CSV

code:**csv**

Basic information about the format

CSV stands for "coma separated value" and it is very simple molecule format.

```
id,mol,registeting_user,note
1,C,anonymous@chemicalize.com,this is a rather common element
2,[H],h.canvenids@chemicalize.com,"I bet this is more common, how could you miss it?"
3,[He],pjc_janssen@chemicalize.com,This is boring il ne reagit pas avec quoi que ce soit!
```

In this file we have 3 molecules, and every of them has the following information:

- ID
- registering_user
- note

The molecule sources are in smiles. After import we get the following structures and properties:

- A simple Carbon, with:
  - ID = 1
  - registering_user = anonymous@chemicalize.com
  - note = this is a rather common element
- A simple Hydrogen, with:
  - ID = 2
  - registering_user = h.canvenids@chemicalize.com
  - note = I bet this is more common, how could you miss it?
- A simple Helium, with:
  - ID = 3
  - registering_user = pjc_janssen@chemicalize.com
  - note = This is boring il ne reagit pas avec quoi que ce soit!

But the user can specify molecule during import which header to use. For example this file:

```
id,CHEMICAL_DATA,name
1,c1ccccc1CC(N)C,amphetamin
2,c1ccccc1,benzene
```

Can be imported with the following settings:

```
csv:strucCHEMICAL_DATA
```

With this MolImporter recognise that CHEMICAL_DATA filed holds the structure.

Import options

Headers

Automatically recognized molecule headers

Molecule can have any ChemAxon supported formats, but they must be written in one line. The recognized molecule headers are:

- mol
- molecule
- structure
- struc
- smiles
- cxsmiles
- smarts
- cxsmarts
- inchi

User defined header

User can define which header to use as identifier of the molecule column when importing structure. This can be done with the "**struc**" parameter.

For example this file:

```
id,CHEMICAL_DATA,name
1,c1ccccc1CC(N)C,amphetamin
2,c1ccccc1,benzene
```

Can be imported with the following settings:

```
csv:strucCHEMICAL_DATA
```

With this MolImporter recognise that
CHEMICAL_DATA filed holds the structure.

Headless import

User can import CSV molecules without header, in this case csv importer must be informed that all rows are data (for this use "**headless**" keyword), and the which colum has the chemical structure. This can be done by defining the zero-based index of the structure column. For example the following file

```
7,12,4,ccCCcc,rt,gh,jk
23,1,56,COO,rf,gg,kk
```

Can be imported as:

```
csv:headless,struc3
```

This would import the following structure:

- ccCCcc (as smiles) with the following properties:
  - column_0 = 7
  - column_1 = 12
  - column_2 = 4
  - column_3 = rt
  - column_4 = gh
  - column_5 = jk
- COO (as smiles) with the following properties:
  - column_0 = 23
  - column_1 = 1
  - column_2 = 56
  - column_3 = rf
  - column_4 = gg
  - column_5 = kk

Override column names

During import user can dynamically ovverride column names. For this he has to set the names in order. (Every definition starts with an "**f**" and spearated by coma".) For example this file:

```
result,hour
S.[He],11:15:00
[He],11:10:00
```

can be imported as:

- S.[He]
  - TIME = 11:15:00

- [He]
  - TIME = 11:10:00

With the following params:

```
csv:fMOL,fTIME
```

In the above example the renamed headers contained an autoreconizable header name, so we did not have to specifiy molecule colum. But this can be than as it is described in Header section with the "**struc**" keyword.?

Molecule format

User can specify what is the format of the molecules in the molecule comulmn with the "**input** " keyword. For example fro names use:

```
csv:inputname
```

Export options

Define Molecule column name:

User can set the name of the molecule column with "**struc**" keyword, like:

```
csv:strucMY_MOL_COLUMN
```

Define headless export

User can export molecules without headers with the "**headless**" keyword, like:

```
csv:headless
```

Define export format

User can define which format to use when export molecule with the "**format**" keyword, like:

```
csv:formatsmarts
```

Define exported column header names

It is possible to define the name of the exported clumns every name must start with an "**s**" like:

```
csv:sname,smol,suser
```