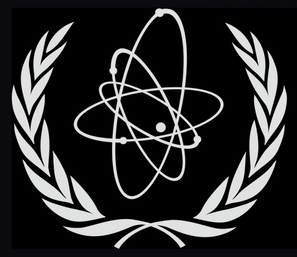


IAEA
International Atomic Energy Agency

CollisionDB and ALADDIN2

Support for nuclear fusion energy research

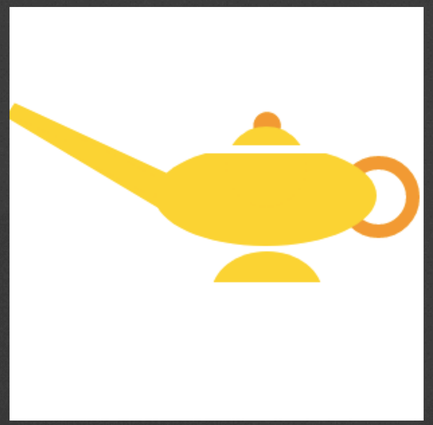
Machine Learning Coffee Morning

IAEA Headquarters, Vienna, Austria
10 November 2022

Christian Hill, Kalle Heinola, Dipti,
International Atomic Energy Agency

CollisionDB and ALADDIN2

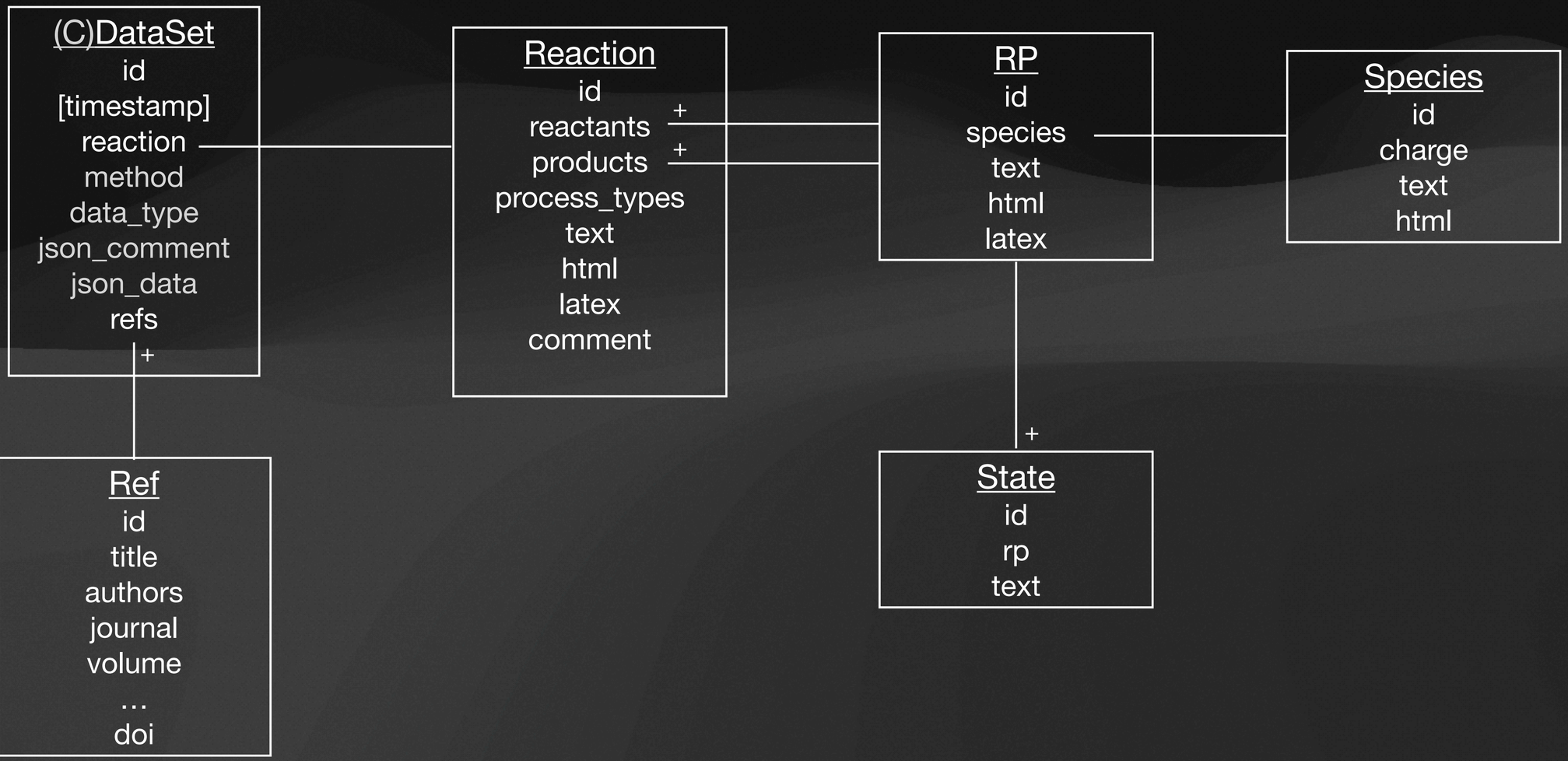
Plasma collisional processes



- Collisional cross sections, differential cross sections or rate coefficients
- Searchable online at https://db-amdis.org/collisiondb/
- CollisionDB: 106,372 data sets, ALADDIN2: 15,987 data sets
- API
- Python Package for interacting with the API, data exploration, data transformation, plotting, etc.

CollisionDB and ALADDIN2

Relational Database Schema



How FAIR are we?

Findability, Accessibility, Interoperability, Reusability

- ✓ All data associated with rich metadata, and assigned a persistent, globally-unique identifier
- ✓ Data retrievable by search and identifier from a browser and through an API
- ✓ Data provided in a formal, accessible, standardised format, with provenance (e.g. DOI)
- ✓ Data provided under a clear licence, with relevant attributes and unambiguous metadata (units, origin, pertinent experimental or computational details)

https://www.go-fair.org/

CollisionDB

Downloaded data format


```
{"qid": "D104629", "reaction": "e- + H2 X(1\u03a3+g);v=2 \u2192 H2+ + 2e-", "p
rocess_types": {"EIN": "Ionization"}, "data_type": "cross section", "refs": {"
B35": {"doi": "10.1016/j.adt.2021.101424"}}, "json_comment": {"comment": "Cros
