

Capitulo_3

Flavio Codeço Coelho
(Dated: September 12, 2019)

CONTENTS

I. Identificando Entidades Nomeadas	1
I.1. Testando o NER do Spacy	2
I.2. Treinando Um identificador de Entidades a partir do DHBB	3

I. IDENTIFICANDO ENTIDADES NOMEADAS

Neste capítulo vamos treinar um identificador de Entidades (NER) nomeadas usando a biblioteca [Spacy](#). A partir deste capítulo vamos importar também funções que já criamos anteriormente, e que encontram-se reproduzidas em [dhbbmining.py](#)

```
In [38]: import os, glob
import spacy
from spacy import displacy
from sqlalchemy import create_engine
from dhbbmining import *
import ipywidgets as widgets
```

Para utilizar o spacy em um corpus na lingua portuguesa, vamos primeiro importar o modelo linguístico do português

```
In [19]: nlp = spacy.load("pt_core_news_sm")
```

Em seguida podemos carregar os verbetes biográficos que salvamos no nosso banco SQLite.

```
In [17]: eng = create_engine("sqlite:///minha_tabela.sqlite")
#dhbb = pd.read_sql_table('resultados')
biograficos = pd.read_sql_query('select * from resultados where_
↳natureza="biográfico"',
con=eng)
biograficos.head()
```

```
[17]:
```

	index	arquivo	title	natureza	sexo	\
0	0	1.text	COELHO, Machado	biográfico	m	
1	1	10.text	ABÍLIO, Armando	biográfico	m	
2	2	100.text	ALEIXO, Pedro	biográfico	m	
3	3	1000.text	CAMPOS, Eduardo	biográfico	m	
4	4	1001.text	CAMPOS, Eleazar Soares	biográfico	m	

```

cargos \
0 \n - dep. fed. DF 1927-1929 \n - dep. fed. DF ...
1 \n - dep. fed. PB 1995-1999\n - dep. fed. PB ...
2 \n - const. 1934\n - dep. fed. MG 1935-1937\n ...
3 \n - dep. fed. PE 1995\n - dep. fed. PE 1998-...
4 \n - magistrado\n - interv. MA 1945-1946\n
```

corpo

```

0 \n\nJosé Machado Coelho de Castroz nasceu em ...
1 \n\nArmando Abílio Vieiraz nasceu em Itaporan...
2 \n\nPedro Aleixo nasceu em São Caetano, dist...
3 \n\nEduardo Henrique Accioly Campos nasceu e...
4 \n\nEleazar Soares Campos nasceu em São Luís...

```

Para começar a utilizar o Spacy, precisamos primeiro processar o texto. Nesta passagem várias análises linguísticas são realizadas.

```

In [35]: doc = nlp(biograficos.corpo[0].strip())
         type(doc)

```

```

[35]: spacy.tokens.doc.Doc

```

```

In [36]: for i, token in enumerate(doc):
         print(token.text, token.lemma_, token.pos_, token.tag_, token.dep_,
               token.shape_, token.is_alpha, token.is_stop)
         if i>5:
             break

```

```

ñ ñ PUNCT PU|@PU punct ñ False False
José José PROPN PROPN nsubj Xxxx True False
Machado Machado PROPN PROPN flat:name Xxxxx True False
Coelho Coelho PROPN PROPN flat:name Xxxxx True False
de de ADP PRP|@N< case xx True True
Castro Castro PROPN PROP|@P< nmod Xxxxx True False
z z PUNCT PU|@PU punct z False False

```

I.1. Testando o NER do Spacy

Como o Spacy já inclui algum suporte à língua portuguesa, antes de pensar em treinar nosso próprio NER, podemos avaliar a performance do existente.

