Choose your favourite job sector/sub-sector, then use TF/IDF to extract important keywords. Visualise them in word cloud chart (hint: you can use the online tool https://wordart.com/create or similar websites)**.

In [1]:
```
import matplotlib.pyplot as plt
from sklearn.feature_extraction.text import TfidfVectorizer
from nltk.corpus import stopwords
from wordcloud import WordCloud

def update(Classification="Information & Communication Technology"):
    stop_words = list(stopwords.words('english'))
    requirements = df[df['Classification'] == Classification]['Requirement'].dropna()
    original_documents = [x.strip() for x in requirements]

    max_features = 500
    tfidf = TfidfVectorizer(analyzer='word', ngram_range=(1, 1), min_df=1, stop_words=stop_words, max_features=max_feat

    features = tfidf.fit_transform(original_documents)
    corpus_tf_idf = features

    sum_words = corpus_tf_idf.sum(axis=0)
    words_freq = [(word, sum_words[0, idx]) for word, idx in tfidf.vocabulary_.items()]
    words_freq_dict = {x: y for x,y in words_freq}
    words_freq_dict.items()

    keywords = [word for word, freq in words_freq]
    keyword_freqs = [freq for word, freq in words_freq]

    # Create a word cloud
    wordcloud = WordCloud(width=800, height=400, background_color='white').generate_from_frequencies(dict(words_freq))

    # Plot the word cloud
    plt.figure(figsize=(16, 16))
    plt.imshow(wordcloud, interpolation='bilinear')
    plt.axis('off')
    plt.title('TF-IDF Analysis of Requirements', size=14)
    plt.show()

interact(update, Classification="Information & Communication Technology")
```

→ When I run this code in the Jupyter, it shows the above Wordcloud without any errors, but after I uploaded it to Github, it says “Unable to render code block”. Therefore, I include this attachment separately. All the codes for this question are included in my assignment. Thanks.