In [ ]:

```
## input text article
article_text="Just what is agility in the context of software engineering work? Ivar Jacobs
```

In [ ]:

```
import re
import nltk
from nltk.tokenize import word_tokenize
```

Type *Markdown* and LaTeX: α^2

In [ ]:

```
article_text = article_text.lower()
article_text
```

In [ ]:

```
# remove spaces, punctuations and numbers
clean_text = re.sub('[^a-zA-Z]', ' ', article_text)
clean_text = re.sub('\s+', ' ', clean_text)
clean_text
```

In [ ]:

```
# split into sentence list
sentence_list = nltk.sent_tokenize(article_text)
sentence_list
```

In [ ]:

```
tokens=word_tokenize(article_text)
```

In [ ]:

```
tokens
```

In [ ]:

```
## run this cell once to download stopwords
# import nltk
#nltk.download('stopwords')
```

In [ ]:

```
stopwords = nltk.corpus.stopwords.words('english')

word_frequencies = {}
for word in tokens:
    if word not in stopwords:
        if word not in word_frequencies:
            word_frequencies[word] = 1
        else:
            word_frequencies[word]+= 1
```

In [ ]:

```
word_frequencies
```

In [ ]:

```
import string
from nltk.probability import FreqDist
removethese=set(stopwords+list(string.punctuation)+list(string.digits))
```

In [ ]:

```
word_freq=FreqDist(word_frequencies)
```

In [ ]:

```
word_freq.plot(30,title='freqplot')
```

In [ ]:

```
from collections import Counter
```

In [ ]:

```
dictionary=Counter(word_freq)
```

In [ ]:

```
import matplotlib.pyplot as plt
from wordcloud import WordCloud
```

In [ ]:

```
cloud=WordCloud(max_font_size=80,colormap='hsv').generate_from_frequencies(dictionary)
```

In [ ]:

```
plt.figure(figsize=(16,12))
plt.imshow(cloud,interpolation='bilinear')
plt.axis('off')
plt.show()
```

In [ ]:

```

```