



# Ro-dou

O Robô minerador do Diário Oficial da União



## A Ferramenta

“Libertar as pessoas da tarefa manual de pesquisa no DOU.”

- Software Livre
  - Utiliza Airflow como motor
  - Configuração através de YAML
- 

# Utilizando YAMLS

```
alteracoes_estrutura_exemplo_evento_semana_inovacao.yaml M •
dags > ro_dou > dag_confs > alteracoes_estrutura_exemplo_evento_semana_inovacao.yaml
1 dag:
2   id: alteracoes_estrutura_exemplo_evento_semana_inovacao
3   description: Realiza busca no DOU por alterações na estrutura do SIORG.
4   search:
5     terms:
6       - estrutura regimental
7       - organização básica
8       - aprova o regimento interno
9       - quantitativo de cargos
10      - institui o comitê
11   dou_sections:
12     - SECAO_1
13   report:
14     emails:
15       - nitai.silva@economia.gov.br
16     subject: "[SIORG] Pesquisa por Mudanças Organizacionais"
17
```

# Resultado no e-mail

The screenshot shows the ProtonMail web interface. The top navigation bar includes the ProtonMail logo, a search bar, and links for 'Special offer', 'Upgrade', 'Contatos', and 'Configurações'. The user's name 'Nital Bezerra' and email 'nitalbezerra@protonmail.com' are visible in the top right. The left sidebar shows the 'Caixa de entrada' (Inbox) with 953 messages and other folders like 'Rascunhos', 'Enviados', 'Com estrela', 'Arquivo', 'Spam', and 'Lixeira'. The main content area displays search results for '[SIORG] Pesquisa por Mudanças Organizacionais - DOU de 2021-11-10'. The results are organized into sections, each starting with 'Seção 1' and a link to a document. The first section is titled '.Resultados para: estrutura regimental' and contains three entries: 1. 'PORTARIA GABAER Nº 171/GC1, DE 9 DE NOVEMBRO DE 2021' regarding the Aeronáutica command. 2. 'PORTARIA DECEA Nº 149/DGCEA, DE 8 DE NOVEMBRO DE 2021' regarding the DEPARTAMENTO DE CONTROLE DO ESPAÇO AÉREO. 3. 'RESOLUÇÃO Nº 938, DE 5 DE OUTUBRO DE 2021' regarding the IN CRA. The second section is titled '.Resultados para: aprova o regimento interno' and contains one entry: 'PAUTA DE JULGAMENTOS' regarding the Regimento Interno do CARF. The bottom status bar shows '32431 MB / 500.00 MB' and '4.0.10'.

ProtonMail

Pesquisar mensagens

Special offer Upgrade Contatos Configurações Nital Bezerra nitalbezerra@protonmail.com NB

Nova mensagem

Caixa de entrada 953

Rascunhos

Enviados

Com estrela

Arquivo

Spam

Lixeira

Todos os e-mails

PASTAS

[SIORG] Pesquisa por Mudanças Organizacionais - DOU de 2021-11-10

.Resultados para: estrutura regimental

Seção 1

[PORTARIA GABAER Nº 171/GC1, DE 9 DE NOVEMBRO DE 2021](#)

COMANDANTE DA AERONÁUTICA, no uso das atribuições que lhe conferem os incisos I e XIV do art. 23 da **Estrutura ... Regimental** do Comando da Aeronáutica, aprovada pelo Decreto nº 6.834, de 30 de abril de 2009, e considerando âmbito do Comando da Aeronáutica, a aplicação do inciso XIV do ar...

10/11/2021

Seção 1

[PORTARIA DECEA Nº 149/DGCEA, DE 8 DE NOVEMBRO DE 2021](#)

DEPARTAMENTO DE CONTROLE DO ESPAÇO AÉREO, de conformidade com o previsto no artigo 19, inciso I, da **Estrutura ... Regimental** do Comando da Aeronáutica, aprovada pelo Decreto nº 6.834, de 30 de abril de 2009, alteradoPORTARIA DECEA Nº 149/DGCEA, DE 8 DE NOVEMBRO DE 2021 Altera os itens I,...

