In [9]:
```python
import time
import json
from selenium import webdriver
from selenium.webdriver.common.by import By
from selenium.webdriver.chrome.service import Service
from selenium.webdriver.chrome.options import Options
from selenium.webdriver.common.action_chains import ActionChains
from selenium.webdriver.support.ui import WebDriverWait
from selenium.webdriver.support import expected_conditions as EC
```

In [31]:
```python
# Set up Chrome options
path_to_extension = "C:/Users/sandr/AppData/Local/Google/Chrome/User Data/Defaul

chrome_options = Options()
# chrome_options.add_argument("--headless")  # Ensure GUI is off
chrome_options.add_argument("--no-sandbox")
chrome_options.add_argument("--disable-dev-shm-usage")
chrome_options.add_argument("--disable-gpu")
chrome_options.add_argument(f'--load-extension={path_to_extension}')

# Initialize Chrome WebDriver
driver = webdriver.Chrome(options=chrome_options)

# Define the base URLs for each deck list page
urls = [
    "https://onepiecetopdecks.com/deck-list/english-format-op1-and-st1to4-meta-d
    "https://onepiecetopdecks.com/deck-list/en-format-op02-paramount-war-decklis
    "https://onepiecetopdecks.com/deck-list/en-format-op03-mighty-enemy-decklist
    "https://onepiecetopdecks.com/deck-list/en-format-op04-kingdom-of-intrigue-d
    "https://onepiecetopdecks.com/deck-list/en-format-op05-awakening-of-the-new-
    "https://onepiecetopdecks.com/deck-list/en-format-op-06-wings-of-the-captain
    "https://onepiecetopdecks.com/deck-list/english-eb-01-memorial-set-op-07-500
]

def dismiss_consent_modal():
    try:
        wait = WebDriverWait(driver, 5)  # Wait max 5 seconds for modal to appea
        do_not_consent_button = wait.until(
            EC.element_to_be_clickable((By.CSS_SELECTOR, "button.fc-button.fc-ct
        )
        do_not_consent_button.click()
        print("Dismissed consent modal.")
        time.sleep(2)  # Allow time for modal to disappear
    except Exception:
        print("No consent modal detected.")

def get_deck_links(driver):
    deck_links = []
    a_tags = driver.find_elements(By.TAG_NAME, 'a')
    for a_tag in a_tags:
        href = a_tag.get_attribute('href')
        if '/deck-list/' in href and 'deckgen' in href:
            deck_links.append(href)
    print(f"Found {len(deck_links)} deck links.")
    return deck_links

def parse_deck(deck_url):
    try:
```

```
        # driver.get(deck_url)
        # time.sleep(3)
        deck_data = deck_url.split('&dg=')[1].split('&cs=')[0]

        # Parse the deck data string
        deck_items = []
        for item in deck_data.split('a'):
            if item:
                count = int(item.split('n')[0])
                card_code = item.split('n')[1]
                deck_items.append({card_code: count})

        return deck_items
    except Exception as e:
        print(f"Error parsing deck from url: {deck_url}")
        return []

def scrape_decks(base_url):
    deck_links = []
    driver.get(base_url)

    time.sleep(3)
    dismiss_consent_modal()

    while True:
        time.sleep(3)
        deck_links.extend(get_deck_links(driver))

        try:
            next_button = driver.find_element(By.CSS_SELECTOR, "button.dt-paging
            driver.implicitly_wait(5)
            print(next_button.get_attribute('class'))
            if 'disabled' in next_button.get_attribute('class'):
                break
            print("Navigating to next page...")
            driver.execute_script("window.scrollTo(0, document.body.scrollHeight
            time.sleep(5) # lazy loading
            driver.execute_script("window.scrollTo(0, document.body.scrollHeight
            time.sleep(1) # lazy loading

            ActionChains(driver).move_to_element(next_button).click(next_button)
        except Exception as e:
            print(f"Error navigating to next page: {e}")
            break
        time.sleep(1)

    return deck_links
```

