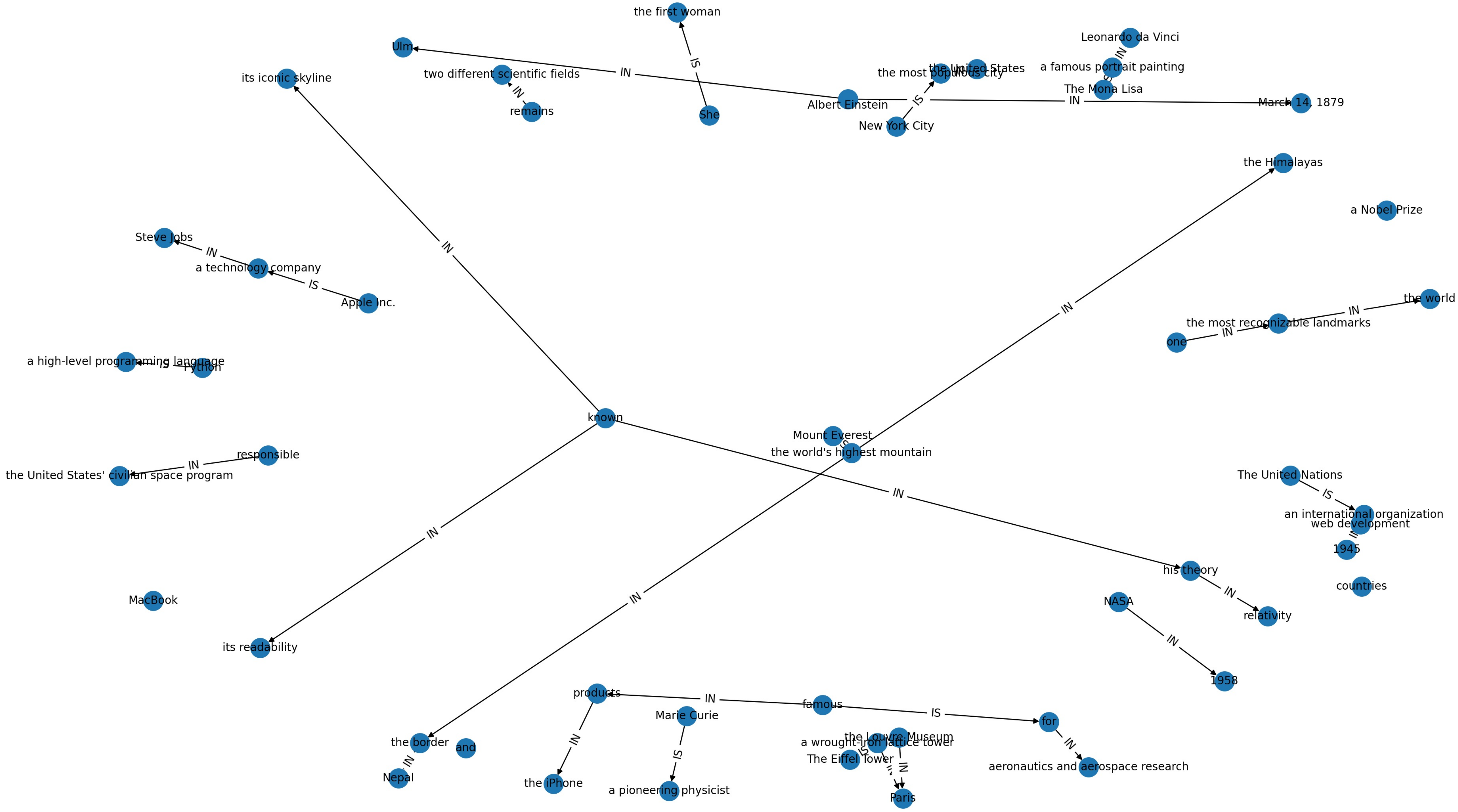
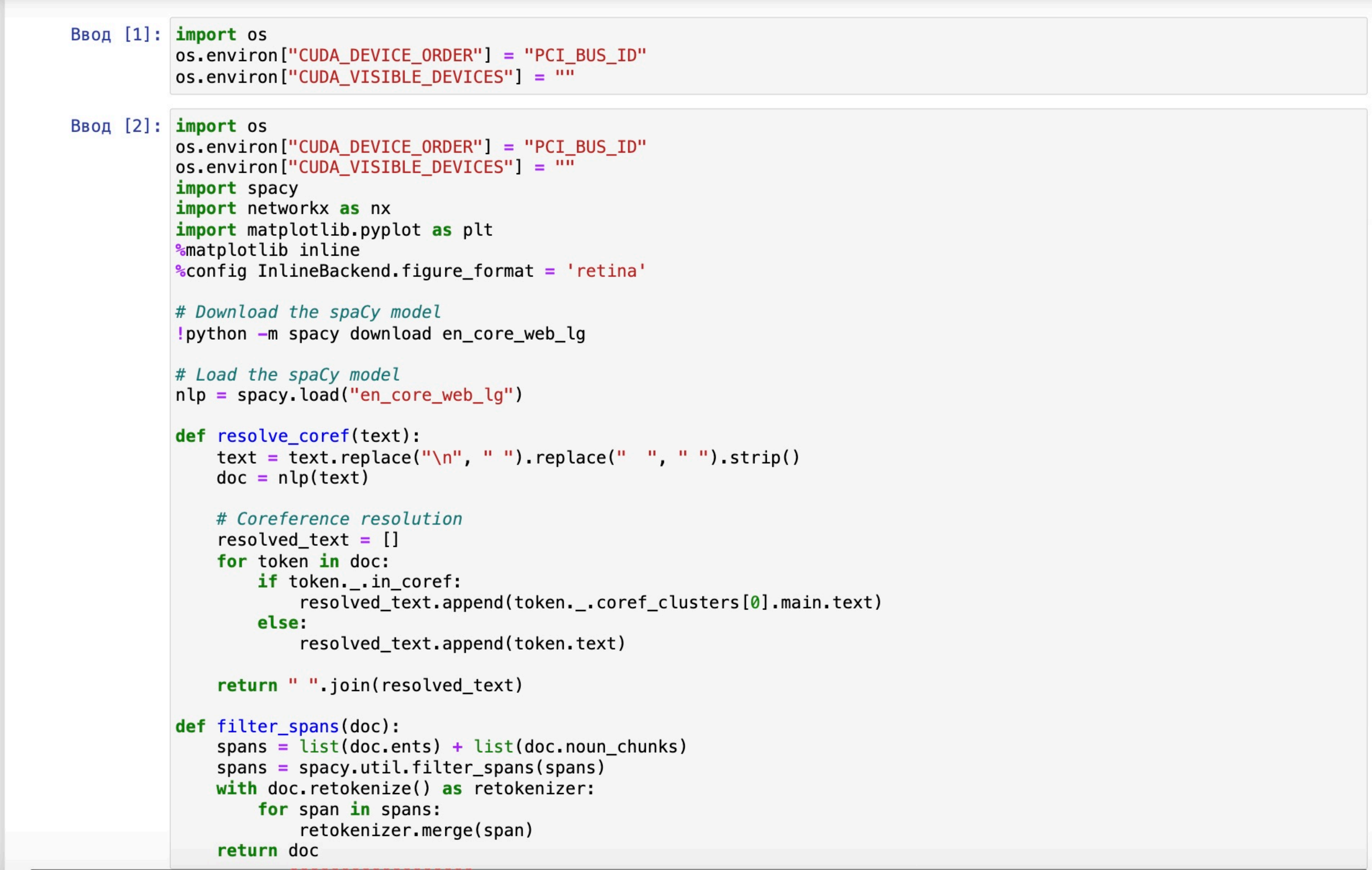

the first woman
Leonardo da Vinci
Ulm
IS
two different scientific fields
its iconic skyline
the most populous city
the United States
a famous portrait painting
IN
The Mona Lisa
IN
IN
IS
IN
remains
Albert Einstein
She
March 14, 1879
New York City
the Himalayas
a Nobel Prize
Steve Jobs
IN
IN
a technology company
IS
Apple Inc.
