RandomEntity"], {5}]

{ "Ahold Chunky Tuna Salad",
  Soup, chicken broth, ready-to-serve, APPLEBEE'S, 9 oz house sirloin steak,
  Cento Hot Jalapeno Peppers Diced, Vienna Beef Polish Sausage
}

EntityValue["Food", "Properties"] // RandomSample[#, 10] &

{ "chromium content per typical item size",
  phenylalanine content per typical item size, absolute alpha-linolenic acid content,
  alcohol content per typical serving size, sugar alcohol content per typical item size,
  relative parinaric acid content, omega three fatty acid content per typical item size,
  docosahexaenoic acid (DHA) content per typical item size,
  relative aspartic acid content, nickel content per typical item size
}

refriedBeans = Entity["Food", {
  preparation -> refried ($Failed),
  food type -> bean (food type)
}]

Food

- preparation : refried
- food type : bean
FoodData examples

EntityValue queries

EntityValue[refriedBeans, \textit{relative protein content}]
0.057554 g/g

EntityValue[\textit{Food}→\textit{FatContent}→\textit{skim (fat content)}, \textit{food type}→\textit{milk (food type)}], "PropertyAssociation"

\{\textit{relative calorie content} → 0.38 Cal/g, \textit{relative carbohydrate calorie content} → 0.215172 Cal/g\}

dinner = EntityGroup[{
  EntityInstance[\textit{Food}→\textit{ContainsExactly[\textit{spaghetti (food type)}]}], "Servings"→2],
  EntityInstance[\textit{Food}→\textit{ContainsExactly[\textit{meatball (food type)}]}], "Servings"→1],
  EntityInstance[\textit{Food}→\textit{Bertoli (brand name)}→\textit{alfredo (flavor)}→\textit{sauce (food type)}], "Servings"→1.5]
}]

EntityValue[dinner, \{\textit{absolute carbohydrate content}, \textit{absolute calorie content}, \textit{absolute protein content}\}, "PropertyAssociation"

\{\textit{absolute carbohydrate content} → 118.974 g, \textit{absolute calorie content} → 963.7 Cal, \textit{absolute protein content} → 40.602 g\}
FoodData examples

Fruits weighted by volume

Tamarind, mangosteen, boysenberry, cherry, lime, passion fruit, sugar apple, grapefruit, melon, pomegranate, pineapple, apple, passion apple, pear, pitanga, naranjilla, plum orange, gooseberry, grape, currant, persimmon, fig, tangerine, persimmon, cherry, apricot, gooseberry, grape, nectarine, persimmon, melon, pomegranate, pineapple, apple, passion apple, pear, pitanga, naranjilla, plum orange, gooseberry, grape, currant.

Grains weighted by surface area

Millet, tortilla, bread, corn, amaranth grain, spaghetti, spaghett, teff, tapioca pearl, amaranth grain.
**BarCode**

entity1 =

```
Entity["Food", universal product codes] \[BarcodeRecognize[...]]
```

```
foods [universal product codes] 03800310201
```

```
EntityValue[entity1, image, "EntityAssociation"]
```

```
\{Kellogg's Pop-Tarts Blueberry Frosted Toaster Pastries\} →
```

```
EntityValue[First@EntityList@entity1, "NutritionLabel"]
```

<table>
<thead>
<tr>
<th><strong>Nutrition Facts</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>serving size</strong> 52 g</td>
</tr>
<tr>
<td><strong>total calories</strong> 200</td>
</tr>
<tr>
<td><strong>% daily value</strong></td>
</tr>
<tr>
<td><strong>total fat</strong> 5 g</td>
</tr>
<tr>
<td>saturated fat</td>
</tr>
<tr>
<td>trans fat</td>
</tr>
<tr>
<td><strong>cholesterol</strong> 0 mg</td>
</tr>
<tr>
<td><strong>sodium</strong> 170 mg</td>
</tr>
<tr>
<td><strong>total carbohydrates</strong> 38 g</td>
</tr>
<tr>
<td>dietary fiber</td>
</tr>
<tr>
<td>sugar</td>
</tr>
<tr>
<td><strong>protein</strong> 2 g</td>
</tr>
<tr>
<td>vitamin A</td>
</tr>
<tr>
<td>thiamin</td>
</tr>
<tr>
<td>niacin</td>
</tr>
<tr>
<td>iron</td>
</tr>
<tr>
<td>riboflavin</td>
</tr>
<tr>
<td>niacin</td>
</tr>
<tr>
<td>vitamin B12</td>
</tr>
</tbody>
</table>

*percent daily values are based on a 2000 calorie diet*
Unit and formula system: Unit conversion

UnitConvert["Dalton"]
1.660539 \times 10^{-27} \text{ kg}

UnitConvert["DietaryCalories", "SI"]
4.18 \text{ kJ}

UnitConvert[Quantity[.9, "Atmospheres"], "SI"]
91.1925 \text{ kPa}

UnitConvert[Quantity[55, "mph"], "SI"]
\frac{15367}{625} \text{ m/s}

UnitConvert[Quantity[55.0, "mph"], "SI"]
24.5872 \text{ m/s}

UnitConvert[Quantity[55.0, "mph"], "Furlongs"/"Fortnights"]
147840. \text{ furlongs/fortnight}
Unit and formula system: FormulaLookup and FormulaData

FormulaLookup["Arrhenius equation"]
{FirstOrderArrheniusEquation, SecondOrderArrheniusEquation}

FormulaData["FirstOrderArrheniusEquation"]
\[ k = e^{-\frac{E_A}{RT}} \]

InputForm[%]
QuantityVariable["k", "FirstOrderRateConstant"] == E^((Quantity[-1, "MolarGasConstant" QuantityVariable["T", "Temperature"]]) * QuantityVariable["A", "FirstOrderArrheniusFactor"])

FormulaData["FirstOrderArrheniusEquation", 
{"T" -> 298 K, "ActivationEnergy" -> 100 kJ/mol, "A" -> 10.12 / s}]

\[ k = 2.96441 \times 10^{-6} \text{ per second} \]
Unit and formula system: Advanced formula lookup

RequiredPhysicalQuantities

FormulaLookup[All, RequiredPhysicalQuantities -> {"Volume", "Mass"}] // Column
Buoyancy
FreeWaterDeficit
MassDensity
{BoseCondensationTemperature, Volume}
{FermiEnergy, Volume}
{SackurTetrodeEquation, InternalEnergy}
{MeanFreePath, Volume, Diameter}
{MeanFreePath, Volume, Radius}
{SackurTetrodeEquation, InternalEnergy, Temperature}
{SackurTetrodeEquation, ThermalDeBroglieWavelength, Temperature}

RequiredPhysicalQuantities and ExcludedPhysicalQuantities

FormulaLookup["moment of inertia formula"] // Column
ParallelAxisTheorem
ConeMomentOfInertia
CuboidMomentOfInertia
CylinderMomentOfInertia
DiskMomentOfInertia
EllipticalLaminaMomentOfInertia
PointMassMomentOfInertia
SolidEllipsoidMomentOfInertia
SphereMomentOfInertia
ThinRodMomentOfInertia
TriangularPlateMomentOfInertia

FormulaLookup["moment of inertia formula",
  RequiredPhysicalQuantities -> {"Radius"},
  ExcludedPhysicalQuantities -> {"Height"}] // Column
DiskMomentOfInertia
SphereMomentOfInertia
Unit and formula system: In Wolfram|Alpha

WolframAlpha["arrhenius equation"]

Assuming the input is a formula | Use "arrhenius" as a popular curve instead

Assuming first order Arrhenius equation | Use second order Arrhenius equation instead

Calculate [first order rate constant]

• pre-exponential factor: \(1.0 \times 10^{12}\) per second
• activation energy: \(100\) kJ/mol
• temperature: \(298\) K

Input interpretation:

[first order Arrhenius equation]

Equation:

\[
k = A \exp\left(-\frac{E_a}{RT}\right)
\]

\(k\) first order rate constant

\(A\) pre-exponential factor

\(E_a\) activation energy

\(T\) temperature

\(R\) molar gas constant (\(\approx 8.31446\) J/(mol K))

Input values:

<table>
<thead>
<tr>
<th>pre-exponential factor</th>
<th>(1 \times 10^{12}) per second</th>
</tr>
</thead>
<tbody>
<tr>
<td>activation energy</td>
<td>(100) kJ/mol (kilojoules per mole)</td>
</tr>
<tr>
<td>temperature</td>
<td>(298) K (kelvins)</td>
</tr>
</tbody>
</table>

Results:

<table>
<thead>
<tr>
<th>first order rate constant</th>
<th>(2.964 \times 10^{-6}) per second</th>
</tr>
</thead>
<tbody>
<tr>
<td>characteristic time</td>
<td>(93.71) hours</td>
</tr>
</tbody>
</table>
\[ \ln(K) = \frac{-E_a}{R} \frac{1}{T} + \ln(A) \]

(slope = \(-\frac{E_a}{R}\), intercept = \(\ln(A)\))
ChemicalFunctionalGroup summary

- curated collection consisting of 322 chemical functional groups
- to be released in a forthcoming version of Mathematica
**ChemicalFunctionalGroup entities**

EntityValue["ChemicalFunctionalGroup", "Entities"] // Take[#, 50] &

{ chloride, fluoride, ester, thio, nitro, carbodiimide, carboxylate, formate, nitroso, triple bond, oxime, alkoxide, ammonium, aminocarbonyl, azido, aziridinyl, boronate, borono, bromide, bromoethynyl, carbonyl, carboxyl, chloroethyl, cuprate, cyanato, cyano, diazenyl, dithioester, double bond, fluoroethyl, formyl, fourfold conjugated double bond, furanyl, hydrogen borono, hydroperoxy, hydrosulfiny1, hydroxy, hydroxyimino, imino, iodide, iodoethyl, isocyanato, isothiazolyl, isothiocyanato, isoxazolyl, ketenyl, mercapto, methyl, methylene, organo lithium }

