In [35]:
```
import sys
print(sys.version)
```

```
3.12.12 (main, Oct 10 2025, 08:52:57) [GCC 11.4.0]
```

Install and Import Required NLP Libraries (NLTK & spaCy)

In [36]:
```
# Install required libraries
!pip install nltk spacy matplotlib seaborn pandas
!python -m spacy download en_core_web_sm
```

```
Requirement already satisfied: nltk in /usr/local/lib/python3.12/dist-packages
(3.9.1)
Requirement already satisfied: spacy in /usr/local/lib/python3.12/dist-packages
(3.8.11)
Requirement already satisfied: matplotlib in /usr/local/lib/python3.12/dist-pac
kages (3.10.0)
Requirement already satisfied: seaborn in /usr/local/lib/python3.12/dist-packag
es (0.13.2)
Requirement already satisfied: pandas in /usr/local/lib/python3.12/dist-package
s (2.2.2)
Requirement already satisfied: click in /usr/local/lib/python3.12/dist-packages
(from nltk) (8.3.1)
Requirement already satisfied: joblib in /usr/local/lib/python3.12/dist-package
s (from nltk) (1.5.3)
Requirement already satisfied: regex>=2021.8.3 in /usr/local/lib/python3.12/dis
t-packages (from nltk) (2025.11.3)
Requirement already satisfied: tqdm in /usr/local/lib/python3.12/dist-packages
(from nltk) (4.67.1)
Requirement already satisfied: spacy-legacy<3.1.0,>=3.0.11 in /usr/local/lib/py
thon3.12/dist-packages (from spacy) (3.0.12)
Requirement already satisfied: spacy-loggers<2.0.0,>=1.0.0 in /usr/local/lib/py
thon3.12/dist-packages (from spacy) (1.0.5)
Requirement already satisfied: murmurhash<1.1.0,>=0.28.0 in /usr/local/lib/pyth
on3.12/dist-packages (from spacy) (1.0.15)
Requirement already satisfied: cymem<2.1.0,>=2.0.2 in /usr/local/lib/python3.1
2/dist-packages (from spacy) (2.0.13)
Requirement already satisfied: preshed<3.1.0,>=3.0.2 in /usr/local/lib/python
3.12/dist-packages (from spacy) (3.0.12)
Requirement already satisfied: thinc<8.4.0,>=8.3.4 in /usr/local/lib/python3.1
2/dist-packages (from spacy) (8.3.10)
Requirement already satisfied: wasabi<1.2.0,>=0.9.1 in /usr/local/lib/python3.1
2/dist-packages (from spacy) (1.1.3)
Requirement already satisfied: srsly<3.0.0,>=2.4.3 in /usr/local/lib/python3.1
2/dist-packages (from spacy) (2.5.2)
Requirement already satisfied: catalogue<2.1.0,>=2.0.6 in /usr/local/lib/python
3.12/dist-packages (from spacy) (2.0.10)
Requirement already satisfied: weasel<0.5.0,>=0.4.2 in /usr/local/lib/python3.1
2/dist-packages (from spacy) (0.4.3)
Requirement already satisfied: typer-slim<1.0.0,>=0.3.0 in /usr/local/lib/pytho
n3.12/dist-packages (from spacy) (0.21.1)
Requirement already satisfied: numpy>=1.19.0 in /usr/local/lib/python3.12/dist-
packages (from spacy) (2.0.2)
Requirement already satisfied: requests<3.0.0,>=2.13.0 in /usr/local/lib/python
3.12/dist-packages (from spacy) (2.32.4)
Requirement already satisfied: pydantic!=1.8,!=1.8.1,<3.0.0,>=1.7.4 in /usr/loc
al/lib/python3.12/dist-packages (from spacy) (2.12.3)
Requirement already satisfied: jinja2 in /usr/local/lib/python3.12/dist-package
s (from spacy) (3.1.6)
Requirement already satisfied: setuptools in /usr/local/lib/python3.12/dist-pac
kages (from spacy) (75.2.0)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.12/dis
t-packages (from spacy) (25.0)
Requirement already satisfied: contourpy>=1.0.1 in /usr/local/lib/python3.12/di
st-packages (from matplotlib) (1.3.3)
```

```
