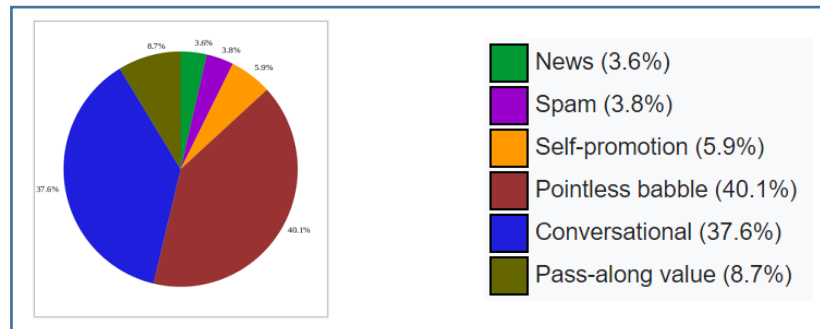




Twitter es una red social muy versátil que puede ser utilizada para muchos propósitos. El uso más común suele ser la conversación, pero hay usuarios que, por ejemplo, usan esta plataforma como herramienta de trabajo para compartir noticias, contenidos e ideas acerca de una gran cantidad de temas. El siguiente gráfico tomado de Wikipedia muestra la distribución del tipo de *tweets* escritos en la red social. Como se puede observar, las categorías “conversación” (37%) y “cháchara sin sentido” (40%) copan 4 de cada 5 *tweets*:

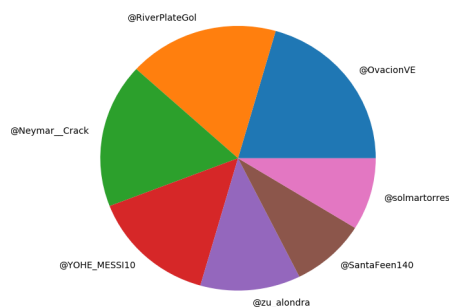


En este ejercicio desarrollaremos un sistema que permitirá realizar un análisis simple sobre una colección de 1000 *tweets* en formato JSON recuperados mediante la consulta ['messi', 'ronaldo'] el 28 de junio de 2017. Para cada tweet se dispone de cuatro informaciones (autor, fecha, retweets, texto). He aquí un fragmento de un fichero de entrada con la información correspondiente a los dos primeros tweets:

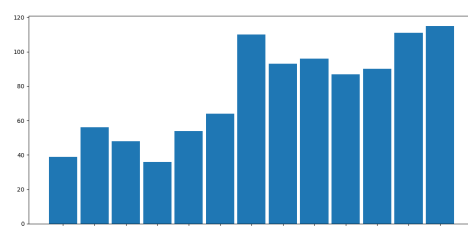
```
[
  {
    "texto": "RT @ImMESS10nante: El tatuador de Messi Roberto López lo tiene claro ...",
    "autor": "@lamboglianare01",
    "fecha": "Wed Jun 28 15:49:54 +0000 2017",
    "retweets": 4251
  },
  {
    "texto": "RT @mariomassaccesi: #Ahora #Urgente #Rosario Operativo de la #AFIP ...",
    "autor": "@MafiaoRepublica",
    "fecha": "Wed Jun 28 15:49:54 +0000 2017",
    "retweets": 12
  },
]
```

Fragmento del fichero de entrada

El ejercicio consiste en la implementación de un módulo Python con una serie de funciones que permitirán leer los datos de entrada, extraer información de *tweets* y obtener ciertos indicadores de actividad a partir de una colección de *tweets*. Las siguientes figuras muestran dos de los posibles resultados del análisis:



Usuarios más activos y *tweets* escritos por ellos



Evolución temporal de *tweets*