Abaixo vamos construir uma visualização interativa da marcação de entidades nos verbetes do DHBB.

```

In [63]: from IPython.display import display,clear_output
         from ipywidgets import interact

```

```

In [73]: @interact(e=(0, len(biograficos)))
         def mostra_ner(e=0):
             text = biograficos.iloc[e].corpo.strip()
             doc = nlp(text)
             displacy.render(doc, style="ent", jupyter=True)
             clear_output(wait=True)

```

```

interactive(children=(IntSlider(value=0, description='e', max=6724), Output()), _dom_classes=('widget-in

```

Além da visualização, podemos extrair as entidades presentes em um verbete:

```

In [74]: for ent in doc.ents:
         print(ent.text, ent.start_char, ent.end_char, ent.label_)

```

```

José Machado Coelho de Castroz 1 31 PER
Lorena 42 48 LOC
SP 50 52 LOC

```

Ginásio Diocesano de São Paulo 67 97 LOC
 Faculdade de Ciências Jurídicas 127 158 ORG
 Sociais 161 168 LOC
 Cunha 220 225 LOC
 SP 227 229 LOC
 Rio de
 Janeiro 263 277 LOC
 Distrito Federal 285 301 LOC
 Distrito Federal 357 373 LOC
 Câmara 488 494 LOC
 Getúlio Vargas 575 589 PER
 Ligado 591 597 LOC
 Washington Luís 654 669 PER
 Guanabara 682 691 LOC
 Alcântara 951 960 LOC
 Lisboa 965 971 LOC
 Getúlio 1050 1057 PER
 Vargas 1058 1064 PER
 São 1125 1128 LOC
 Paulo 1129 1134 LOC
 São 1206 1209 LOC
 Paulo 1210 1215 LOC
 Partido Social Democrático 1231 1257 ORG
 PSD 1259 1262 ORG
 Assembléia Nacional Constituinte 1284 1316 ORG
 ANC 1318 1321 ORG
 Carta
 1484 1490 PER
 Constituinte 1523 1535 MISC
 Congresso 1539 1548 LOC
 Comissão Permanente de Obras Públicas da Câmara Federal 1571 1626 ORG
 Câmara 1732 1738 ORG
 Companhia de Cimento Vale do Paraíba 1784 1820 ORG
 Rio de Janeiro 1834 1848 LOC

I.2. Treinando Um identificador de Entidades a partir do DHBB

Identificadores de entidades são algoritmos treinados em corpora manualmente anotados. Como cada corpora possui um conjunto particular de entidades, para uma performance ótima o ideal é treinarmos o modelo no Próprio DHBB. Para este fim utilizaremos os dicionários já disponíveis no DHBB, juntamente com o índice construído no capítulo 2 para recuperar o contexto de cada entrada dos dicionários.

```
In [2]: from whoosh import index
        import os
        from whoosh.qparser import QueryParser
        from whoosh import qparser
```