Requirement already satisfied: cycler>=0.10 in /usr/local/lib/python3.12/dist-p
ackages (from matplotlib) (0.12.1)
Requirement already satisfied: fonttools>=4.22.0 in /usr/local/lib/python3.12/d
ist-packages (from matplotlib) (4.61.1)
Requirement already satisfied: kiwisolver>=1.3.1 in /usr/local/lib/python3.12/d
ist-packages (from matplotlib) (1.4.9)
Requirement already satisfied: pillow>=8 in /usr/local/lib/python3.12/dist-pack
ages (from matplotlib) (11.3.0)
Requirement already satisfied: pyparsing>=2.3.1 in /usr/local/lib/python3.12/di
st-packages (from matplotlib) (3.3.1)
Requirement already satisfied: python-dateutil>=2.7 in /usr/local/lib/python3.1
2/dist-packages (from matplotlib) (2.9.0.post0)
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.12/dist-p
ackages (from pandas) (2025.2)
Requirement already satisfied: tzdata>=2022.7 in /usr/local/lib/python3.12/dis
t-packages (from pandas) (2025.3)
Requirement already satisfied: annotated-types>=0.6.0 in /usr/local/lib/python
3.12/dist-packages (from pydantic!=1.8,!=1.8.1,<3.0.0,>=1.7.4->spacy) (0.7.0)
Requirement already satisfied: pydantic-core==2.41.4 in /usr/local/lib/python
3.12/dist-packages (from pydantic!=1.8,!=1.8.1,<3.0.0,>=1.7.4->spacy) (2.41.4)
Requirement already satisfied: typing-extensions>=4.14.1 in /usr/local/lib/pyth
on3.12/dist-packages (from pydantic!=1.8,!=1.8.1,<3.0.0,>=1.7.4->spacy) (4.1
5.0)
Requirement already satisfied: typing-inspection>=0.4.2 in /usr/local/lib/pytho
n3.12/dist-packages (from pydantic!=1.8,!=1.8.1,<3.0.0,>=1.7.4->spacy) (0.4.2)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.12/dist-packa
ges (from python-dateutil>=2.7->matplotlib) (1.17.0)
Requirement already satisfied: charset_normalizer<4,>=2 in /usr/local/lib/pytho
n3.12/dist-packages (from requests<3.0.0,>=2.13.0->spacy) (3.4.4)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.12/dist-p
ackages (from requests<3.0.0,>=2.13.0->spacy) (3.11)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.12/
dist-packages (from requests<3.0.0,>=2.13.0->spacy) (2.5.0)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.12/
dist-packages (from requests<3.0.0,>=2.13.0->spacy) (2026.1.4)
Requirement already satisfied: blis<1.4.0,>=1.3.0 in /usr/local/lib/python3.12/
dist-packages (from thinc<8.4.0,>=8.3.4->spacy) (1.3.3)
Requirement already satisfied: confection<1.0.0,>=0.0.1 in /usr/local/lib/pytho
n3.12/dist-packages (from thinc<8.4.0,>=8.3.4->spacy) (0.1.5)
Requirement already satisfied: cloudpathlib<1.0.0,>=0.7.0 in /usr/local/lib/pyt
hon3.12/dist-packages (from weasel<0.5.0,>=0.4.2->spacy) (0.23.0)
Requirement already satisfied: smart-open<8.0.0,>=5.2.1 in /usr/local/lib/pytho
n3.12/dist-packages (from weasel<0.5.0,>=0.4.2->spacy) (7.5.0)
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.12/dis
t-packages (from jinja2->spacy) (3.0.3)
Requirement already satisfied: wrapt in /usr/local/lib/python3.12/dist-packages
(from smart-open<8.0.0,>=5.2.1->weasel<0.5.0,>=0.4.2->spacy) (2.0.1)
Collecting en-core-web-sm==3.8.0
  Downloading https://github.com/explosion/spacy-models/releases/download/en_co
re_web_sm-3.8.0/en_core_web_sm-3.8.0-py3-none-any.whl (12.8 MB)
     ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 12.8/12.8 MB 92.0 MB/s eta 0:00:0
0
✔ Download and installation successful
You can now load the package via spacy.load('en_core_web_sm')
```

