
NAME

InfoFingerprintsSDFiles.pl - List information about fingerprints data in SDFFile(s)

SYNOPSIS

InfoFingerprintsSDFiles.pl SDFFile(s)...

```
InfoFingerprintsSDFiles.pl [-a, --all] [--AverageBitDensity] [--BitDensity] [-c, --count] [-d, --detail InfoLevel]
[--DataCheck] [-e, --empty] [--FingerprintsField FieldLabel] [--FingerprintsFormatMode Internal | Specify] [
--FingerprintsString Hexadecimal | Binary | RawBinary] [--FingerprintsType] [--FingerprintsStringType] [
--FingerprintsSize] [-h, --help] [--OnBits] [-w, --WorkingDir dirname] SDFFile(s)...
```

DESCRIPTION

List information about fingerprints data in *SDFFile(s)*: number of compounds containing fingerprints data, type and size of fingerprint, bit density and average bit density of bit-based fingerprints, and so on.

Multiple SDFFile names are separated by spaces. The valid file extensions are *.sdf* and *.sd*. All other file names are ignored. All the SD files in a current directory can be specified either by **.sdf* or the current directory name.

OPTIONS

-a, --all

List all the available information

--AverageBitDensity

List average bit density of bit-based fingerprints data

--BitDensity

List bit density of bit-based fingerprints data in each row

-c, --count

List number of rows containing fingerprints data. This is default behavior.

-d, --detail *InfoLevel*

Level of information to print about lines being ignored. Default: *1*. Possible values: *1*, *2* or *3*

--DataCheck

Validate fingerprints data specified using **--FingerprintsCol** and list information about missing and invalid data.

-e, --empty

List number of rows containing no fingerprints data.

--FingerprintsField *FieldLabel*

Fingerprints field label to use during calculations of similarity matrices for *SDFFile(s)*. Default value: *first data field label containing the word Fingerprints in its label*.

--FingerprintsFormatMode *Internal | Specify*

Specify format of fingerprints data in *SDFFile(s)*: use default format which MayaChemTools fingerprint generation scripts use to write out fingerprints data or explicitly specify format of fingerprints. Possible values: *Internal | Specify*. Default value: *Internal*.

Internal fingerprints string format consists of four parts delimited by semicolon: `<Type:StringType:Size:String>`. For example:

```
"PathLength:Binary:512:010011..."
"MDLKeys166FP:Binary:166:010011..."
"MDLKeys166Count:Vector:166:0 1 2..."
```

For *Specify* value of **--FingerprintsFormatMode** option, **--FingerprintsString** is used to interpret fingerprints string.

--FingerprintsString *Hexadecimal | Binary | RawBinary*

Format of fingerprints string during *Specify* value of **--FingerprintsFormatMode** option. Possible values:

Hexadecimal, Binary, or RawBinary. Default value: *none*; its value must be explicitly specified.

- FingerprintsType
List types of fingerprints data
- FingerprintsStringType
List types of fingerprint strings
- FingerprintsSize
List size of fingerprints
- h, --help
Print this help message
- OnBits
List number of on bits in bit-based fingerprints data for each rowx.
- w, --WorkingDir *DirName*
Location of working directory. Default: current directory

EXAMPLES

To count compounds containing fingerprints data present in a data field with Fingerprint substring in its label, type:

```
% InfoFingerprintsSDFiles.pl -a SampleFP1.sdf
```

To list all available information about fingerprints data in any internal format present in a data field with Fingerprint substring in its label, type:

```
% InfoFingerprintsSDFiles.pl -a SampleFP1.sdf
```

To list all available information about fingerprints data in any internal format present in a data field named PathLengthFingerprints, type:

```
% InfoFingerprintsSDFiles.pl -a --FingerprintsField PathLengthFingerprints  
SampleFP2.sdf
```

To list all available information about fingerprints data in hexadecimal bit-string format present in a data field name PathLengthFingerprints, type:

```
% InfoFingerprintsSDFiles.pl -a --FingerprintsField PathLengthFingerprints  
--FingerprintsFormatMode Specify --FingerprintsString Hexadecimal  
SampleFP2.sdf
```

AUTHOR

Manish Sud <msud@san.rr.com>

SEE ALSO

InfoFingerprintsTextFiles.pl, PathLengthFingerprints.pl

COPYRIGHT

Copyright (C) 2004-2008 Manish Sud. All rights reserved.

This file is part of MayaChemTools.

MayaChemTools is free software; you can redistribute it and/or modify it under the terms of the GNU Lesser General Public License as published by the Free Software Foundation; either version 3 of the License, or (at your option) any later version.