10/11/2021

Seção 1

[RESOLUÇÃO Nº 938, DE 5 DE OUTUBRO DE 2021](#)

outubro de 1984, por seu Presidente, no uso das atribuições que lhe são conferidas pelo art. 17 da **Estrutura ... Regimental** do IN CRA, aprovada pelo Decreto nº 10.252, de 20 de fevereiro de 2020, concomitante o art. 17 da Estrutura Regimental do IN CRA, aprovada pelo Decreto nº 10.252, de ...

10/11/2021

Seção 1

[DECRETO Nº 10.853, DE 9 DE NOVEMBRO DE 2021](#)

Nº 10.853, DE 9 DE NOVEMBRO DE 2021 Altera o Decreto nº 5.417, de 13 de abril de 2005, que aprova a **Estrutura ... Regimental** e o Quadro Demonstrativo dos Cargos em Comissão e das Funções Gratificadas do Comando das Funções Gratificadas do Comando da Marinha do Ministério da Defesa. O PRE...

10/11/2021

.Resultados para: aprova o regimento interno

Seção 1

[PAUTA DE JULGAMENTOS](#)

o **Regimento Interno do CARF. ... o Regimento Interno do CARF.**

10/11/2021

Seção 1

[PAUTA DE JULGAMENTOS](#)

32431 MB / 500.00 MB 4.0.10

# Parâmetros de configuração

```
all_parameters_example.yaml M x
dag_confs > all_parameters_example.yaml
1 dag:
2   id: all_parameters_example
3   description: DAG exemplo utilizando todos os demais parâmetros.
4   tags:
5     - projeto_a
6     - departamento_x
7   schedule_interval: 0 8 * * MON-FRI # sintaxe CRON
8   search:
9     terms:
10      - dados abertos
11      - governo aberto
12      - lei de acesso à informação
13      - cibernética
14      - hacker ativismo
15     field: TUDO
16     is_exact_search: On
17     date: MES
18     dou_sections:
19       - SECAO_1
20       - EDICAO_SUPLEMENTAR
21     report:
22     emails:
23       - dest1@economia.gov.br
24       - dest2@economia.gov.br
25     attach_csv: True
26     subject: Assunto do Email
27
```

# Parâmetros extras

## Diário Oficial da União

PESQUISA

Informe o termo que deseja pesquisar nas Edições do Diário Oficial  | PESQUISA AVANÇADA

|  |   |   |   |
|--|---|---|---|
| <b>TIPO DE PESQUISA</b>                              | <b>ONDE PESQUISAR</b>                     | <b>DATA</b> <span>Novo!</span>                    | <b>JORNAL</b>                               |
| <input checked="" type="radio"/> Qualquer resultado  | <input checked="" type="radio"/> Tudo     | <input checked="" type="radio"/> Qualquer período | <input checked="" type="checkbox"/> Todos   |
| <input type="radio"/> Resultado exato                | <input type="radio"/> No título           | <input type="radio"/> Edição do Dia               | <input type="checkbox"/> Seção 1            |
|  | <input type="radio"/> No conteúdo         | <input type="radio"/> Última semana               | <input type="checkbox"/> Seção 2            |
| <b>FORMA DE PESQUISA</b> *                           | <b>ORDENAÇÃO</b> <span>Novo!</span>       | <input type="radio"/> Último mês                  | <input type="checkbox"/> Seção 3            |
| <input checked="" type="radio"/> Pesquisa Ato-a-Ato  | <input checked="" type="radio"/> Por data | <input type="radio"/> Último ano                  | <input type="checkbox"/> Edição Extra       |
| <input type="radio"/> Pesquisa na Versão Certificada | <input type="radio"/> Por relevância      | <input type="radio"/> Personalizado               | <input type="checkbox"/> Edição Suplementar |
| <input type="radio"/> Diário Completo Certificado    |   |   |   |

**PESQUISAR**

\* Para consultas anteriores a 2018, selecione uma das opções **Pesquisa na Versão Certificada** ou **Diário Completo Certificado**.