⚠ Restart to reload dependencies

If you are in a Jupyter or Colab notebook, you may need to restart Python in order to load all the package's dependencies. You can do this by selecting the 'Restart kernel' or 'Restart runtime' option.

In [38]:
```
# Import libraries
import nltk
import spacy
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

In [39]:
```
# Download NLTK resources
nltk.download('punkt')
nltk.download('averaged_perceptron_tagger')
```

```
[nltk_data] Downloading package punkt to /root/nltk_data...
[nltk_data]   Package punkt is already up-to-date!
[nltk_data] Downloading package averaged_perceptron_tagger to
[nltk_data]     /root/nltk_data...
[nltk_data]   Package averaged_perceptron_tagger is already up-to-
[nltk_data]       date!
```

Out[39]: True

Load and Display the Essay Text

In [40]:
```
essay_text = """
Academic research plays a crucial role in advancing knowledge and fostering ir
Researchers analyze data systematically to evaluate hypotheses and develop the
The study emphasizes critical thinking, evidence-based reasoning, and methodol
Effective communication of findings contributes to scholarly discourse and pol
"""
```

Tokenize the Essay into Words (Using NLTK)

In [41]:
```
from nltk.tokenize import word_tokenize

tokens = word_tokenize(essay_text)
print(tokens)
```

```
['Academic', 'research', 'plays', 'a', 'crucial', 'role', 'in', 'advancing', 'k
nowledge', 'and', 'fostering', 'innovation', '.', 'Researchers', 'analyze', 'da
ta', 'systematically', 'to', 'evaluate', 'hypotheses', 'and', 'develop', 'theor
ies', '.', 'The', 'study', 'emphasizes', 'critical', 'thinking', ',', 'evidenc
e-based', 'reasoning', ',', 'and', 'methodological', 'rigor', '.', 'Effective',
'communication', 'of', 'findings', 'contributes', 'to', 'scholarly', 'discours
e', 'and', 'policy', 'development', '.']
```

Apply POS Tagging Using NLTK

In [42]:
```
nltk_pos_tags = nltk.pos_tag(tokens)
```

```python
# Display tagged words
for word, tag in nltk_pos_tags:
    print(f"{word} -> {tag}")
```

```
Academic -> NNP
research -> NN
plays -> VBZ
a -> DT
crucial -> JJ
role -> NN
in -> IN
advancing -> VBG
knowledge -> NN
and -> CC
fostering -> JJ
innovation -> NN
. -> .
Researchers -> NNP
analyze -> VBP
data -> NNS
systematically -> RB
to -> TO
evaluate -> VB
hypotheses -> NNS
and -> CC
develop -> VB
theories -> NNS
. -> .
The -> DT
study -> NN
emphasizes -> VBZ
critical -> JJ
thinking -> NN
, -> ,
evidence-based -> JJ
reasoning -> NN
, -> ,
and -> CC
methodological -> JJ
rigor -> NN
. -> .
Effective -> JJ
communication -> NN
of -> IN
findings -> NNS
contributes -> NNS
to -> TO
scholarly -> JJ
discourse -> NN
and -> CC
policy -> NN
development -> NN
. -> .
```

Apply POS Tagging Using spaCy (Universal Tags)

In [43]:
```
nlp = spacy.load("en_core_web_sm")
doc = nlp(essay_text)

spacy_pos_tags = [(token.text, token.pos_) for token in doc]

for word, tag in spacy_pos_tags:
    print(f"{word} -> {tag}")
```

 -> SPACE
Academic -> ADJ
research -> NOUN
plays -> VERB
a -> DET
crucial -> ADJ
role -> NOUN
in -> ADP
advancing -> VERB
knowledge -> NOUN
and -> CCONJ
fostering -> VERB
innovation -> NOUN
. -> PUNCT

 -> SPACE
Researchers -> NOUN
analyze -> VERB
data -> NOUN
systematically -> ADV
to -> PART
evaluate -> VERB
hypotheses -> NOUN
and -> CCONJ
develop -> VERB
theories -> NOUN
. -> PUNCT

 -> SPACE
The -> DET
study -> NOUN
emphasizes -> VERB
critical -> ADJ
thinking -> NOUN
, -> PUNCT
evidence -> NOUN
- -> PUNCT
based -> VERB
reasoning -> NOUN
, -> PUNCT
and -> CCONJ
methodological -> ADJ
rigor -> NOUN
. -> PUNCT

 -> SPACE
Effective -> ADJ
communication -> NOUN
of -> ADP
findings -> NOUN
contributes -> VERB
to -> ADP
scholarly -> ADJ
discourse -> NOUN

```
and -> CCONJ
policy -> NOUN
development -> NOUN
. -> PUNCT

 -> SPACE
```

Compare NLTK and spaCy POS Tags

In [44]:
```
comparison_df = pd.DataFrame({
    "Word": [word for word, _ in nltk_pos_tags],
    "NLTK_Tag": [tag for _, tag in nltk_pos_tags],
    "spaCy_Tag": [token.pos_ for token in doc if not token.is_punct]
})