s sections for non-dissociative ionization of H2 using the semi-classical Gryz
inski method"}, "json_data": {"method": "GM", "columns": [{"name": "E", "units
": "eV"}, {"name": "sigma", "units": "cm2"}], "unc_perc": 50.0}}
-------------------------------------------------------------------------
1.443e+01 0.0
1.467e+01 2.207e-19
1.493e+01 6.208e-19
1.518e+01 1.169e-18
1.544e+01 1.906e-18
1.571e+01 2.835e-18
1.598e+01 3.915e-18
1.625e+01 5.11e-18
1.653e+01 6.413e-18
1.682e+01 7.838e-18
1.710e+01 9.4e-18
1.740e+01 1.109e-17
1.770e+01 1.286e-17
1.800e+01 1.47e-17
1.831e+01 1.66e-17
1.862e+01 1.855e-17
1.894e+01 2.053e-17
1.927e+01 2.254e-17
```

CollisionDB

Download data format: JSON metadata


```
qid:              "D104629"
reaction:         "e− + H2 X(1Σ+g);v=2 → H2+ + 2e−"
process_types:
  EIN:            "Ionization"
data_type:        "cross section"
refs:
  B35:
    doi:          "10.1016/j.adt.2021.101424"
json_comment:
  comment:        "Cross sections for non-d…ssical Gryzinski method"
json_data:
  method:         "GM"
  columns:
    0:
      name:       "E"
      units:      "eV"
    1:
      name:       "sigma"
      units:      "cm2"
  unc_perc:       50
```

CollisionDB

Download data format: JSON metadata


```
qid:            "D104629"
reaction:       "e- + H2 X(1Σ+g);v=2 → H2+ + 2e-"
process_types:
  EIN:          "Ionization"
data_type:      "cross section"
refs:
  B35:
    doi:        "10.1016/j.adt.2021.101424"
json_comment:
  comment:      "Cross sections for non-d…ssical Gryzinski method"
json_data:
  method:       "GM"
  columns:
    0:
      name:     "E"
      units:    "eV"
    1:
      name:     "sigma"
      units:    "cm2"
  unc_perc:     50
```