IN
the world
the most recognizable landmarks
IN
one
a high-level programming language
IS
Python
known
Mount Everest
IS
the world's highest mountain
responsible
IN
the United States' civilian space program
IN
The United Nations
IS
an international organization
web development
IN
IN
1945
his theory
countries
MacBook
IN
NASA
relativity
IN
its readability
IN
1958
products
IN
famous
IS
for
Marie Curie
the Louvre Museum
a wrought-iron lattice tower
the border
and
IN
IN
IS
IS
IN
IN
The Eiffel Tower
aeronautics and aerospace research
Nepal
the iPhone
a pioneering physicist
Paris

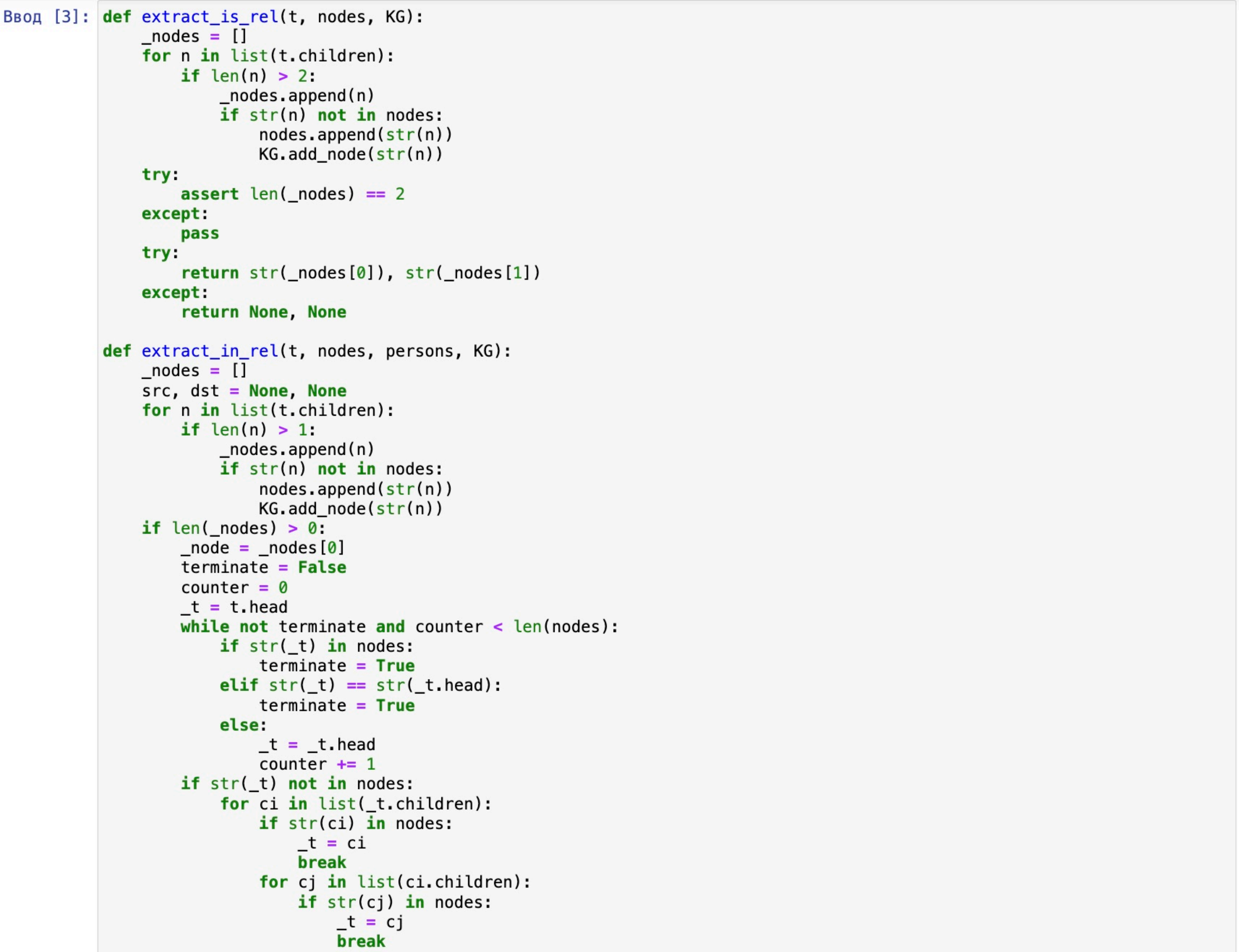


Ввод [1]:

```
import os
os.environ["CUDA_DEVICE_ORDER"] = "PCI_BUS_ID"
os.environ["CUDA_VISIBLE_DEVICES"] = ""
```

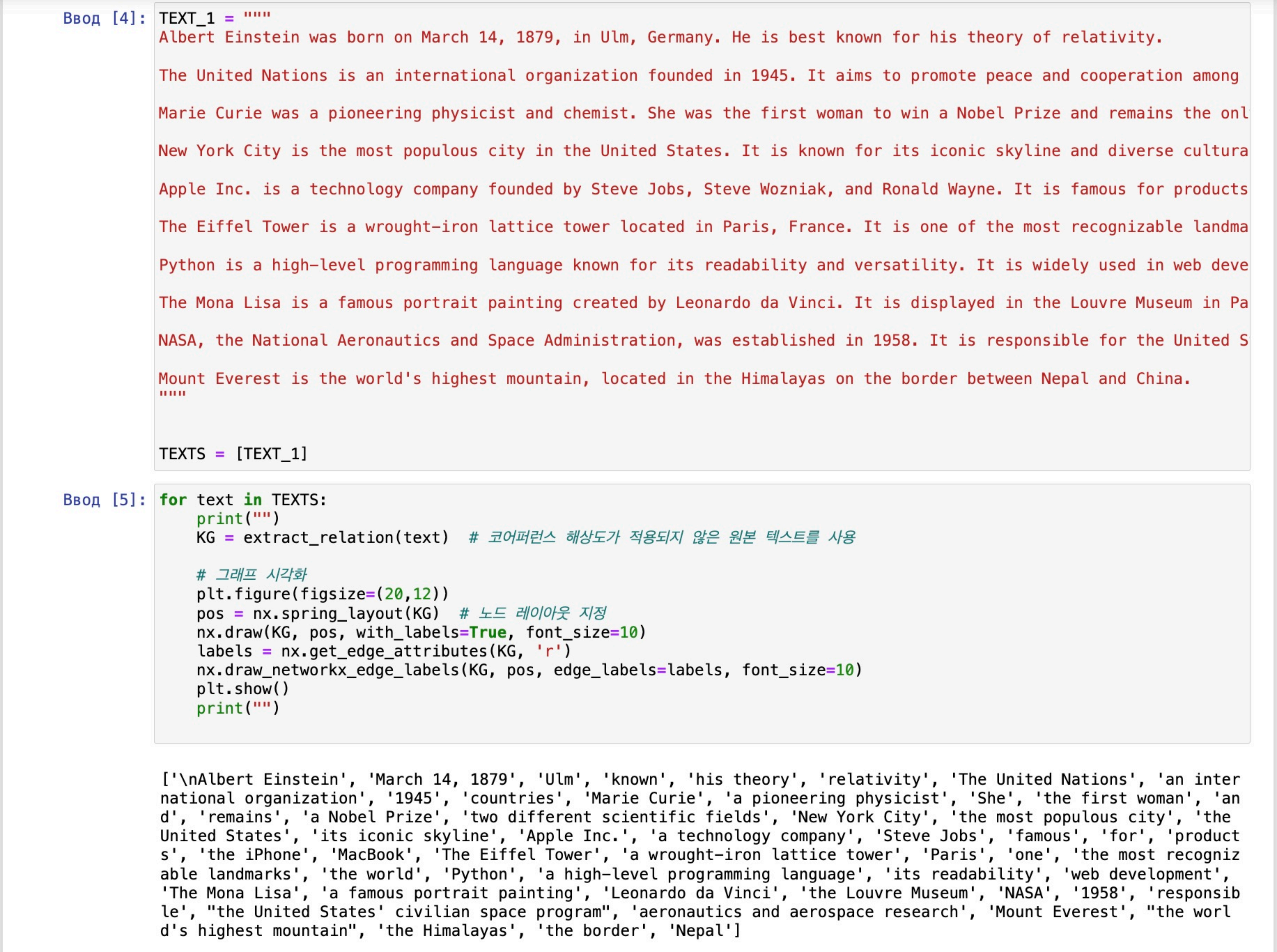
Ввод [2]:

```
import os
os.environ["CUDA_DEVICE_ORDER"] = "PCI_BUS_ID"
os.environ["CUDA_VISIBLE_DEVICES"] = ""
import spacy
import networkx as nx
import matplotlib.pyplot as plt
%matplotlib inline
%config InlineBackend.figure_format = 'retina'

# Download the spaCy model
!python -m spacy download en_core_web_lg

# Load the spaCy model
nlp = spacy.load("en_core_web_lg")

def resolve_coref(text):
    text = text.replace("\n", " ").replace("  ", " ").strip()
    doc = nlp(text)

    # Coreference resolution
    resolved_text = []
    for token in doc:
        if token._.in_coref:
            resolved_text.append(token._.coref_clusters[0].main.text)
        else:
            resolved_text.append(token.text)

    return " ".join(resolved_text)

def filter_spans(doc):
    spans = list(doc.ents) + list(doc.noun_chunks)
    spans = spacy.util.filter_spans(spans)
    with doc.retokenize() as retokenizer:
        for span in spans:
            retokenizer.merge(span)
    return doc
```



Ввод [3]:

```
def extract_is_rel(t, nodes, KG):
    _nodes = []
    for n in list(t.children):
        if len(n) > 2:
            _nodes.append(n)