comparison_df
```

Out[44]:

	Word	NLTK_Tag	spaCy_Tag
0	Academic	NNP	SPACE
1	research	NN	ADJ
2	plays	VBZ	NOUN
3	a	DT	VERB
4	crucial	JJ	DET
5	role	NN	ADJ
6	in	IN	NOUN
7	advancing	VBG	ADP
8	knowledge	NN	VERB
9	and	CC	NOUN
10	fostering	JJ	CCONJ
11	innovation	NN	VERB
12	.	.	NOUN
13	Researchers	NNP	SPACE
14	analyze	VBP	NOUN
15	data	NNS	VERB
16	systematically	RB	NOUN
17	to	TO	ADV
18	evaluate	VB	PART
19	hypotheses	NNS	VERB
20	and	CC	NOUN
21	develop	VB	CCONJ
22	theories	NNS	VERB
23	.	.	NOUN
24	The	DT	SPACE
25	study	NN	DET
26	emphasizes	VBZ	NOUN
27	critical	JJ	VERB
28	thinking	NN	ADJ
29	,	,	NOUN
30	evidence-based	JJ	NOUN

	Word	NLTK_Tag	spaCy_Tag
31	reasoning	NN	VERB
32	,	,	NOUN
33	and	CC	CCONJ
34	methodological	JJ	ADJ
35	rigor	NN	NOUN
36	.	.	SPACE
37	Effective	JJ	ADJ
38	communication	NN	NOUN
39	of	IN	ADP
40	findings	NNS	NOUN
41	contributes	NNS	VERB
42	to	TO	ADP
43	scholarly	JJ	ADJ
44	discourse	NN	NOUN
45	and	CC	CCONJ
46	policy	NN	NOUN
47	development	NN	NOUN
48	.	.	SPACE

Identify Nouns (Concepts) and Verbs (Arguments)

In [45]:
```
# Using spaCy tags for clarity
nouns = [token.text.lower() for token in doc if token.pos_ == "NOUN"]
verbs = [token.text.lower() for token in doc if token.pos_ == "VERB"]

print("Nouns (Academic Concepts):", nouns)
print("Verbs (Arguments):", verbs)
```

```
Nouns (Academic Concepts): ['research', 'role', 'knowledge', 'innovation', 'res
earchers', 'data', 'hypotheses', 'theories', 'study', 'thinking', 'evidence',
'reasoning', 'rigor', 'communication', 'findings', 'discourse', 'policy', 'deve
lopment']
Verbs (Arguments): ['plays', 'advancing', 'fostering', 'analyze', 'evaluate',
'develop', 'emphasizes', 'based', 'contributes']
```

Calculate Frequency of Nouns and Verbs

In [46]:
```
from collections import Counter
```

```
noun_freq = Counter(nouns)
verb_freq = Counter(verbs)

print("Noun Frequencies:\n", noun_freq)
print("\nVerb Frequencies:\n", verb_freq)
```

```
Noun Frequencies:
 Counter({'research': 1, 'role': 1, 'knowledge': 1, 'innovation': 1, 'researche
rs': 1, 'data': 1, 'hypotheses': 1, 'theories': 1, 'study': 1, 'thinking': 1,
'evidence': 1, 'reasoning': 1, 'rigor': 1, 'communication': 1, 'findings': 1,
'discourse': 1, 'policy': 1, 'development': 1})

Verb Frequencies:
 Counter({'plays': 1, 'advancing': 1, 'fostering': 1, 'analyze': 1, 'evaluate':
1, 'develop': 1, 'emphasizes': 1, 'based': 1, 'contributes': 1})
```

Organize Frequencies into Structured DataFrames

In [47]:
```
noun_df = pd.DataFrame(noun_freq.items(), columns=["Noun", "Frequency"])
verb_df = pd.DataFrame(verb_freq.items(), columns=["Verb", "Frequency"])

noun_df, verb_df
```

Out[47]:
```
(            Noun  Frequency
 0       research          1
 1           role          1
 2      knowledge          1
 3     innovation          1
 4    researchers          1
 5           data          1
 6     hypotheses          1
 7       theories          1
 8          study          1
 9       thinking          1
 10      evidence          1
 11     reasoning          1
 12         rigor          1
 13  communication          1
 14       findings          1
 15     discourse          1
 16        policy          1
 17   development          1,
          Verb  Frequency
 0       plays          1
 1   advancing          1
 2   fostering          1
 3     analyze          1
 4    evaluate          1
 5     develop          1
 6  emphasizes          1
 7       based          1
 8  contributes          1)
```

Visualize Noun and Verb Frequencies

In [50]:
```
plt.figure(figsize=(12,5))

plt.subplot(1,2,1)
sns.barplot(x="Frequency", y="Noun", data=noun_df)
plt.title("Noun Frequency (Academic Concepts)")

plt.subplot(1,2,2)
sns.barplot(x="Frequency", y="Verb", data=verb_df)
plt.title("Verb Frequency (Arguments)")

plt.tight_layout()
plt.show()
```