

Python file magic

May 8, 2007 / Gavin Jackson

[code](#)[linux](#)[programming](#)[python](#)

File magic is used on Linux and Unix systems as a binary "fingerprint" to tell what type a file actually is - this provides a much better result than MIME or filename extensions.

The "magic" pattern is a text file definition that specifies header/footer and content binary patterns for file types on Ubuntu the file is located /usr/share/file/magic.

Example of an entry in the magic file:

```
#-----  
  
# rtf: file(1) magic for Rich Text Format (RTF)  
#  
  
# Duncan P. Simpson, D.P.Simpson@dcs.warwick.ac.uk  
#  
0 string {\rtf Rich Text Format data,  
>5 byte x version %c,  
>6 string \ansi ANSI  
>6 string \mac Apple Macintosh  
>6 string \pc IBM PC, code page 437  
>6 string \pca IBM PS/2, code page 850
```

We can interface with file magic using python by installing the python-magic package (apt-get install python-magic).

```
import magic  
  
ms = magic.open(magic.MAGIC_NONE)  
ms.load()  
type = ms.file("/path/to/some/file")  
print type  
  
f = file("/path/to/some/file", "r")  
buffer = f.read(4096)  
f.close()  
  
type = ms.buffer(buffer)  
print type  
  
ms.close()
```

Downloaded from <https://www.gavinj.net/post/python-file-magic>

Generated September 4, 2026. Copyright Gavin Jackson. All rights reserved.