O primeiro passo é abrirmos o nosso índice.

```
In [3]: if os.path.exists('indexdir'):
        indice = index.open_dir('indexdir')
```

```
In [5]: indice.doc_count
```

```
[5]: 6724
```

```
In [10]: def busca(consulta):
        qp = QueryParser("corpo", indice.schema)
        qp.add_plugin(qparser.EveryPlugin())
        query = qp.parse(consulta)

        with indice.searcher() as searcher:
            results = [(dict(hit),hit.highlights("corpo")) for hit in searcher.
→search(query,
            limit=None)]
        return results
```

```
In [12]: resultados = busca('"filho de")[0]
```

```
[12]: ({'caminho': '3822.text',
        'corpo': '\n\nPaulo Nogueira Filho nasceu na cidade de São Paulo no dia 16
de\nnovembro de 1898, filho de Paulo de Almeida Nogueira e de Ester\nNogueira.
Seu avô, José Paulino Nogueira, foi coronel da Guarda\nNacional, fazendeiro,
comerciante e iniciador de grandes empreendimentos\nfinanceiros.\n\nDescendente
de famílias abastadas, proprietárias de fazendas e\nindústrias, Paulo Nogueira
Filho conviveu desde a infância com\ndestacadas personalidades do mundo
econômico e político da época, como\nManuel de Campos Sales, José Gomes Pinheiro
Machado, Carlos Guimarães,\nJosé Maria Whitaker e outros. Recebeu educação
cuidadosa, freqüentando\ndurante dois anos uma escola na Europa. Concluiu os
estudos secundários\nno Colégio São Bento, na capital paulista, e, depois de
estudar durante\nsete meses no curso preparatório Henrique Greenen, ingressou em
1914 na\nFaculdade de Direito de São Paulo. Desde o início do curso,
desenvolveu\nintensa atividade política, iniciada na campanha pelo serviço
militar\nobrigatório e o voto secreto, sob a liderança de Olavo Bilac e
Miguel\nPereira. Em decorrência de sua participação nesse movimento, ainda
como\nprimeiranista ingressou no Voluntariado de Manobras organizado
pelo\nExército e, em seguida, integrou o Batalhão Acadêmico formado na
sua\nfaculdade. Em 1916, criou junto com um grupo de colegas o grêmio\nliterário
Álvares de Azevedo que, mobilizando recursos privados,\nconstruiu uma escola
noturna dirigida por estudantes e destinada à\nalfabetização de adultos.\n\nO
torpedeamento do navio brasileiro Paraná por um submarino alemão em 7\nde
abril de 1917, durante a Primeira Guerra Mundial, gerou intensa onda\nde
protestos de cunho nacionalista. Enquanto Rui Barbosa, no Rio, e\nJúlio de
Mesquita, em São Paulo, escreviam veementes artigos de\nprotesto, estudantes
organizaram comícios e, no dia 12, uma multidão\nempastelou o Diário Alemão.
Nesse contexto, foi fundada em São Paulo a\nLiga Nacionalista, seguidora das
idéias de Olavo Bilac e presidida pelo\nprofessor Frederico Steidel. Paulo
Nogueira Filho foi um dos primeiros\nfilhos dessa organização. Ainda
estudante, apoiou a candidatura de\nLuís Pereira Barreto para o Senado estadual,
vitoriosa depois de uma\ncampanha que realizou inúmeras caravanas pelo interior
e desafiou o\ndomínio político exercido pelo Partido Republicano Paulista (PRP)
no\nestado.\n\nPor ocasião da epidemia de gripe de 1918, o governo facultou a
aprovação\nde estudantes por decreto, mas a Liga Nacionalista exigiu que
seus\nassociados prestassem os exames finais. Descontentes, Paulo
Nogueira\nFilho e outros estudantes lançaram, em 28 de janeiro de 1919, o
chamado\nManifesto dos 60, formalizando o rompimento com a Liga e
criticando\nduramente seu presidente. Nos meses seguintes, Paulo Nogueira
apoiou\nativamente a campanha de Rui Barbosa para a presidência da
República,\ntornada vaga com a morte de Francisco de Paula Rodrigues Alves.
Apesar\nda vitória de Epitácio Pessoa no pleito realizado em 13 de abril
de\n1919, a participação na campanha eleitoral estimulou o liberalismo de\nPaulo
Nogueira Filho, bem como sua admiração pelo candidato derrotado.\n\nEm 1921,
```

casou-se com Regina Sales de Oliveira Coutinho, descendente de uma família paulista tradicional. No ano seguinte, decidiu dedicar-se aos negócios, reorganizou a Companhia de Tecidos de Seda Santa Branca, onde sua propriedade, elaborando um regimento interno para a empresa baseado nos códigos de trabalho vigentes nos países de capitalismo avançado.

A formação do Partido Democrático

Paulo Nogueira Filho viajou à Europa em 1924 para tomar conhecimento de técnicas mais avançadas de tecelagem. De volta ao Brasil, simpatizou com o movimento revolucionário deflagrado em 5 de julho de 1924 na capital paulista sob o comando do general Isidoro Dias Lopes, mas não chegou a aderir. Mesmo assim, intensificou sua atividade política sob a influência dos movimentos que combatiam a oligarquia dominante em São Paulo, vinculada ao PRP. No mês seguinte ao seu retorno, sob o estado de sítio que vigorou durante todo o governo de Artur Bernardes, convocou uma reunião em sua casa para organizar uma sociedade secreta que teria como objetivo criar um novo partido político e um jornal. Surgiu assim, em 31 de agosto de 1924, a Sociedade Invisível de Ação Política (SIAP), cujo termo de fundação foi assinado por Joaquim de Abreu Sampaio Vidal, José Beneton Prado, Paulo Vicente de Azevedo, Valdemar Ferreira, José Maria de Camargo Aranha, Tácito de Almeida, Tomás Lessa, Rubens Borba Alves de Moraes, Mário de Andrade, Francisco Malta Cardoso, Clóvis Ribeiro, Antônio Vicente de Azevedo e pelo próprio Paulo Nogueira Filho.