Length[%]

322
ChemicalFunctionalGroup properties and classes

EntityValue["ChemicalFunctionalGroup", "Properties"]

\{ brutto formula, color structure diagram, number of generic locants, 
exclusively carbon and hydrogen containing, number of membered atoms, name, 
contains nitrogen, contains oxygen, contains phosphorus, tautomerism causative, 
representative, representative color structure diagram, InChi of representative, 
InChIKey of representative, representative structure graph, contains silicon, 
chemical structure graph, substance class name, contains sulfur, valence of group fragment \}

EntityClassList["ChemicalFunctionalGroup"]

\{ chemical functional groups, exclusively carbon and hydrogen containing functional groups, 
potential stereogenic functional groups, 
potential tautomerism causing functional groups, oxygen containing functional groups, 
nitrogen containing functional groups, sulfur containing functional groups, 
phosphorus containing functional groups, silicon containing functional groups, 
electron donating functional groups, electron withdrawing functional groups, 
radical stabilizing functional groups, metal organic functional groups \}
ChemicalFunctionalGroup examples

\[
data = \text{Entity["ChemicalFunctionalGroup", "Ester"]}["PropertyAssociation"];
\]
\[
\text{Grid[List@@@ Normal[data], Dividers \rightarrow All, Alignment \rightarrow Left]}
\]
<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>brutto formula</td>
<td>C02R2</td>
</tr>
<tr>
<td>color structure diagram</td>
<td></td>
</tr>
<tr>
<td>number of generic locants</td>
<td>2</td>
</tr>
<tr>
<td>exclusively carbon and hydrogen containing</td>
<td>False</td>
</tr>
<tr>
<td>number of membered atoms</td>
<td>3</td>
</tr>
<tr>
<td>name</td>
<td>ester</td>
</tr>
<tr>
<td>contains nitrogen</td>
<td>False</td>
</tr>
<tr>
<td>contains oxygen</td>
<td>True</td>
</tr>
<tr>
<td>contains phosphorus</td>
<td>False</td>
</tr>
<tr>
<td>tautomerism causative</td>
<td>False</td>
</tr>
<tr>
<td>representative</td>
<td>methyl acetate</td>
</tr>
<tr>
<td>representative color structure diagram</td>
<td></td>
</tr>
<tr>
<td>InChI of representative</td>
<td>InChI=1S/C3H6O2/c1-3(4)5-2/h1-2H3</td>
</tr>
<tr>
<td>InChIKey of representative</td>
<td>KXXVLQRXCPHEJC–UHFFFAOYSA–N</td>
</tr>
<tr>
<td>representative structure graph</td>
<td></td>
</tr>
<tr>
<td>contains silicon</td>
<td>False</td>
</tr>
<tr>
<td>chemical structure graph</td>
<td></td>
</tr>
<tr>
<td>substance class name</td>
<td>carbonacid esters</td>
</tr>
<tr>
<td>contains sulfur</td>
<td>False</td>
</tr>
<tr>
<td>valence of group fragment</td>
<td>bivalent</td>
</tr>
</tbody>
</table>
ChemicalFunctionalGroup examples

Show hydrocarbon functional groups:

EntityList[exclusively carbon and hydrogen containing functional groups (Failed)]

{triple bond, double bond, fourfold conjugated double bond, methyl,
methylene, phenyl, substituted carbon, threefold conjugated double bond,
twofold conjugated double bond, methantetrayl, methyldyne, propadienyl}

Queries based on property values via implicitly defined entities:

Entity["ChemicalFunctionalGroup",
   {"MemberedAtomCount" -> TakeLargest[5]]]"Entities"

{(1-carboxy)octa-1,3,5,7-tetraenyl, (oxycarboxy)octa-2,4,6,8-tetraenyl,
(1-formyl)octa-1,3,5,7-tetraenyl, 1-oxonona-1,3,5,7-tetraenyl, (1-carboxy)hexa-1,3,5-triennyl}

EntityValue[%, "MemberedAtomCount"]
{12, 11, 11, 10, 10}

Entity["ChemicalFunctionalGroup",
   {"MemberedAtomCount" -> TakeSmallest[5]]]"Entities"

{chloride, fluoride, thio, alkoxide, bromide}

EntityValue[%, "MemberedAtomCount"]
{1, 1, 1, 1, 1}

Entity["ChemicalFunctionalGroup",
   {"MemberedAtomCount" -> Interval[[1, 2]]]"Entities"

{alkoxide, azo, bromide, carbonyl, chloride, cuprate, cyano,
dioxy, dithio, double bond, fluoride, hydroxy, iodide, mercapto,
methantetrayl, methyldyne, N-carbonimidoyl, nitroso, organo lithium,
organo magnesium halide, organo zinc, oxy, quatary oxysilanyl,
quaternary ammonium, secondary amino, secondary iminium, secondary phosphino,
substituted carbon, sulfinyl, tertiary amino, tertiary ammonium, tertiary iminium,
tertiary phosphino, tetralkylsilyl, tetralkylstannyl, thio, thioxo, triple bond}
ChemicalFunctionalGroup examples

Queries based on class membership via implicitly defined entities:

Entity["ChemicalFunctionalGroup",
    {EntityProperty["ChemicalFunctionalGroup", "SiliconContaining"] -> True][
    "Entities"
]
    { quartary oxysilyl, tertiary oxysilyl, tetraalkylsilyl }

EntityValue[%, representative color structure diagram]

{ \begin{array}{c}
   \text{O} & \text{Si} \\
   & \text{Si} \end{array} } , \begin{array}{c}
   \text{Si} & \text{O} \\
   \text{Si} & \text{H} \end{array} , \begin{array}{c}
   \text{Si} \end{array} 
\end{array}
"PubChem" service

- provides information on the biological activities of small molecules
- provides a fast chemical structure similarity search tool
- API accessible via ServiceConnect in the Wolfram Language
"PubChem" examples

Find compound by identifier:

ServiceExecute["PubChem", "CompoundDescription", 
  {"InChI" -> "InChI=1S/ClO2/c2-1-3"}]

<table>
<thead>
<tr>
<th>CompoundID</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>24870</td>
<td>CHLORINE DIOXIDE</td>
</tr>
</tbody>
</table>

Find compounds by name:

ServiceExecute["PubChem", "CompoundDescription", {"Name" -> "lipitor"}]

<table>
<thead>
<tr>
<th>CompoundID</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>60822</td>
<td>ATORVASTATIN CALCIUM</td>
</tr>
<tr>
<td>60823</td>
<td>atorvastatin</td>
</tr>
<tr>
<td>656846</td>
<td>Torvast</td>
</tr>
<tr>
<td>15378998</td>
<td>ATORVASTATIN CALCIUM</td>
</tr>
</tbody>
</table>

2 levels | 1 rows
2 levels | 4 rows
"PubChem" examples

Search compounds by substructure:

```
ServiceExecute["PubChem", "CompoundDescription", {
  "InChI" -> "InChI=1S/C12H4Cl4O2/c13-5-1-9-10(2-6(5)14)18-12-4-8(16)7(15)3-11(12)17-9/h1-4H",
  Method -> "SubstructureSearch"}
]
```

<table>
<thead>
<tr>
<th>CompoundID</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>15625</td>
<td>Dioxin</td>
</tr>
<tr>
<td>156712</td>
<td>C-132,3,7,8–TCDD</td>
</tr>
<tr>
<td>158818</td>
<td>AC1L4KNU</td>
</tr>
<tr>
<td>21684359</td>
<td>SCHEMEL10692049</td>
</tr>
<tr>
<td>24849081</td>
<td>LS=193908</td>
</tr>
<tr>
<td>66779894</td>
<td>SCHEMEL723868</td>
</tr>
<tr>
<td>67147272</td>
<td>SCHEMEL1747904</td>
</tr>
<tr>
<td>67412859</td>
<td>SCHEMEL2449396</td>
</tr>
<tr>
<td>67418390</td>
<td>SCHEMEL2464843</td>
</tr>
<tr>
<td>67700425</td>
<td>SCHEMEL8815834</td>
</tr>
<tr>
<td>67893929</td>
<td>SCHEMEL9618549</td>
</tr>
<tr>
<td>69696353</td>
<td>SCHEMEL6013737</td>
</tr>
<tr>
<td>69780693</td>
<td>SCHEMEL6436813</td>
</tr>
<tr>
<td>69865253</td>
<td>SCHEMEL6658823</td>
</tr>
<tr>
<td>69940835</td>
<td>SCHEMEL6930992</td>
</tr>
<tr>
<td>70002866</td>
<td>SCHEMEL7159644</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

2 levels | 26 rows
Search by superstructure:

```plaintext
ServiceExecute["PubChem", "CompoundDescription",
{"InChI" -> "mesoxalic acid (chemical)"},
Method -> "SuperstructureSearch"]
```

<table>
<thead>
<tr>
<th>CompoundID</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>712</td>
<td>formaldehyde</td>
</tr>
<tr>
<td>284</td>
<td>formic acid</td>
</tr>
<tr>
<td>962</td>
<td>water</td>
</tr>
<tr>
<td>283</td>
<td>formate</td>
</tr>
<tr>
<td>5462310</td>
<td>Carbon</td>
</tr>
<tr>
<td>760</td>
<td>glyoxylic acid</td>
</tr>
<tr>
<td>7880</td>
<td>GLYOXAL</td>
</tr>
<tr>
<td>10132</td>
<td>Ketomalonic acid</td>
</tr>
<tr>
<td>24602</td>
<td>DEUTERIUM OXIDE</td>
</tr>
<tr>
<td>159832</td>
<td>Monoxygen</td>
</tr>
<tr>
<td>783</td>
<td>Hydrogen</td>
</tr>
<tr>
<td>961</td>
<td>hydroxide</td>
</tr>
<tr>
<td>1038</td>
<td>hydron</td>
</tr>
<tr>
<td>9084</td>
<td>Propanedioate</td>
</tr>
<tr>
<td>24523</td>
<td>DEUTERIUM</td>
</tr>
<tr>
<td>24824</td>
<td>TRITIUM</td>
</tr>
<tr>
<td>...</td>
<td></td>
</tr>
</tbody>
</table>