https://github.com/xnx/pyvalem

https://amdis.iaea.org/databases/processes

https://www.doi.org/

https://db-amdis.org/collisiondb/theoretical-methods/

https://github.com/xnx/pyqn

CollisionDB

Upload data format

- Validation
- Unit conversion / standardization
- Canonicalize reaction
- DOI resolution
- Numerical checks
- Physics checks

```
comment="Cross sections for non-d…ssical Gryzinski method"
method="GM"
doi=["10.1016/j.adt.2021.101424"]
data_type="cross section"
uncertainty="50%"
columns=["E, eV", "sigma, a0^2"]
reaction="e- + H2 X(1SIGMA+g);v=2 -> H2+ + 2e-"
process_types=["EIN"]

E              sigma
1.443e+01      0.0
1.467e+01      2.207e-19
1.493e+01      6.208e-19
...
```

```
qid:            "D104629"
reaction:       "e− + H2 X(1Σ+g);v=2 → H2+ + 2e−"
process_types:
  EIN:          "Ionization"
data_type:      "cross section"
refs:
  B35:
    doi:        "10.1016/j.adt.2021.101424"
json_comment:
  comment:      "Cross sections for non-d…ssical Gryzinski method"
json_data:
  method:       "GM"
  columns:
    0:
      name:     "E"
      units:    "eV"
    1:
      name:     "sigma"
      units:    "cm2"
  unc_perc:     50
```

PyCollisionDB: API library

Example

In [1]:
```
from pycollisiondb.pycollisiondb import PyCollision
import matplotlib.pyplot as plt
%matplotlib notebook
```

In [2]:
```
# Fetch datasets from the server over the API.
# Proton-impact ionization of H.

query = {'reactants': ['H+', 'H 1s'],
         'process_types': ['HIN'],
         'data_type': 'cross section'}
pycoll = PyCollision.get_datasets(query=query)
```

In [3]:
```
pycoll.datasets
```

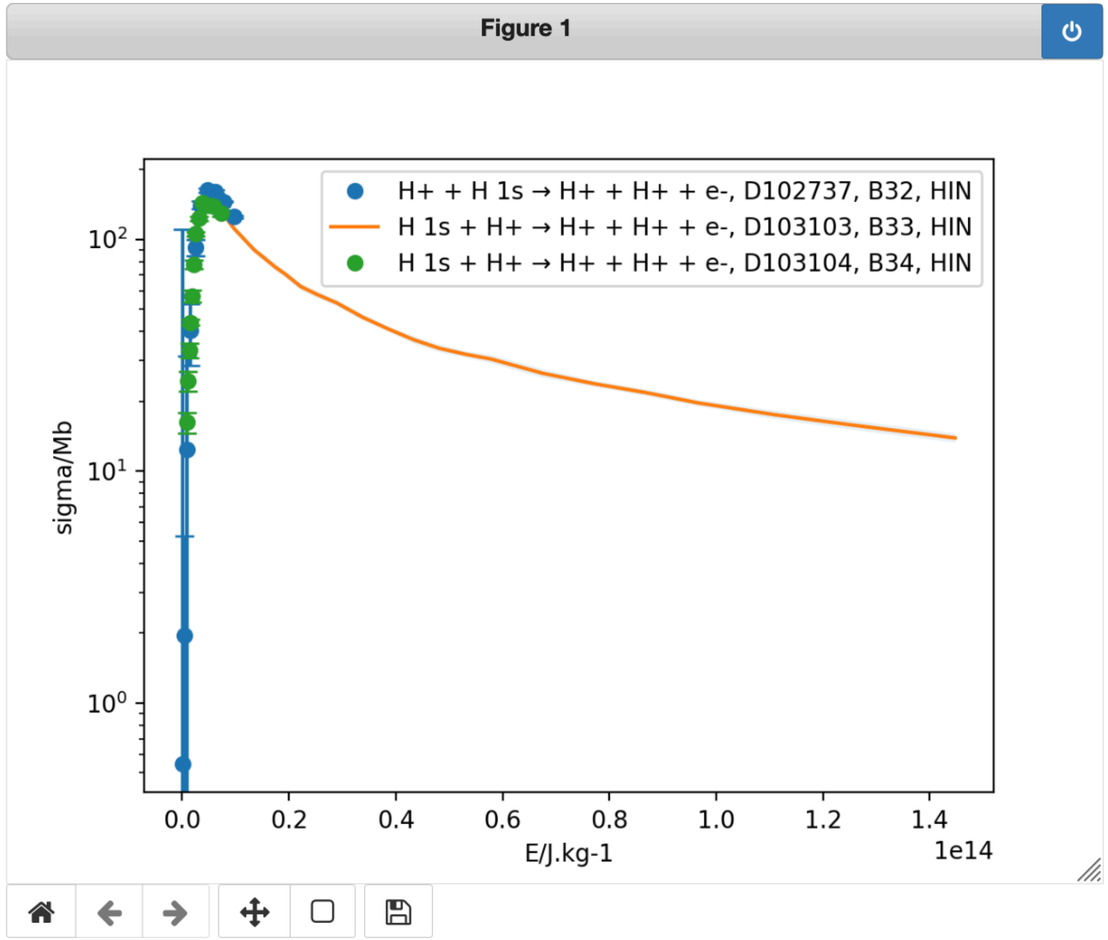
Out[3]:
```
{102737: D102737: H+ + H 1s → H+ + H+ + e-,
 103103: D103103: H 1s + H+ → H+ + H+ + e-,
 103104: D103104: H 1s + H+ → H+ + H+ + e-}
```