O programa dessa organização defendia a eleição do presidente da República pelo Congresso Nacional, o controle parlamentar sobre as atividades dos ministros, a limitação do estado de sítio aos casos de guerra, a realização de uma reforma tributária e a adoção de leis reguladoras das relações entre operários e patrões.

Em virtude do estado de sítio e da vigilância policial, a SIAP se limitou a reuniões esporádicas, mas tornou-se uma das principais vertentes na articulação de um novo partido paulista. Em 1925, Paulo Nogueira Filho, o conselheiro Antônio da Silva Prado e outros buscaram aglutinar a crescente oposição ao regime vigente, entrando em contato com o deputado José Adriano Marrey Júnior - que proclamava na Câmara Estadual a necessidade de formar um partido popular - e Valdemar Ferreira - que tentava criar um partido liberal. Da aglutinação dessas correntes surgiu o Partido Democrático (PD) de São Paulo, cuja comissão organizadora se reuniu pela primeira vez na casa de Antônio Prado em 24 de fevereiro de 1926 e elegeu Paulo Nogueira Filho para a secretaria geral. A nova agremiação defendia as instituições da democracia liberal, lutando especialmente pela reforma eleitoral e o voto secreto, mas era deliberadamente vaga na definição das principais reivindicações trabalhistas, defendendo todas as medidas que interessam à questão social.

Durante o primeiro congresso do novo partido, realizado em 21 de março de 1927, Paulo Nogueira Filho defendeu, com êxito, a criação de um jornal vinculado à causa democrática e propôs um estatuto que incluía a adoção do voto secreto para a escolha dos dirigentes partidários. Apesar de contar em sua direção com uma clara maioria de fazendeiros e profissionais de alta extração social, não se diferenciando nesse aspecto das correntes tradicionais, o PD encontrou excepcional receptividade em São Paulo, conseguindo cerca de 20 mil filiações em menos de três meses. Em 13 de junho de 1927 foi assinado o manifesto de lançamento da Sociedade Anônima Diário Nacional, empresa dirigida por Paulo Nogueira Filho e Marrey Júnior e responsável pelo jornal do partido cujo primeiro número foi publicado em 14 de julho seguinte.

Pouco depois, Paulo Nogueira e Francisco Morato viajaram ao Rio de Janeiro, então Distrito Federal, para discutir com o líder gaúcho Joaquim Francisco de Assis Brasil a criação de um partido de âmbito nacional dentro da linha programática do PD. Os entendimentos prosseguiram até a formação, em 21 de setembro de 1927, do Partido Democrático Nacional, cujo diretório provisório ficou composto por Assis Brasil, Paulo de Moraes Barros, Adolfo Bergamini, Francisco