2 levels | 158 rows
"PubChem" examples

Get detailed properties of compounds:

```
ServiceExecute["PubChem", "CompoundProperties",
{"Name" -> {"aspirin", "acetaminophen", "ibuprofen"}}]
```

<table>
<thead>
<tr>
<th>CompoundID</th>
<th>MolecularFormula</th>
<th>MolecularWeight</th>
<th>CanonicalSMILES</th>
<th>IsomericSMILES</th>
</tr>
</thead>
<tbody>
<tr>
<td>2244</td>
<td>C9H8O4</td>
<td>180.2</td>
<td>CO(=O)OC1=CC=CC=CC1(=O)O</td>
<td>CO(=O)OC1=CC=CC=CC1(=O)O</td>
</tr>
<tr>
<td>1983</td>
<td>C8H9NO2</td>
<td>151.2</td>
<td>CO(=O)NC1=CC=CC(=O)=C(=O)O</td>
<td>CO(=O)NC1=CC=CC(=O)=C(=O)O</td>
</tr>
<tr>
<td>3672</td>
<td>C13H18O2</td>
<td>206.3</td>
<td>CO(C)OC1=CC=CC1=CC1(=O)(=O)O</td>
<td>CO(C)OC1=CC=CC1=CC1(=O)(=O)O</td>
</tr>
</tbody>
</table>

2 levels  | 3 rows

```
res = ServiceExecute["PubChem", "CompoundSDF", {"Name" -> "aspirin"}]
```

```
res["Graphics3D", 1]
```
"ChemSpider" service

- free chemical structure database providing access to more than 43 million structures, properties, and associated information
- integrates compounds from ~500 data sources
- owned by the Royal Society of Chemistry
- API accessible via ServiceConnect in the Wolfram Language
- requires sign up
"ChemSpider" examples

Search for a structure by name or keyword:

Security Token: 24602a49-40f9-483b-b6fe-fc6d0e5a0fb7

ServiceExecute["ChemSpider", "Search", "Query" → "phenol"]

<table>
<thead>
<tr>
<th>ID</th>
<th>971</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 level</td>
<td>1 elements</td>
</tr>
</tbody>
</table>

Get basic or extended information about compounds:

ServiceExecute["ChemSpider", "CompoundInformation", "ID" → 971]

<table>
<thead>
<tr>
<th>CSD</th>
<th>971</th>
</tr>
</thead>
<tbody>
<tr>
<td>InChI</td>
<td>InChI=1S/C8H8O/c7-8-4-2-1-3-5-6/h1-5,7H</td>
</tr>
<tr>
<td>InChIKey</td>
<td>ISWSIDIOBOJEZ-UHFFFAOYSA-N</td>
</tr>
<tr>
<td>SMILES</td>
<td>c1ccc(c(c1)O)</td>
</tr>
<tr>
<td>1 level</td>
<td>4 elements</td>
</tr>
</tbody>
</table>

ServiceExecute["ChemSpider", "ExtendedCompoundInformation", "ID" → "27472491"]

<table>
<thead>
<tr>
<th>CSD</th>
<th>27472491</th>
</tr>
</thead>
<tbody>
<tr>
<td>MF</td>
<td>C_{(16)H_{(19)}N_{(3)}}O_{(5)}S</td>
</tr>
<tr>
<td>SMILES</td>
<td>CO([C@@H][C@@@H][S(1)][C@@][C@@O][C2=O]NC(O)[C@H]c3ccc(cc3)O)N(C(=O)O)O</td>
</tr>
<tr>
<td>InChI</td>
<td>InChI=1/C16H19N3O5S1/c1-16(21)11(15(23)24)19-13(22)10(14(19)25-16)18-12(21)2(17)7</td>
</tr>
<tr>
<td>InChIKey</td>
<td>LSQZJLSUYDOPKJ-RMLAFQJ</td>
</tr>
<tr>
<td>AverageMass</td>
<td>365.4042</td>
</tr>
<tr>
<td>MolecularWeight</td>
<td>365.4042</td>
</tr>
<tr>
<td>MonocotopicMass</td>
<td>365.104553</td>
</tr>
<tr>
<td>NominalMass</td>
<td>365</td>
</tr>
<tr>
<td>ALOGP</td>
<td>0</td>
</tr>
<tr>
<td>XLOGP</td>
<td>0</td>
</tr>
<tr>
<td>CommonName</td>
<td>(2S,5S,6S)-6-(((2S)-2-Amino-2-(4-hydroxyphenyl)acetyl)amino)-3,3-dimethyl-7-oxo-4-thiazolecarboxylic acid</td>
</tr>
<tr>
<td>1 level</td>
<td>12 elements</td>
</tr>
</tbody>
</table>
Get a structure diagram for a given compound ID:

ServiceExecute["ChemSpider", "CompoundThumbnail", "ID" → 3970]

\[
\begin{align*}
\text{H}_2\text{N}&-\text{O} & \text{CH}_3
\end{align*}
\]
"ChemSpider" examples

For a given InChI code, get the corresponding MOL file:

```
mol = ServiceExecute("ChemSpider", "GetIdentifier", {
  "Identifier" → "MOL",
  "InChI" → "InChI=1S/C6H3BrClF/c7-4-1-5(8)3-6(9)2-4/h1-3H"
})
```

```
03121618462D

9 9 0 0 0 0 0 0 0 0 0 1 V2000
1.9950 -1.1518 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0.0000 -2.3037 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
1.9950 -3.4555 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0.6635 -1.1518 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
2.6584 -2.3037 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0.6635 -3.4555 0.0000 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0.0000 0.0000 0.0000 Br 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
3.9899 -2.3037 0.0000 Cl 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0.0000 -4.6073 0.0000 F 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
1 4 2 0 0 0 0
1 5 1 0 0 0 0
2 4 1 0 0 0 0
2 6 2 0 0 0 0
3 5 2 0 0 0 0
3 6 1 0 0 0 0
4 7 1 0 0 0 0
5 8 1 0 0 0 0
6 9 1 0 0 0 0
M END
```

1 level | 1 elements
Use the retrieved MOL string for further computations:

```
ImportString[mol["MOL"], "MOL"]
```

Count all heavy atoms:

```
Tally[ImportString[mol["MOL"], {"MOL", "VertexTypes"}]]
```

```
{{C, 6}, {Br, 1}, {Cl, 1}, {F, 1}}
```
"ChemSpider" examples: Spectra

Retrieve data on all available spectra:

```plaintext
allSpecInfo = ServiceExecute["ChemSpider", "AllSpectraInformation"]
```

<table>
<thead>
<tr>
<th>SpectrumID</th>
<th>ID</th>
<th>SpectrumType</th>
<th>FileName</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>36</td>
<td>235</td>
<td>HNMR</td>
<td>Benzaldehyde=HNMR.jdx</td>
<td>Benzaldehyde HNMR Solv</td>
</tr>
<tr>
<td>37</td>
<td>235</td>
<td>CNMR</td>
<td>BenzaldehydeCNMR.jdx</td>
<td>Benzaldehyde CNMR Solv</td>
</tr>
<tr>
<td>65</td>
<td>172</td>
<td>HNMR</td>
<td>ethanal_h.jdx</td>
<td>This data was acquired on t</td>
</tr>
<tr>
<td>66</td>
<td>172</td>
<td>CNMR</td>
<td>ethanal_c.jdx</td>
<td>This data was acquired on t</td>
</tr>
<tr>
<td>87</td>
<td>8438</td>
<td>HNMR</td>
<td>3,4-dihydroxybenzaldehyde=HNMR.jdx</td>
<td>{}</td>
</tr>
<tr>
<td>88</td>
<td>2006805</td>
<td>HNMR</td>
<td>5-methylfurfurylamine=HNMR.jdx</td>
<td>Dr. Jean-claude Bradley's</td>
</tr>
<tr>
<td>90</td>
<td>13859497</td>
<td>HNMR</td>
<td>piperonal.jdx</td>
<td>Dr. Jean-Claude Bradley's</td>
</tr>
<tr>
<td>91</td>
<td>6145</td>
<td>HNMR</td>
<td>piperonal.jdx</td>
<td>Dr. Jean-Claude Bradley's</td>
</tr>
<tr>
<td>95</td>
<td>5566</td>
<td>HNMR</td>
<td>3,4-dihydroxybenzaldehyde=HNMR.jdx</td>
<td>{}</td>
</tr>
<tr>
<td>96</td>
<td>7525</td>
<td>HNMR</td>
<td>resepine_1.Href.jdx</td>
<td>This data was provided by</td>
</tr>
<tr>
<td>98</td>
<td>7387</td>
<td>R</td>
<td>etanline.jdx</td>
<td>public domain</td>
</tr>
<tr>
<td>100</td>
<td>6794</td>
<td>R</td>
<td>detailn.jdx</td>
<td>public domain</td>
</tr>
<tr>
<td>101</td>
<td>7641</td>
<td>R</td>
<td>mylene.jdx</td>
<td>public domain</td>
</tr>
<tr>
<td>102</td>
<td>13869451</td>
<td>R</td>
<td>mclanI.jdx</td>
<td>public domain</td>
</tr>
<tr>
<td>103</td>
<td>236</td>
<td>UV-Vis</td>
<td>benzepi.jdx</td>
<td>public domain</td>
</tr>
<tr>
<td>104</td>
<td>59008</td>
<td>UV-Vis</td>
<td>thymoblue.jdx</td>
<td>public domain</td>
</tr>
</tbody>
</table>

