In [1]:
```
import nltk
from nltk.corpus import stopwords
from nltk.tokenize import word_tokenize
```

In [2]:
```
def cos_sim(Query, Statement):
    X_list = word_tokenize(Query)
    Y_list = word_tokenize(Statement)
    sw = stopwords.words('english')
    l1 =[];l2 =[]
    X_set = {w for w in X_list if not w in sw}
    Y_set = {w for w in Y_list if not w in sw}
    rvector = X_set.union(Y_set)
    for w in rvector:
        if w in X_set: l1.append(1)
        else: l1.append(0)
        if w in Y_set: l2.append(1)
        else: l2.append(0)
    c = 0
    for i in range(len(rvector)):
        c+= l1[i]*l2[i]
        cosine = c / float((sum(l1)*sum(l2))**0.5)
    return cosine
```

In [3]:
```
q="gold silver truck"
d1="Shipment of gold damaged in fire"
d2="Delivery of silver arrived in a silver truck"
d3="Shipment of gold arrived in a truck"
```

In [4]:
```
cos_sim(q,d1)
```

Out[4]:
```
0.2886751345948129
```

In [5]:
```
cos_sim(q,d2)
```

Out[5]:
```
0.5773502691896258
```

In [6]:
```
cos_sim(q,d3)
```

Out[6]:
```
0.5773502691896258
```

In [ ]:
```

```