In [4]:
```
pycoll.convert_units({'sigma': 'Mb', 'E': 'J.kg-1'})
```

PyCollisionDB: API library

Example

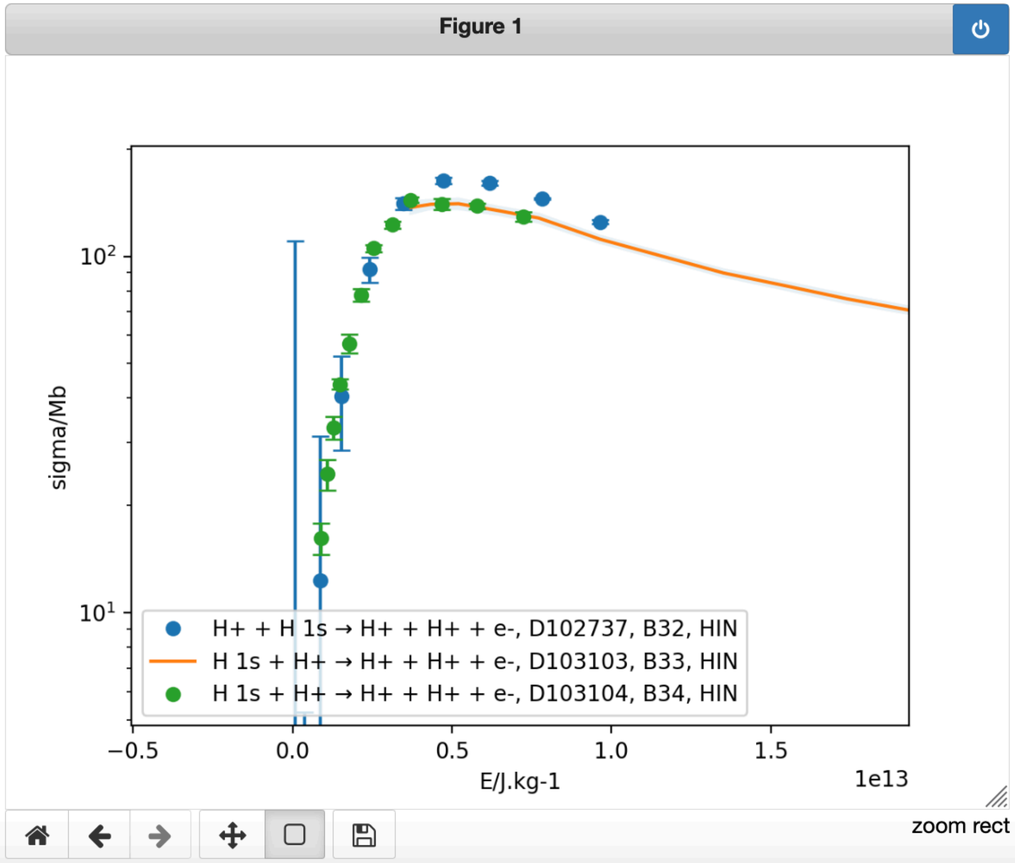
In [7]:
```
fig, ax = plt.subplots()
pycoll.plot_all_datasets(ax, label=('reaction', 'qid', 'refs', 'process_types'))
plt.legend()
```



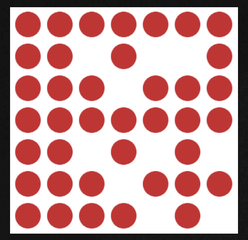
PyCollisionDB: API library

Example: Interactive plots

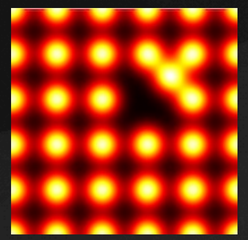
```
In [7]: fig, ax = plt.subplots()
        pycoll.plot_all_datasets(ax, label=('reaction', 'qid', 'refs', 'process_types'))
        plt.legend()
```



Other Databases



CascadesDB: molecular dynamics simulations of collision cascades for nuclear materials (https://cascadesdb.iaea.org/)



DefectDB: DFT calculations of radiation-induced defects in nuclear materials (https://db-amdis.org/defectdb)

AMBDAS: Bibliographic database for atomic and molecular data in fusion (https://amdis.iaea.org/db/ambdas)

hcdb: Hierarchical database of heterogeneous data sets related to CRP outputs (https://db-amdis.org/hcdb)