2 levels | 8635 rows
```plaintext
allSpecInfo[All, "SpectrumType"] // Tally

<table>
<thead>
<tr>
<th>SpectrumType</th>
<th>Tally</th>
</tr>
</thead>
<tbody>
<tr>
<td>HNMR</td>
<td>1503</td>
</tr>
<tr>
<td>CNMR</td>
<td>1116</td>
</tr>
<tr>
<td>IR</td>
<td>5407</td>
</tr>
<tr>
<td>UV-Vis</td>
<td>144</td>
</tr>
<tr>
<td>NR</td>
<td>30</td>
</tr>
<tr>
<td>EI</td>
<td>104</td>
</tr>
<tr>
<td>2D1H1H-COSY</td>
<td>12</td>
</tr>
<tr>
<td>2D1H13CD</td>
<td>37</td>
</tr>
<tr>
<td>APCI+</td>
<td>2</td>
</tr>
<tr>
<td>R</td>
<td>47</td>
</tr>
<tr>
<td>MALDI+</td>
<td>12</td>
</tr>
<tr>
<td>2D1H13CLR</td>
<td>5</td>
</tr>
<tr>
<td>APPI−</td>
<td>1</td>
</tr>
<tr>
<td>O+ve</td>
<td>16</td>
</tr>
<tr>
<td>ESI+</td>
<td>159</td>
</tr>
<tr>
<td>2D1H1H-ESY</td>
<td>2</td>
</tr>
</tbody>
</table>

2 levels | 20 rows
```
"ChemSpider" examples: Spectra

Retrieve spectrum 6127:

specData = Select[allSpecInfo, #ID == "6127" &]

<table>
<thead>
<tr>
<th>SpectrumID</th>
<th>ID</th>
<th>SpectrumType</th>
<th>FileName</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>6726</td>
<td>6127</td>
<td>CNMR</td>
<td>c13nir.jdx</td>
<td>Thanks to contributions from the following partners:</td>
</tr>
<tr>
<td>6727</td>
<td>6127</td>
<td>IR</td>
<td>ir.jdx</td>
<td>Thanks to contributions from the following partners:</td>
</tr>
<tr>
<td>6728</td>
<td>6127</td>
<td>EI</td>
<td>mass.jdx</td>
<td>Thanks to contributions from the following partners:</td>
</tr>
<tr>
<td>6729</td>
<td>6127</td>
<td>HNMR</td>
<td>nmr.jdx</td>
<td>Thanks to contributions from the following partners:</td>
</tr>
<tr>
<td>6730</td>
<td>6127</td>
<td>UV-Vis</td>
<td>uv.jdx</td>
<td>Thanks to contributions from the following partners:</td>
</tr>
<tr>
<td>8583</td>
<td>6127</td>
<td>HNMR</td>
<td>nmr_assigned.jdx</td>
<td>Thanks to contributions from the following partners:</td>
</tr>
<tr>
<td>8584</td>
<td>6127</td>
<td>CNMR</td>
<td>c13nir_assigned.jdx</td>
<td>Thanks to contributions from the following partners:</td>
</tr>
</tbody>
</table>

2 levels | 7 rows

Plot the infrared spectrum:

thumbnailImage = ServiceExecute["ChemSpider", "CompoundThumbnail", "ID" -> 6127]

\[
\begin{align*}
\text{H}_3\text{C} & \quad \text{O} \\
\text{Cl} & \\
\end{align*}
\]

rawIR = Import["http://www.chemspider.com/FilesHandler.ashx?type=blob&id=" <> specData[2, 1], "JCAMP-DX"];

acs2016-small.nb | 67
lp = ListPlot[MovingAverage[rawIR[[1]], 15],
Joined -> True, PlotRange -> {{4200, 500}, All},
ScalingFunctions -> {"Reverse", Identity}, Frame -> True,
GridLines -> Automatic, Filling -> Top, ColorFunction -> "BlueGreenYellow",
Epilog -> Inset[Framed[thumbnailImage], {-3000, Center}]]
"ChemSpider" examples: Spectra

Find peaks in the 13C-NMR spectrum using Wolfram Language commands

rawCNMR = Import["http://www.chemspider.com/FilesHandler.ashx?type=blob&id=" <> specData[1, 1], "JCAMP-DX"];

(* ### OBSERVE FREQUENCY=75 *)
tsCNMR = TimeSeries[#1[[1]] / 75, #1[[2]]] & /@ rawCNMR[[1]]

Data points: 16384]

peaksCNMR = FindPeaks[TimeSeriesResample[t CNMR], 0, 0, 10^7]

TimeSeries[Time: 20.6 to 179.
Data points: 8]

peakList = Grid[MapIndexed[{"Peak " ~~ ToString[#2[[1]]], #1[[1]]} &, Normal[peaksCNMR]], Frame -> All, Alignment -> "];

ACS2016-small.nb 69
(* The pure compound should exhibit only peaks #3 and #7: byproducts or other impurities are present here. Peaks #4 to #6 is CDCl3 (solvent). *)

ListPlot[tsCNMR, PlotRange -> All, ScalingFunctions -> {"Reverse", Identity}, Frame -> True, GridLines -> Automatic, FrameTicks -> {Automatic, None}, Joined -> True, Epilog -> {{Red, PointSize[0.01], Point[{-1, 1} # & /@ peaksCNMR["Path"]]}, Inset[Framed[thumbnailImage], {-100, Center}], Inset[peakList, {-200, Center}]}, ImageSize -> Large]
"OpenPHACTS" service

- brings together data from multiple publicly available sources of pharmacological and physicochemical information
- integrated into a consolidated database
- API accessible via ServiceConnect in the Wolfram Language
- requires sign up
"OpenPHACTS" examples

Application Key: 4af923782e7110bd5da003c68995ee6
Application ID: 0024cc5c

Find the compound URI that corresponds to a given identifier:

ServiceExecute["OpenPHACTS", "GetURI",
"SMILES" -> "CCCC(C(N=C(N1)C2=C(OCC)C=CC(S(N3CCN(CC3)C(=O)=O)=C2)=C4C1=O)=NN4C"
]

Get detailed information about a compound:

sorafenib = ServiceExecute["OpenPHACTS", "CompoundInformation", "URI" ->
 "http://www.conceptwiki.org/concept/38932552-111f-4a4e-a46a-4ed1d7bdf9d5" ];

Keys[sorafenib] // Normal

{{URI, InDataset, ReportedAdverseEvent, PrefLabel},
 {URI, InDataset, ReportedAdverseEvent, PrefLabel},
 {URI, DescriptionEn, Description, DrugTypeEn, DrugType,
  GenericNameEn, GenericName, MetabolismEn, Metabolism, ProteinBindingEn,
  ProteinBinding, ToxicityEn, Toxicity, InDataset, DrugInteraction},
 {URI, MwFreebase, InDataset, Type}, {URI, InDataset, Hba, Hbd, Inchi,
  Inchikey, Logp, Molformula, Molweight, Psa, Ro5Violations, Rtb, Smiles}}
### sorafenib[[3]]

<table>
<thead>
<tr>
<th>Field</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>DescriptionEn</td>
<td>Sorafenib (INN), marketed as Nexavar by Bayer, is a drug approved for the treatment of advanced renal cell carcinoma.</td>
</tr>
<tr>
<td>Description</td>
<td>Sorafenib (INN), marketed as Nexavar by Bayer, is a drug approved for the treatment of advanced renal cell carcinoma.</td>
</tr>
<tr>
<td>DrugTypeEn</td>
<td>{Investigational, approved}</td>
</tr>
<tr>
<td>DrugType</td>
<td>{Investigational, approved}</td>
</tr>
<tr>
<td>GenericNameEn</td>
<td>Sorafenib</td>
</tr>
<tr>
<td>GenericName</td>
<td>Sorafenib</td>
</tr>
<tr>
<td>MetabolismEn</td>
<td>Sorafenib is metabolized primarily in the liver, undergoing oxidative metabolism, mediated by CYP3A4, i</td>
</tr>
<tr>
<td>Metabolism</td>
<td>Sorafenib is metabolized primarily in the liver, undergoing oxidative metabolism, mediated by CYP3A4, i</td>
</tr>
<tr>
<td>ProteinBindingEn</td>
<td>99.5% bound to plasma proteins.</td>
</tr>
<tr>
<td>ProteinBinding</td>
<td>99.5% bound to plasma proteins.</td>
</tr>
<tr>
<td>ToxicityEn</td>
<td>The highest dose of sorafenib studied clinically is 800 mg twice daily. The adverse reactions observed</td>
</tr>
<tr>
<td>Toxicity</td>
<td>The highest dose of sorafenib studied clinically is 800 mg twice daily. The adverse reactions observed</td>
</tr>
<tr>
<td>InDataset</td>
<td><a href="http://www.openphacts.org/bio2rdf/drugbank">http://www.openphacts.org/bio2rdf/drugbank</a></td>
</tr>
<tr>
<td>DrugInteraction</td>
<td><img src="http://bio2rdf.org/drugbank_resource:DB00398:DB00755" alt="Diagram" /> TextEn → CDI between Sorafenib and other drugs]</td>
</tr>
</tbody>
</table>

1 level | 15 elements
"OpenPHACTS" examples

Get a list of compounds in a given class:

ServiceExecute["OpenPHACTS", "CompoundClassMembersList", 
{"URI" → "http://purl.obolibrary.org/obo/CHEBI_24431", MaxItems → 5}]