            if str(n) not in nodes:
                nodes.append(str(n))
                KG.add_node(str(n))
    try:
        assert len(_nodes) == 2
    except:
        pass
    try:
        return str(_nodes[0]), str(_nodes[1])
    except:
        return None, None

def extract_in_rel(t, nodes, persons, KG):
    _nodes = []
    src, dst = None, None
    for n in list(t.children):
        if len(n) > 1:
            _nodes.append(n)
            if str(n) not in nodes:
                nodes.append(str(n))
                KG.add_node(str(n))
    if len(_nodes) > 0:
        _node = _nodes[0]
        terminate = False
        counter = 0
        _t = t.head
        while not terminate and counter < len(nodes):
            if str(_t) in nodes:
                terminate = True
            elif str(_t) == str(_t.head):
                terminate = True
            else:
                _t = _t.head
                counter += 1
        if str(_t) not in nodes:
            for ci in list(_t.children):
                if str(ci) in nodes:
                    _t = ci
                    break
                for cj in list(ci.children):
                    if str(cj) in nodes:
                        _t = cj
                        break
```



Ввод [4]:
```
TEXT_1 = """
Albert Einstein was born on March 14, 1879, in Ulm, Germany. He is best known for his theory of relativity.

The United Nations is an international organization founded in 1945. It aims to promote peace and cooperation among

Marie Curie was a pioneering physicist and chemist. She was the first woman to win a Nobel Prize and remains the onl

New York City is the most populous city in the United States. It is known for its iconic skyline and diverse cultura

Apple Inc. is a technology company founded by Steve Jobs, Steve Wozniak, and Ronald Wayne. It is famous for products

The Eiffel Tower is a wrought-iron lattice tower located in Paris, France. It is one of the most recognizable landma