# Lista dinâmica de termos

```
movimentacao_servidores.yaml M x
dags > ro_dou > dag_confs > movimentacao_servidores.yaml
1 dag:
2   id: busca_movimentacao_de_servidores
3   description: Realiza busca no DOU das movimentações de servidores das carreiras da SEGES.
4   schedule_interval: '30 5 * * *'
5   search:
6     terms:
7       from_db_select:
8         sql: "SELECT servidor_nome_contato, carreira FROM sgc.vw_servidor_contato_ativos_2;"
9         conn_id: conn_id_bd
10    dou_sections:
11      - SECAO_2
12      - EDICAO_EXTRA
13    force_rematch: On
14    ignore_signature_match: On
15    report:
16    emails:
17      - pessoas.felizes@economia.gov.br
18    attach_csv: On
19    subject: "[Airflow-bot] Movimentações de Servidores"
20
```

# Ro-dou v0.1

- [github.com/economiagovbr/Ro-dou/](https://github.com/economiagovbr/Ro-dou/)
- Ambiente de Exemplo em Docker!
- Em produção:
  - Servidor com Docker
  - Serviço SMTP

The screenshot displays the GitHub repository page for `economiagovbr/Ro-dou`. The repository is public and has 178 commits. The file browser shows the following structure:

| File                            | Description  | Last Commit  |
|---------------------------------|--|--------------|
| <code>.github/workflows</code>  | <code>fix(ci-tests)</code> : accesses directory to run makefile on every step    | 2 days ago   |
| <code>dag_confs</code>          | <code>chore(yaml examples)</code> : sets name without abbreviation               | 2 days ago   |
| <code>docs/img</code>           | <code>docs(readme.md)</code> : adds YAML examples and a better introduction      | 2 months ago |
| <code>src</code>                | <code>feat(parsers.py)</code> : enables usage of multiline strings as terms list | 2 days ago   |
| <code>tests</code>              | <code>chore(yaml examples)</code> : sets name without abbreviation               | 2 days ago   |
| <code>.gitignore</code>         | <code>ci(gitignore)</code> : ignores tracking pg container volume data           | 2 months ago |
| <code>LICENSE</code>            | Create LICENSE   | 3 months ago |
| <code>Makefile</code>           | <code>fix(makefile)</code> : changes git clone protocol                          | 2 days ago   |
| <code>README.md</code>          | <code>docs(readme)</code> : clarify required Docker Compose version              | yesterday    |
| <code>docker-compose.yml</code> | <code>docs(readme)</code> : adds "Configurando em Produção" section              | 2 days ago   |

The README content is as follows:

## Ro-dou

CI Tests: passing

O Ro-dou é uma ferramenta para gerar dinamicamente DAGs no [Apache Airflow](#) que fazem clipping do Diário Oficial da União (DOU) a partir de arquivos **YAML**. Receba no email todas as publicações que contenham as **palavras chaves** que você definir.

## Recursos

- Frequência **diária**, **semanal** ou **mensal**
- Relatório em `.csv`
- Busca dinâmica das palavras chaves no **DB** ou de uma **Variável**
- Pesquisa em **Seção** específica

## Ambiente de Exemplo



## Potenciais evoluções

- Integração com o “Querido Diário”
  - Desenvolver uma interface web mais simples para edição
  - Disponibilização de um Serviço público
- 

# Obrigado!

nitaibezerra@protonmail.com



@nitaibezerra



@nitaib

CREDITS: This presentation template was created by Slidesgo, including icons by Flaticon, and infographics & images by Freepik.

CREDITS: Illustrations from Storyset