<table>
<thead>
<tr>
<th>URI</th>
<th>ExactMatch</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://purl.obolibrary.org/obo/CHEBI_10">http://purl.obolibrary.org/obo/CHEBI_10</a></td>
<td>KeyAbsent</td>
</tr>
<tr>
<td><a href="http://purl.obolibrary.org/obo/CHEBI_100">http://purl.obolibrary.org/obo/CHEBI_100</a></td>
<td><a href="http://ps.rsc.org/OPS1634725">http://ps.rsc.org/OPS1634725</a></td>
</tr>
<tr>
<td><a href="http://purl.obolibrary.org/obo/CHEBI_10000">http://purl.obolibrary.org/obo/CHEBI_10000</a></td>
<td>KeyAbsent</td>
</tr>
<tr>
<td><a href="http://purl.obolibrary.org/obo/CHEBI_10001">http://purl.obolibrary.org/obo/CHEBI_10001</a></td>
<td>KeyAbsent</td>
</tr>
<tr>
<td><a href="http://purl.obolibrary.org/obo/CHEBI_10002">http://purl.obolibrary.org/obo/CHEBI_10002</a></td>
<td>KeyAbsent</td>
</tr>
</tbody>
</table>

2 levels | 5 rows

Get a list of ChEMBL target types and counts for each type:

ServiceExecute["OpenPHACTS", "TargetTypes"]

<table>
<thead>
<tr>
<th>URI</th>
<th>TargetCount</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://df.ebi.ac.uk/terms/chebi#protein_family">http://df.ebi.ac.uk/terms/chebi#protein_family</a></td>
<td>217</td>
<td>protein_family</td>
</tr>
<tr>
<td><a href="http://df.ebi.ac.uk/terms/chebi#UnclassifiedTarget">http://df.ebi.ac.uk/terms/chebi#UnclassifiedTarget</a></td>
<td>2</td>
<td>unclassified</td>
</tr>
<tr>
<td><a href="http://df.ebi.ac.uk/terms/chebi#ProteinNucleicAcidComplex">http://df.ebi.ac.uk/terms/chebi#ProteinNucleicAcidComplex</a></td>
<td>6</td>
<td>KeyAbsent</td>
</tr>
<tr>
<td><a href="http://df.ebi.ac.uk/terms/chebi#ProteinComplex">http://df.ebi.ac.uk/terms/chebi#ProteinComplex</a></td>
<td>261</td>
<td>protein_complex</td>
</tr>
<tr>
<td><a href="http://df.ebi.ac.uk/terms/chebi#UnknownTarget">http://df.ebi.ac.uk/terms/chebi#UnknownTarget</a></td>
<td>18</td>
<td>unknown</td>
</tr>
<tr>
<td><a href="http://df.ebi.ac.uk/terms/chebi#ProteinComplexGroup">http://df.ebi.ac.uk/terms/chebi#ProteinComplexGroup</a></td>
<td>44</td>
<td>protein_complex_group</td>
</tr>
<tr>
<td><a href="http://df.ebi.ac.uk/terms/chebi#SubCellular">http://df.ebi.ac.uk/terms/chebi#SubCellular</a></td>
<td>9</td>
<td>sub_cellular</td>
</tr>
<tr>
<td><a href="http://df.ebi.ac.uk/terms/chebi#Tissue">http://df.ebi.ac.uk/terms/chebi#Tissue</a></td>
<td>242</td>
<td>tissue</td>
</tr>
<tr>
<td><a href="http://df.ebi.ac.uk/terms/chebi#SingleProtein">http://df.ebi.ac.uk/terms/chebi#SingleProtein</a></td>
<td>6018</td>
<td>single_protein</td>
</tr>
<tr>
<td><a href="http://df.ebi.ac.uk/terms/chebi#SmallMoleculeTarget">http://df.ebi.ac.uk/terms/chebi#SmallMoleculeTarget</a></td>
<td>18</td>
<td>KeyAbsent</td>
</tr>
<tr>
<td><a href="http://df.ebi.ac.uk/terms/chebi#ProteinProteinInteraction">http://df.ebi.ac.uk/terms/chebi#ProteinProteinInteraction</a></td>
<td>21</td>
<td>ppi</td>
</tr>
<tr>
<td><a href="http://df.ebi.ac.uk/terms/chebi#ProteinSelectivityGroup">http://df.ebi.ac.uk/terms/chebi#ProteinSelectivityGroup</a></td>
<td>97</td>
<td>protein_selectivity_group</td>
</tr>
<tr>
<td><a href="http://df.ebi.ac.uk/terms/chebi#Macromolecule">http://df.ebi.ac.uk/terms/chebi#Macromolecule</a></td>
<td>5</td>
<td>KeyAbsent</td>
</tr>
<tr>
<td><a href="http://df.ebi.ac.uk/terms/chebi#Metal">http://df.ebi.ac.uk/terms/chebi#Metal</a></td>
<td>8</td>
<td>KeyAbsent</td>
</tr>
<tr>
<td><a href="http://df.ebi.ac.uk/terms/chebi#CellLineTarget">http://df.ebi.ac.uk/terms/chebi#CellLineTarget</a></td>
<td>1630</td>
<td>KeyAbsent</td>
</tr>
<tr>
<td><a href="http://df.ebi.ac.uk/terms/chebi#ChimericProtein">http://df.ebi.ac.uk/terms/chebi#ChimericProtein</a></td>
<td>4</td>
<td>chimeric_protein</td>
</tr>
</tbody>
</table>

... 5

2 levels | 21 rows
Combine external data with Wolfram Language data and computation

Use OpenPHACTS to get a list of pathways stored in the linked data cache (LDC):

```
Table = ServiceExecute["OpenPHACTS", "PathwayOrganisms"]
```

<table>
<thead>
<tr>
<th>URL</th>
<th>PathwayCount</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://purl.obolibrary.org/obo/NCBITaxon_9608">http://purl.obolibrary.org/obo/NCBITaxon_9608</a></td>
<td>664</td>
<td>Homo sapiens</td>
</tr>
<tr>
<td><a href="http://purl.obolibrary.org/obo/NCBITaxon_10090">http://purl.obolibrary.org/obo/NCBITaxon_10090</a></td>
<td>159</td>
<td>Mus musculus</td>
</tr>
<tr>
<td><a href="http://purl.obolibrary.org/obo/NCBITaxon_10116">http://purl.obolibrary.org/obo/NCBITaxon_10116</a></td>
<td>142</td>
<td>Rattus norvegicus</td>
</tr>
<tr>
<td><a href="http://purl.obolibrary.org/obo/NCBITaxon_4932">http://purl.obolibrary.org/obo/NCBITaxon_4932</a></td>
<td>120</td>
<td>Saccharomyces cerevisiae</td>
</tr>
<tr>
<td><a href="http://purl.obolibrary.org/obo/NCBITaxon_9913">http://purl.obolibrary.org/obo/NCBITaxon_9913</a></td>
<td>115</td>
<td>Bos taurus</td>
</tr>
<tr>
<td><a href="http://purl.obolibrary.org/obo/NCBITaxon_7955">http://purl.obolibrary.org/obo/NCBITaxon_7955</a></td>
<td>93</td>
<td>Danio rerio</td>
</tr>
<tr>
<td><a href="http://purl.obolibrary.org/obo/NCBITaxon_6239">http://purl.obolibrary.org/obo/NCBITaxon_6239</a></td>
<td>57</td>
<td>Caenorhabditis elegans</td>
</tr>
<tr>
<td><a href="http://purl.obolibrary.org/obo/NCBITaxon_9598">http://purl.obolibrary.org/obo/NCBITaxon_9598</a></td>
<td>40</td>
<td>Pan troglodytes</td>
</tr>
<tr>
<td><a href="http://purl.obolibrary.org/obo/NCBITaxon_9615">http://purl.obolibrary.org/obo/NCBITaxon_9615</a></td>
<td>38</td>
<td>Canis familiaris</td>
</tr>
<tr>
<td><a href="http://purl.obolibrary.org/obo/NCBITaxon_9031">http://purl.obolibrary.org/obo/NCBITaxon_9031</a></td>
<td>37</td>
<td>Gallus gallus</td>
</tr>
</tbody>
</table>

Use SemanticInterpretation to interpret right column:

```
Table[[All,"Label"]]} // Normal
```

{Homo sapiens, Mus musculus, Rattus norvegicus, Saccharomyces cerevisiae, Bos taurus, Danio rerio, Caenorhabditis elegans, Pan troglodytes, Canis familiaris, Gallus gallus}
<table>
<thead>
<tr>
<th>Animal Species</th>
<th>Common Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homo sapiens</td>
<td>human</td>
</tr>
<tr>
<td>Mus musculus</td>
<td>house mouse</td>
</tr>
<tr>
<td>Rattus norvegicus</td>
<td>Norway rat</td>
</tr>
<tr>
<td>Saccharomyces cerevisiae</td>
<td>Saccharomyces cerevisiae</td>
</tr>
<tr>
<td>Bos taurus</td>
<td>cow</td>
</tr>
<tr>
<td>Danio rerio</td>
<td>striped danio</td>
</tr>
<tr>
<td>Caenorhabditis elegans</td>
<td>Caenorhabditis elegans</td>
</tr>
<tr>
<td>Pan troglodytes</td>
<td>common chimpanzee</td>
</tr>
<tr>
<td>Canis familiaris</td>
<td>dog</td>
</tr>
<tr>
<td>Gallus gallus</td>
<td>red junglefowl</td>
</tr>
</tbody>
</table>
Combine external data with Wolfram Language data and computation

Use WL to see available information:

cow (species specification) -> "PropertyAssociation"

alternate common names -> {domesticated cattle, domestic cattle (feral)}
alternate scientific names -> {Bos indicus, Bos primigenius}, class -> mammals,
common name -> cow, family -> antelopes, buffalo, bison, cattle, gazelles, goats, sheep...
genus -> Bos , image -> , inraspecies -> Missing[NotAvailable],
kingson -> animals , name -> cow , number of members -> 0,
order -> antelope, giraffe, camels, pigs, hippos..., parent entity -> Bos,
phylum -> chordates , scientific name -> Bos taurus , sibling taxa -> { gaur , grunting ox},
...banteng , Entity[Species, Species:BosPrimigenius Species:BosSauveli]},
species -> cow , reference -> Linnaeus, 1758 , sub-entities -> {}, taxonomic sequence ->
{ {animals , chordates } , mammals } , antelopes, giraffe, camels, pigs, hippos...,
{ antelopes, buffalo, bison, cattle, gazelles, goats, sheep... , Bos , cow },
taxonomy graph -> 
cow = `cow (species specification)`["Image"]