Python is a high-level programming language known for its readability and versatility. It is widely used in web deve

The Mona Lisa is a famous portrait painting created by Leonardo da Vinci. It is displayed in the Louvre Museum in Pa

NASA, the National Aeronautics and Space Administration, was established in 1958. It is responsible for the United S

Mount Everest is the world's highest mountain, located in the Himalayas on the border between Nepal and China.
"""


TEXTS = [TEXT_1]
```

Ввод [5]:
```
for text in TEXTS:
    print("")
    KG = extract_relation(text)  # 코어퍼런스 해상도가 적용되지 않은 원본 텍스트를 사용

    # 그래프 시각화
    plt.figure(figsize=(20,12))
    pos = nx.spring_layout(KG)  # 노드 레이아웃 지정
    nx.draw(KG, pos, with_labels=True, font_size=10)
    labels = nx.get_edge_attributes(KG, 'r')
    nx.draw_networkx_edge_labels(KG, pos, edge_labels=labels, font_size=10)
    plt.show()
    print("")
```

```
['\nAlbert Einstein', 'March 14, 1879', 'Ulm', 'known', 'his theory', 'relativity', 'The United Nations', 'an inter
national organization', '1945', 'countries', 'Marie Curie', 'a pioneering physicist', 'She', 'the first woman', 'an
d', 'remains', 'a Nobel Prize', 'two different scientific fields', 'New York City', 'the most populous city', 'the
United States', 'its iconic skyline', 'Apple Inc.', 'a technology company', 'Steve Jobs', 'famous', 'for', 'product
s', 'the iPhone', 'MacBook', 'The Eiffel Tower', 'a wrought-iron lattice tower', 'Paris', 'one', 'the most recogniz
able landmarks', 'the world', 'Python', 'a high-level programming language', 'its readability', 'web development',
'The Mona Lisa', 'a famous portrait painting', 'Leonardo da Vinci', 'the Louvre Museum', 'NASA', '1958', 'responsib
le', "the United States' civilian space program", 'aeronautics and aerospace research', 'Mount Everest', "the worl
d's highest mountain", 'the Himalayas', 'the border', 'Nepal']
```