In [32]:

```
all_deck_lists = []

for url in urls:
    print(f"Scraping deck lists from {url}...")
    deck_links = scrape_decks(url)
    for deck_link in deck_links:
        # print(f"Scraping deck data from {deck_link}...")
        deck_data = parse_deck(deck_link)
        if deck_data:
            all_deck_lists.append(deck_data)
        # time.sleep(1)
```

```
print("Scraping complete.")
```

```
Scraping deck lists from https://onepiecetopdecks.com/deck-list/english-format-op
1-and-st1to4-meta-decks/...
Dismissed consent modal.
Found 100 deck links.
dt-paging-button next
Navigating to next page...
Found 70 deck links.
dt-paging-button disabled next
Scraping deck lists from https://onepiecetopdecks.com/deck-list/en-format-op02-pa
ramount-war-decklist/...
No consent modal detected.
Found 100 deck links.
dt-paging-button next
Navigating to next page...
Found 77 deck links.
dt-paging-button disabled next
Scraping deck lists from https://onepiecetopdecks.com/deck-list/en-format-op03-mi
ghty-enemy-decklist/...
No consent modal detected.
Found 100 deck links.
dt-paging-button next
Navigating to next page...
Found 100 deck links.
dt-paging-button next
Navigating to next page...
Found 14 deck links.
dt-paging-button disabled next
Scraping deck lists from https://onepiecetopdecks.com/deck-list/en-format-op04-ki
ngdom-of-intrigue-decklist/...
No consent modal detected.
Found 100 deck links.
dt-paging-button next
Navigating to next page...
Found 100 deck links.
dt-paging-button next
Navigating to next page...
Found 100 deck links.
dt-paging-button next
Navigating to next page...
Found 28 deck links.
dt-paging-button disabled next
Scraping deck lists from https://onepiecetopdecks.com/deck-list/en-format-op05-aw
akening-of-the-new-era/...
No consent modal detected.
Found 100 deck links.
dt-paging-button next
Navigating to next page...
Found 100 deck links.
dt-paging-button next
Navigating to next page...
Found 100 deck links.
dt-paging-button next
Navigating to next page...
Found 12 deck links.
dt-paging-button disabled next
Error parsing deck from url: https://onepiecetopdecks.com/deck-list/en-format-op0
5-awakening-of-the-new-era/deckgen?dn=Red%20Green%20Law&date=1/28/2024&cn=Oceania
&au=Jeremy&pl=17th%20Place&tn=Oceania%20Final&hs=TAK(512)&dg=1nOP01-002a4nOP01-00
6a4nOP01-016a4nST01-006a4nOP01-025a2nST01-012a3nOP01-013a3nST01-011a4nOP02-005a2n
OP02-015a%20nOP02-018a%20nOP04-002a3nOP05-010a4nST02-009a4nOP01-047a4nST02-007a3n
```

```
ST02-004a2nOP02-040&cs=243
Scraping deck lists from https://onepiecetopdecks.com/deck-list/en-format-op-06-w
ings-of-the-captain-decks/...
No consent modal detected.
Found 50 deck links.
dt-paging-button next
Navigating to next page...
Found 50 deck links.
dt-paging-button next
Navigating to next page...
Found 50 deck links.
dt-paging-button next
Navigating to next page...
Found 50 deck links.
dt-paging-button next
Navigating to next page...
Found 50 deck links.
dt-paging-button next
Navigating to next page...
Found 50 deck links.
dt-paging-button next
Navigating to next page...
Found 32 deck links.
dt-paging-button disabled next
Scraping deck lists from https://onepiecetopdecks.com/deck-list/english-eb-01-mem
orial-set-op-07-500-years-into-the-future-decks/...
No consent modal detected.
Found 50 deck links.
dt-paging-button next
Navigating to next page...
Found 50 deck links.
dt-paging-button next
Navigating to next page...
Found 50 deck links.
dt-paging-button next
Navigating to next page...
Found 50 deck links.
dt-paging-button next
Navigating to next page...
Found 50 deck links.
dt-paging-button next
Navigating to next page...
Found 50 deck links.
dt-paging-button disabled next
Scraping complete.
```

In [34]:
```python
# Save the collected deck lists to a JSON file
with open('data/one_piece_deck_lists.json', 'w') as f:
    json.dump(all_deck_lists, f, indent=2)
```