ImageIdentify[cow]

grazing land

ImageIdentify[cow, All, 20, "Probability"]

<table>
<thead>
<tr>
<th>Abstract Entity</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>grazing land</td>
<td>0.776094</td>
</tr>
<tr>
<td>grassland</td>
<td>0.776212</td>
</tr>
<tr>
<td>parcel</td>
<td>0.777421</td>
</tr>
<tr>
<td>biome</td>
<td>0.776416</td>
</tr>
<tr>
<td>geographical area</td>
<td>0.777421</td>
</tr>
<tr>
<td>biotic community</td>
<td>0.776416</td>
</tr>
<tr>
<td>region</td>
<td>0.777437</td>
</tr>
<tr>
<td>group</td>
<td>0.776446</td>
</tr>
<tr>
<td>location</td>
<td>0.784891</td>
</tr>
<tr>
<td>abstract entity</td>
<td>0.77679</td>
</tr>
<tr>
<td>placental mammal</td>
<td>0.146045</td>
</tr>
<tr>
<td>mammal</td>
<td>0.146376</td>
</tr>
<tr>
<td>vertebrate</td>
<td>0.151195</td>
</tr>
<tr>
<td>chordate</td>
<td>0.151217</td>
</tr>
<tr>
<td>hoofed mammal</td>
<td>0.109682</td>
</tr>
<tr>
<td>animal</td>
<td>0.15193</td>
</tr>
<tr>
<td>bovid</td>
<td>0.0639287</td>
</tr>
<tr>
<td>ruminant</td>
<td>0.0650192</td>
</tr>
<tr>
<td>even-toed ungulate</td>
<td>0.0693848</td>
</tr>
<tr>
<td>horse (domestic)</td>
<td>0.0399478</td>
</tr>
</tbody>
</table>

ImageIdentify[cow, "Animal", 20, "Probability"]

<table>
<thead>
<tr>
<th>Cattle Breed</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Charolais</td>
<td>0.0842726</td>
</tr>
<tr>
<td>Brown Swiss</td>
<td>0.0215915</td>
</tr>
<tr>
<td>greyhound</td>
<td>0.0089906</td>
</tr>
<tr>
<td>Durham</td>
<td>0.00796973</td>
</tr>
<tr>
<td>whippet</td>
<td>0.00660723</td>
</tr>
<tr>
<td>longhorn</td>
<td>0.0060638</td>
</tr>
<tr>
<td>Staffordshire bull terrier</td>
<td>0.00301595</td>
</tr>
<tr>
<td>Italian greyhound</td>
<td>0.00238337</td>
</tr>
<tr>
<td>Hereford</td>
<td>0.00226538</td>
</tr>
<tr>
<td>beef cattle</td>
<td>0.10067</td>
</tr>
<tr>
<td>zebu</td>
<td>0.0682376</td>
</tr>
<tr>
<td>basset hound</td>
<td>0.0657347</td>
</tr>
<tr>
<td>gemsbok</td>
<td>0.0309806</td>
</tr>
<tr>
<td>moufflon</td>
<td>0.0235162</td>
</tr>
<tr>
<td>dairy cattle</td>
<td>0.0215915</td>
</tr>
<tr>
<td>shire horse</td>
<td>0.0203784</td>
</tr>
<tr>
<td>Clydesdale horse</td>
<td>0.00868348</td>
</tr>
<tr>
<td>collie dog</td>
<td>0.00416024</td>
</tr>
<tr>
<td>Alaskan malamute</td>
<td>0.00407353</td>
</tr>
<tr>
<td>bull terrier</td>
<td>0.00355402</td>
</tr>
</tbody>
</table>
Google images

Search for "brown swiss cow" on Google Images.
"bos taurus" in Wolfram|Alpha

WolframAlpha["bos taurus"]

Input interpretation:
cow (animal)

Scientific name:
Bos taurus

Alternate scientific names:
Bos indicus | Bos primigenius

Alternate common names:
domesticated cattle | domestic cattle (feral)

Taxonomy:

<table>
<thead>
<tr>
<th>kingdom</th>
<th>Animalia (animals)</th>
</tr>
</thead>
<tbody>
<tr>
<td>phylum</td>
<td>Chordata (chordates)</td>
</tr>
<tr>
<td>class</td>
<td>Mammalia (mammals)</td>
</tr>
<tr>
<td>order</td>
<td>Artiodactyla (antelope, giraffe, camels, pigs, hippos...)</td>
</tr>
<tr>
<td>family</td>
<td>Bovidae (antelopes, buffalo, bison, cattle, gazelles, goats, sheep...)</td>
</tr>
<tr>
<td>genus</td>
<td>Bos</td>
</tr>
<tr>
<td>species</td>
<td>Bos taurus (cow)</td>
</tr>
</tbody>
</table>

Biological properties:

Basic properties:

maximum recorded lifespan | (20 to 25) years

Sensory organs:

- number of tastebuds | 25 000 (human: 9000)
- eyeball diameter | 1.6 inches (human: 0.94 in)
- eyeball volume | 1.9 in³ (cubic inches) (human: 0.4 in³)
- photoreceptors per square millimeter | 360 000 (human: 160 000)
- eardrum surface area | 86 mm² (square millimeters) (human: 55 mm²)

Internal organs:

- brain weight | 1.2 lb (pounds) (human: 2.6 to 3.3 lb)
- single hemisphere brain surface area | 500 cm² (square centimeters) (human: 1100 cm²)
<table>
<thead>
<tr>
<th>Trait</th>
<th>Value</th>
<th>(human: value range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>cardiac output</td>
<td>2800 in³/min</td>
<td>(340 to 370 in³/min)</td>
</tr>
<tr>
<td>liver weight</td>
<td>(29 to 48) lb</td>
<td>(human: 4.6 lb)</td>
</tr>
<tr>
<td>lung weight</td>
<td>(80 to 130) lb</td>
<td></td>
</tr>
<tr>
<td>stomach capacity</td>
<td>67 gallons</td>
<td></td>
</tr>
<tr>
<td>digestive tract capacity</td>
<td>94 gallons</td>
<td></td>
</tr>
<tr>
<td>length of whole intestine</td>
<td>170 feet</td>
<td>(20 to 26 ft)</td>
</tr>
<tr>
<td>small intestine capacity</td>
<td>17 gallons</td>
<td></td>
</tr>
<tr>
<td>large intestine and rectum capacity</td>
<td>7.4 gallons</td>
<td></td>
</tr>
<tr>
<td>appendix capacity</td>
<td>2.6 gallons</td>
<td></td>
</tr>
</tbody>
</table>

(typical values, except where otherwise noted)
Animalia
Chordata
Mammalia
Artiodactyla
Bovidae

(nodes sorted by taxonomic diversity)

Wikipedia page hits history:

(in hits per day)

(based on weekly averages of daily hits to English-language "cow" page)
Chemistry functionality in Wolfram|Alpha: Chemical reactions

\[ \text{WolframAlpha} \left[ \text{"Al} + \text{O}_2 \rightarrow \text{Al}_2\text{O}_3\" \right] \]

Input interpretation:
\[ \text{Al (aluminum) } + \text{O}_2 (oxygen) \rightarrow \text{Al}_2\text{O}_3 \text{ (aluminum oxide)} \]

Balanced equation:
\[ 4 \text{ Al} + 3 \text{O}_2 \rightarrow 2 \text{Al}_2\text{O}_3 \]

Structures:
\[ \text{Al} \quad \overset{\text{oxygen}}{\centerdot} \quad \overset{\text{Al}_2\text{O}_3}{\text{Al}} \]

Names:
\[ \text{aluminum } + \text{oxygen } \rightarrow \text{aluminum oxide} \]

Reaction thermodynamics:

**Enthalpy:**
\[ \Delta H^0_{\text{rxn}} = -3352 \text{ kJ/mol} - 0 \text{ kJ/mol} = -3352 \text{ kJ/mol} \text{ (exothermic)} \]

**Entropy:**
\[ \Delta S^0_{\text{rxn}} = 102 \text{ J/(molK)} - 728.2 \text{ J/(molK)} = -626.2 \text{ J/(molK)} \text{ (exoentropic)} \]

Equilibrium constant:
\[ K_c = \frac{[\text{Al}_2\text{O}_3]^2}{[\text{Al}]^4[\text{O}_2]^3} \]

Chemical names and formulas:

<table>
<thead>
<tr>
<th></th>
<th>aluminum</th>
<th>oxygen</th>
<th>aluminum oxide</th>
</tr>
</thead>
<tbody>
<tr>
<td>formula</td>
<td>Al</td>
<td>(\text{O}_2)</td>
<td>(\text{Al}_2\text{O}_3)</td>
</tr>
<tr>
<td>name</td>
<td>aluminum</td>
<td>oxygen</td>
<td>aluminum oxide</td>
</tr>
<tr>
<td>IUPAC name</td>
<td>aluminum</td>
<td>molecular oxygen</td>
<td>oxo–oxoalumanyloxyalumane</td>
</tr>
</tbody>
</table>

Substance properties:

