```
"""
Assignment No 5
Name – vaibhav shete
Batch – B3
Roll No - 53
Assignment Title: Implement regular expression function to find URL, IP address, Date, PAN number in textual data using python libraries
"""

import spacy
import re

# Load the spaCy English language model
nlp = spacy.load("en_core_web_sm")

# Define regular expressions
url_pattern = re.compile(r'https?://\S+|www\.\S+')

ip_address_pattern = re.compile(r'\b(?:\d{1,3}\.){3}\d{1,3}\b')

date_pattern = re.compile(r'\d{4}-\d{2}-\d{2}')
pan_number_pattern = re.compile(r'[A-Z]{5}[0-9]{4}[A-Z]')

def extract_entities(text):
    # Tokenize the text using spaCy
    doc = nlp(text)

    # Find entities using regular expressions
    urls = re.findall(url_pattern, text)
    ip_addresses = re.findall(ip_address_pattern, text)
    dates = re.findall(date_pattern, text)
    pan_numbers = re.findall(pan_number_pattern, text)

    # Extract spaCy entities
    entities = [(ent.text, ent.label_) for ent in doc.ents]

    return {
        'urls': urls,
        'ip_addresses': ip_addresses,
        'dates': dates,
        'pan_numbers': pan_numbers,
        'spaCy_entities': entities
    }

# Example usage
text_data = """
Here is a sample text with a URL: https://www.example.com.
Also, an IP address: 192.168.1.1. The date is 2023-01-01,
and a PAN number is ABCDE1234F.
"""

results = extract_entities(text_data)

print("URLs:", results['urls'])
print("IP Addresses:", results['ip_addresses'])
print("Dates:", results['dates'])
print("PAN Numbers:", results['pan_numbers'])
print("Entities:", results['spaCy_entities'])

# OUTPUT -
"""
URLs: ['https://www.example.com.']
IP Addresses: ['192.168.1.1']
Dates: ['2023-01-01']
PAN Numbers: ['ABCDE1234F']
Entities: [('IP', 'ORG'), ('192.168.1.1', 'CARDINAL'), ('2023-01-01', 'DATE'), ('PAN', 'ORG')]

"""
```