<table>
<thead>
<tr>
<th></th>
<th>aluminum</th>
<th>oxygen</th>
<th>aluminum oxide</th>
</tr>
</thead>
<tbody>
<tr>
<td>molar mass</td>
<td>26.9815386 g/mol</td>
<td>31.9988 g/mol</td>
<td>101.961 g/mol</td>
</tr>
<tr>
<td>phase</td>
<td>solid (at STP)</td>
<td>gas (at STP)</td>
<td>solid (at STP)</td>
</tr>
<tr>
<td>Property</td>
<td>Value 1</td>
<td>Value 2</td>
<td>Value 3</td>
</tr>
<tr>
<td>--------------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>melting point</td>
<td>660.4 °C</td>
<td>-218 °C</td>
<td>2040 °C</td>
</tr>
<tr>
<td>boiling point</td>
<td>2460 °C</td>
<td>-183 °C</td>
<td></td>
</tr>
<tr>
<td>density</td>
<td>2.7 g/cm³</td>
<td>0.001429 g/cm³</td>
<td>(at 0 °C)</td>
</tr>
<tr>
<td>solubility in water</td>
<td>insoluble</td>
<td></td>
<td></td>
</tr>
<tr>
<td>surface tension</td>
<td>0.817 N/m</td>
<td>0.01347 N/m</td>
<td></td>
</tr>
<tr>
<td>dynamic viscosity</td>
<td>$1.5 \times 10^{-4}$ Pas (at 760 °C)</td>
<td>$2.055 \times 10^{-5}$ Pas (at 25 °C)</td>
<td></td>
</tr>
<tr>
<td>odor</td>
<td>odorless</td>
<td>odorless</td>
<td>odorless</td>
</tr>
</tbody>
</table>
Chemistry functionality in Wolfram|Alpha: Chemical quantities

WolframAlpha["500mg of silver nitrate"]

Assuming milligrams for "mg" | Use megagrams instead

Input interpretation:
500 mg of silver nitrate

Basic properties for 500 mg:

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>mass</td>
<td>0.5 grams</td>
</tr>
<tr>
<td>molar amount</td>
<td>2.94 mmol (millimoles)</td>
</tr>
<tr>
<td>equivalents</td>
<td>0.0029 eq (equivalents)</td>
</tr>
<tr>
<td></td>
<td>of silver(I) (ion)</td>
</tr>
<tr>
<td></td>
<td>0.0029 eq (equivalents)</td>
</tr>
<tr>
<td></td>
<td>of nitrate (ion)</td>
</tr>
<tr>
<td>(at STP)</td>
<td></td>
</tr>
</tbody>
</table>

Thermodynamic properties for 500 mg:

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>heat capacity $c_p$</td>
<td>solid 0.274 J/K</td>
</tr>
<tr>
<td>free energy of formation $\Delta_f G^\circ$</td>
<td>solid $-0.09831$ kJ</td>
</tr>
<tr>
<td>heat of formation $\Delta_f H^\circ$</td>
<td>solid $-0.3662$ kJ</td>
</tr>
<tr>
<td>standard entropy $S^\circ$</td>
<td>solid 0.4147 J/K</td>
</tr>
<tr>
<td>latent heat of fusion</td>
<td>0.0345 kJ (kilojoules)</td>
</tr>
<tr>
<td>heat of solution $\Delta_{\text{sol}} H^\circ$</td>
<td>0.06649 kJ (kilojoules)</td>
</tr>
</tbody>
</table>

Phase change energies for 500 mg from 25 °C:

<table>
<thead>
<tr>
<th>Energy required</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>to heat to melting point</td>
<td>0.0512 kJ (kilojoules)</td>
</tr>
<tr>
<td>to convert to liquid</td>
<td>0.0345 kJ (kilojoules)</td>
</tr>
<tr>
<td>to heat to melting point and convert to liquid</td>
<td>0.0857 kJ (kilojoules)</td>
</tr>
</tbody>
</table>

Mass composition for 500 mg:

<table>
<thead>
<tr>
<th>Element</th>
<th>Mass</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>N (nitrogen)</td>
<td>41.227 mg (8.2%)</td>
<td></td>
</tr>
<tr>
<td>O (oxygen)</td>
<td>0.141 g  (28.3%)</td>
<td></td>
</tr>
<tr>
<td>Ag (silver)</td>
<td>0.317 g  (63.5%)</td>
<td></td>
</tr>
</tbody>
</table>
**Structure diagram:**

\[ \text{Ag}^+ \]

\[ \text{O}^- \quad \text{N}^+ \quad \text{O}^- \]

**Chemical names and formulas:**

<table>
<thead>
<tr>
<th>formula</th>
<th>AgNO₃</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hill formula</td>
<td>AgNO₃</td>
</tr>
<tr>
<td>name</td>
<td>silver nitrate</td>
</tr>
<tr>
<td>IUPAC name</td>
<td>silver nitrate</td>
</tr>
</tbody>
</table>

**Substance properties:**

<table>
<thead>
<tr>
<th>property</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>molar mass</td>
<td>169.873 g/mol</td>
</tr>
<tr>
<td>phase (at STP)</td>
<td>solid</td>
</tr>
<tr>
<td>melting point</td>
<td>212 °C</td>
</tr>
<tr>
<td>solubility in water</td>
<td>soluble</td>
</tr>
<tr>
<td>odor</td>
<td>odorless</td>
</tr>
</tbody>
</table>
Chemistry functionality in Wolfram|Alpha: Chemical solutions

\[ \text{WolframAlpha["concentrated acetic acid"]} \]

**Input interpretation:**

concentrated (17 M) acetic acid

**Physical properties:**

- **Surface tension:** 27.73 mN/m (millinewtons per meter)
- **Dynamic viscosity:** 1.727 mPas (millipascal seconds)
- **Refractive index:** 1.374
- **Vapor pressure:** 2.147 kPa (kilopascals)

**Solution properties:**

- **Molarity:** 1.7 daM (dekanormal)
- **Normality:** 17 N (normal)
- **Molality:** 634 molal

**Solute properties per 1 L:**

- **Acetic acid**
  - Molar amount: 17 mol (moles)
  - Mass: 1.02 kg (kilograms)
  - Volume: 973 mL (milliliters)

**Solvent properties per 1 L:**

- **Water**
  - Volume: 26.8 mL (milliliters)
  - Mass: 26.8 grams
  - Molar amount: 1.49 mol (moles)

**Acid–base information:**

- \( K_a \): 0.0000175
- \( pK_a \): 4.76
<table>
<thead>
<tr>
<th>pH</th>
<th>1.76</th>
</tr>
</thead>
<tbody>
<tr>
<td>$[\text{H}_3\text{O}^+]$</td>
<td>0.0173 mol/L (moles per liter)</td>
</tr>
<tr>
<td>pOH</td>
<td>12.2</td>
</tr>
<tr>
<td>$[\text{OH}^-]$</td>
<td>$5.79 \times 10^{-13}$ mol/L (moles per liter)</td>
</tr>
<tr>
<td>% ionization</td>
<td>0.102%</td>
</tr>
</tbody>
</table>

Chemical names and formulas:

- formula: CH$_3$CO$_2$H
- Hill formula: C$_2$H$_4$O$_2$
- name: acetic acid
- IUPAC name: acetic acid

Substance properties:

- molar mass: 60.052 g/mol
- phase: liquid (at STP)
- melting point: 16.2 °C
- boiling point: 117.5 °C
- density: 1.049 g/cm$^3$
- solubility in water: miscible
- surface tension: 0.0288 N/m
- dynamic viscosity: 0.001056 Pas (at 25 °C)
- odor: vinegar-like
- odor threshold: 0.48 ppm
Chemistry functionality in Wolfram|Alpha: Ions

\[ \text{WolframAlpha["weak acid ions"]} \]

Input interpretation:
ionic weak acids

Members:
dihydrogen phosphate | hydrogen carbonate | hydrogen oxalate | hydrogen phosphate | hydrogen sulfate

Structure diagrams:

<table>
<thead>
<tr>
<th>Compound</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>dihydrogen phosphate (ion)</td>
<td>![Structure](dihydrogen phosphate.png)</td>
</tr>
<tr>
<td>hydrogen carbonate (ion)</td>
<td>![Structure](hydrogen carbonate.png)</td>
</tr>
<tr>
<td>hydrogen oxalate (ion)</td>
<td>![Structure](hydrogen oxalate.png)</td>
</tr>
<tr>
<td>hydrogen phosphate (ion)</td>
<td>![Structure](hydrogen phosphate.png)</td>
</tr>
<tr>
<td>hydrogen sulfate (ion)</td>
<td>![Structure](hydrogen sulfate.png)</td>
</tr>
</tbody>
</table>

Ionic radii:

<table>
<thead>
<tr>
<th>Compound</th>
<th>Radius</th>
</tr>
</thead>
<tbody>
<tr>
<td>hydrogen carbonate</td>
<td></td>
</tr>
<tr>
<td>thermochemical radius</td>
<td>156 pm</td>
</tr>
</tbody>
</table>

\[ \text{WolframAlpha} \]
Chemistry functionality in Wolfram|Alpha: Step-by-step solution for Lewis structures

WolframAlpha["lewis dot structure of ATP"]

Input interpretation:

- adenosine triphosphate
- Lewis dot structure

Result:

Draw the Lewis structure of adenosine triphosphate. Start by drawing the overall structure of the molecule, ignoring potential double and triple bonds:

Count the total valence electrons of the carbon ($n_{C,\text{val}} = 4$), hydrogen ($n_{H,\text{val}} = 1$), nitrogen ($n_{N,\text{val}} = 5$), oxygen ($n_{O,\text{val}} = 6$), and phosphorus ($n_{P,\text{val}} = 5$) atoms:

$$10n_{C,\text{val}} + 16n_{H,\text{val}} + 5n_{N,\text{val}} + 13n_{O,\text{val}} + 3n_{P,\text{val}} = 174$$

Calculate the number of electrons needed to completely fill the valence shells for carbon ($n_{C,\text{full}} = 8$), hydrogen ($n_{H,\text{full}} = 2$), nitrogen ($n_{N,\text{full}} = 8$), oxygen ($n_{O,\text{full}} = 8$), and phosphorus ($n_{P,\text{full}} = 8$):

$$10n_{C,\text{full}} + 16n_{H,\text{full}} + 5n_{N,\text{full}} + 13n_{O,\text{full}} + 3n_{P,\text{full}} = 280$$

Subtracting these two numbers shows that $280 - 174 = 106$ bonding electrons are needed. Each bond has two electrons, so in addition to the 49 bonds already present in the diagram we expect to add 4 bonds. To minimize formal charge oxygen wants 2 bonds, nitrogen wants 3 bonds, and carbon wants 4 bonds.
Identity the atoms that want additional bonds and the number of electrons remaining on each atom:

Add 4 bonds by pairing electrons between adjacent highlighted atoms. Additionally, atoms with large electronegativities can minimize their formal charge by forcing atoms with smaller electronegativities on period 3 or higher to expand their valence shells. The electronegativities of the atoms are 2.19 (phosphorus), 2.20 (hydrogen), 2.55 (carbon), 3.04 (nitrogen), and 3.44 (oxygen). Because the electronegativity of phosphorus is smaller than the electronegativity of oxygen, expand the valence shell of phosphorus to 5 bonds in 3 places. Therefore we add a total of 7 bonds to the diagram. Note that the six atom ring is aromatic, so that the single and double bonds may be rearranged:

Answer:
Chemistry functionality in Wolfram|Alpha: Step-by-step solution for chemical balance

\[ \text{WolframAlpha}["0.2 \text{ mol CH}_4 + \text{O}_2 \rightarrow 7 \text{ mL H}_2\text{O} + \text{CO}_2"] \]

**Input interpretation:**

0.200 mol of \( \text{CH}_4 \) (methane) + \( \text{O}_2 \) (oxygen) \( \rightarrow \) 7 mL of \( \text{H}_2\text{O} \) (water) + \( \text{CO}_2 \) (carbon dioxide)

**Balanced equation:**

Balance the chemical equation:

\[ \text{CH}_4 + \text{O}_2 \rightarrow \text{H}_2\text{O} + \text{CO}_2 \]

Add coefficients to all the molecules:

\[ c_1 \text{CH}_4 + c_2 \text{O}_2 \rightarrow c_3 \text{H}_2\text{O} + c_4 \text{CO}_2 \]

The number of C, H, and O atoms on both sides of the reaction must be equal:

- C: \( c_1 = c_4 \)
- H: \( 4c_1 = 2c_3 \)
- O: \( 2c_2 = c_3 + 2c_4 \)

Since the coefficients are only determined up to a multiplicative constant, set \( c_1 = 1 \) and solve for the coefficients:

- \( c_1 = 1 \)
- \( c_2 = 2 \)
- \( c_3 = 2 \)
- \( c_4 = 1 \)

Since the coefficients are all integers with a greatest common denominator equal to 1, substitute the coefficients into the chemical reaction to obtain the balanced equation:

**Answer:**

\[ \text{CH}_4 + 2 \text{O}_2 \rightarrow 2 \text{H}_2\text{O} + \text{CO}_2 \]

**Stoichiometry:**

Find the percent yield of the following reaction given 0.2 mol \( \text{CH}_4 \) and excess:

\[ 0.2 \text{ mol of CH}_4 \text{ (methane)} + \text{O}_2 \text{ (oxygen)} \rightarrow 7 \text{ mL of H}_2\text{O} \text{ (water)} + \text{CO}_2 \text{ (carbon dioxide)} \]

Convert the specified volume of \( \text{H}_2\text{O} \) into moles using the density at STP (1 g/cm\(^3\)) and the molar mass (18.0153 g/mol):

\[ 7 \text{ mL} \left( \frac{1 \text{ cm}^3}{1 \text{ mL}} \right) \left( \frac{1 \text{ g/cm}^3}{1 \text{ g/mol}} \right) \left( \frac{1}{3.886 \text{ mol H}_2\text{O}} \right) = 0.3886 \text{ mol H}_2\text{O} \]
Make a table of the molar quantities of the reagents corresponding to 0.2 mol CH₄. Begin by filling in this molar quantity:

<table>
<thead>
<tr>
<th></th>
<th>CH₄</th>
<th>O₂</th>
<th>H₂O</th>
<th>CO₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2 mol CH₄</td>
<td>0.2 mol</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Write the balanced equation for the reaction:

\[ \text{CH₄} + 2 \text{O}_2 \rightarrow 2 \text{H}_2\text{O} + \text{CO}_2 \]

Use the ratios of coefficients in the balanced equation to compute the molar quantities of the remaining reagents corresponding to 0.2 mol CH₄:

\[
(0.2 \text{ mol CH}_4) \left( \frac{2 \text{ mol O}_2}{1 \text{ mol CH}_4} \right) = 0.4 \text{ mol O}_2
\]

\[
(0.2 \text{ mol CH}_4) \left( \frac{2 \text{ mol H}_2\text{O}}{1 \text{ mol CH}_4} \right) = 0.4 \text{ mol H}_2\text{O}
\]

\[
(0.2 \text{ mol CH}_4) \left( \frac{1 \text{ mol CO}_2}{1 \text{ mol CH}_4} \right) = 0.2 \text{ mol CO}_2
\]

Summarize the results of the previous step in the table:

<table>
<thead>
<tr>
<th></th>
<th>CH₄</th>
<th>O₂</th>
<th>H₂O</th>
<th>CO₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2 mol CH₄</td>
<td>0.2 mol</td>
<td>0.4 mol</td>
<td>0.4 mol</td>
<td>0.2 mol</td>
</tr>
</tbody>
</table>

The theoretical yields of the products will be: 0.4 mol H₂O and 0.2 mol CO₂

Given the actual yield 7 mL (0.3886 mol) of H₂O, the percent yield is:

Answer:

\[
100\% \left( \frac{0.3886 \text{ mol H}_2\text{O}}{0.4 \text{ mol H}_2\text{O}} \right) = 97.14\% \text{ (yield H}_2\text{O)}
\]
Enthalpy:
\[ \Delta H_{\text{rxn}}^0 = -965.2 \text{ kJ/mol} - 74.6 \text{ kJ/mol} = -890.6 \text{ kJ/mol} \text{ (exothermic)} \]

Gibbs free energy:
\[ \Delta G_{\text{rxn}}^0 = -868.6 \text{ kJ/mol} - 412.4 \text{ kJ/mol} = -1281 \text{ kJ/mol} \text{ (exergonic)} \]

Entropy:
\[ \Delta S_{\text{rxn}}^0 = 353.8 \text{ J/(mol K)} - 596 \text{ J/(mol K)} = -242.2 \text{ J/(mol K)} \text{ (exoentropic)} \]

Equilibrium constant:
\[ K_c = \frac{[\text{H}_2\text{O}]^2 [\text{CO}_2]}{[\text{CH}_4][\text{O}_2]^2} \]

Chemical names and formulas:

<table>
<thead>
<tr>
<th></th>
<th>methane</th>
<th>oxygen</th>
<th>water</th>
<th>carbon dioxide</th>
</tr>
</thead>
<tbody>
<tr>
<td>formula</td>
<td>CH$_4$</td>
<td>O$_2$</td>
<td>H$_2$O</td>
<td>CO$_2$</td>
</tr>
<tr>
<td>name</td>
<td>methane</td>
<td>oxygen</td>
<td>water</td>
<td>carbon dioxide</td>
</tr>
<tr>
<td>IUPAC name</td>
<td>methane</td>
<td>molecular oxygen</td>
<td>water</td>
<td>carbon dioxide</td>
</tr>
</tbody>
</table>

Substance properties:

<table>
<thead>
<tr>
<th></th>
<th>methane</th>
<th>oxygen</th>
<th>water</th>
<th>carbon dioxide</th>
</tr>
</thead>
<tbody>
<tr>
<td>molar mass</td>
<td>16.0425 g/mol</td>
<td>31.9988 g/mol</td>
<td>18.0153 g/mol</td>
<td>44.0095 g/mol</td>
</tr>
<tr>
<td>phase</td>
<td>gas (at STP)</td>
<td>gas (at STP)</td>
<td>liquid (at STP)</td>
<td>gas (at STP)</td>
</tr>
<tr>
<td>melting point</td>
<td>-182.47 °C</td>
<td>-218 °C</td>
<td>0 °C</td>
<td>-56.56 °C (at triple point)</td>
</tr>
<tr>
<td>boiling point</td>
<td>-161.48 °C</td>
<td>-183 °C</td>
<td>99.9839 °C</td>
<td>-78.5 °C (at sublimation point)</td>
</tr>
<tr>
<td>density</td>
<td>6.67151 × 10$^{-4}$ g/cm$^3$ (at 20 °C)</td>
<td>0.001429 g/cm$^3$ (at 0 °C)</td>
<td>1 g/cm$^3$</td>
<td>0.00184212 g/cm$^3$ (at 20 °C)</td>
</tr>
<tr>
<td>solubility in water</td>
<td>soluble</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>surface tension</td>
<td>0.0137 N/m</td>
<td>0.01347 N/m</td>
<td>0.0728 N/m</td>
<td></td>
</tr>
<tr>
<td>dynamic viscosity</td>
<td>1.114 × 10$^{-5}$ Pas (at 25 °C)</td>
<td>2.055 × 10$^{-5}$ Pas (at 25 °C)</td>
<td>8.9 × 10$^{-4}$ Pas (at 25 °C)</td>
<td>1.491 × 10$^{-5}$ Pas (at 25 °C)</td>
</tr>
<tr>
<td>odor</td>
<td>odorless</td>
<td>odorless</td>
<td>odorless</td>
<td>odorless</td>
</tr>
</tbody>
</table>
Future work

- finish and release chemical functional groups, foods, ....
- apply machine learning techniques (chemistry and elsewhere)
- more data sets
- big data
Thanks

Michael Trott
Bjorn Zimmermann
Armin Vollmer