Morato, Marrey Júnior, Plínio Casado e João Batista Luzardo, cabendo a Antônio da Silva Prado o cargo de presidente de honra. Paulo Nogueira Filho participou intensamente da organização e desenvolvimento da campanha do PD para as eleições de 24 de fevereiro de 1928, quando seria renovada a Câmara Estadual paulista. A expectativa despertada pelos sucessivos comícios e a intensa mobilização da população durante a campanha foram frustradas pela utilização da fraude em larga escala por parte das correntes dominantes, desfechando um duro golpe nas esperanças de obter mudanças institucionais significativas através das urnas. Em consequência, alguns líderes do partido - entre os quais Paulo Nogueira Filho - tenderam a se aproximar de setores da jovem oficialidade das forças armadas que defendiam a luta armada contra o regime. Quatro dias depois das eleições em São Paulo, Paulo Nogueira Filho, Valdemar Ferreira, Berto Condé e Paulo Pinto de Carvalho integraram a delegação do PD ao Congresso de Bajé, onde três correntes da oposição gaúcha - a federalista, a democrática e a castilhistas dissidente - se uniram, formando o Partido Libertador (PL), que adotou na íntegra o programa do Partido Democrático Nacional e aderiu à idéia revolucionária. Assis Brasil, principal líder do PL, incumbiu Paulo Nogueira Filho de realizar uma série de contatos no Rio Grande do Sul antes de retornar a São Paulo. A primeira viagem foi à fronteira com a Argentina para encontrar os chefes revolucionários Miguel Costa e Isidoro Dias Lopes, exilados nesse país. Paulo Nogueira retornou em seguida a Porto Alegre, onde conversou com Raul Pilla, foi depois a Pelotas contatar Anacleto Firpo e rumou, finalmente para a cidade de Mello, no Uruguai, a fim de rever Assis Brasil. Ainda por essa época, Paulo Nogueira Filho viajou a Buenos Aires para conversar com Luís Carlos Prestes sobre a preparação de um movimento armado. Segundo seu depoimento posterior, Prestes era cético quanto à possibilidade de que o PD e o PL pudessem atingir sozinhos seus objetivos programáticos, mas considerava útil a ação propagandista que poderiam desenvolver. Ao mesmo tempo, mostrou-se disposto a participar do movimento revolucionário, desde que fosse possível prepará-lo de maneira adequada, levando Paulo Nogueira Filho a comprometer-se com a arrecadação dos fundos financeiros necessários. De volta da viagem, Paulo Nogueira Filho publicou artigo no Diário de São Paulo de 19 de abril de 1928 intitulado O pensamento político de Luís Carlos Prestes. Constatou, entretanto, que o clima político em São Paulo era avesso a qualquer movimento revolucionário, pois os deputados eleitos pelo PD preferiam a manutenção da ordem e a maioria do partido pretendia persistir na prática eleitoral. Por outro lado, os entendimentos com Assis Brasil, no Rio Grande do Sul, e Batista Luzardo e Maurício de Lacerda, no Rio, reforçavam a idéia de que não havia condições de preparar a deflagração da luta armada. Essa também era a opinião de Prestes, que escreveu carta a Nogueira Filho afirmando não ter recebido a quantia prometida, criticando a ação do PD e de Assis Brasil, e recomendando uma ênfase especial no desenvolvimento da propaganda antigoverno. O núcleo revolucionário civil mais próximo a Paulo Nogueira Filho se sentia de uma coordenação precária com outros grupos. Esses problemas, entretanto, não impediam o crescimento do proselitismo dos democráticos. Em 19 de julho de 1928, partiu de São Paulo em direção ao Norte uma caravana formada por Assis Brasil, Maurício de Lacerda, Simões Lopes Filho, Valdemar Ferreira, Nereu Ramos, Davidoff Lessa, Roberto Macedo e Vicente de Moraes, recebendo calorosa recepção em diversas cidades, especialmente em Recife, onde uma multidão calculada em 50 mil pessoas foi ouvir os oradores. Considerando impossível organizar nesse momento um movimento revolucionário, Paulo Nogueira Filho exonerou-se da secretaria geral do PD e da direção do Diário Nacional, embarcando para a Europa em novembro de 1928. ## A Revolução de 1930 De volta ao Brasil em 1929, Paulo Nogueira Filho já encontrou a oposição articulada em torno da Aliança Liberal, coligação de âmbito nacional

que lançara a candidatura de Getúlio Vargas para disputar o pleito presidencial de março de 1930 com o candidato situacionista Júlio Prestes, presidente de São Paulo e líder do PRP. Ainda na condição de membro do diretório central do PD, integrou-se à campanha aliancista e assumiu a responsabilidade de reorganizar o *Diário Nacional*, que estava ameaçado por uma crise financeira. Ao mesmo tempo, retomou seus contatos com os tenentes que atuavam em São Paulo sob a chefia de Antônio de Siqueira Campos na preparação de um levante de âmbito nacional. Paulo Nogueira Filho foi um dos principais organizadores da visita de Getúlio Vargas a São Paulo em 4 de janeiro de 1930, quando o candidato aliancista foi recebido e aclamado por uma imensa multidão. A campanha transcorreu em meio a grande tensão. A certa altura, ajudou a mobilizar cerca de duzentos homens para defender a sede do *Diário Nacional*, ameaçada de empastelamento por parte de membros do PRP. Depois, informado sobre a existência de um plano para assassiná-lo, dirigiu-se ao secretário de Segurança do governo estadual, Antônio Carlos Sales Júnior, para responsabilizar as autoridades por qualquer tipo de atentado ou violência contra sua pessoa. Durante toda a campanha eleitoral, Nogueira Filho continuou participando do grupo que, sob a chefia de Siqueira Campos, preparava a revolução em São Paulo. Com a vitória de Júlio Prestes nas eleições, os preparativos foram acelerados, com a distribuição para os conspiradores de granadas de mão fabricadas por Josias Carneiro Leão. Após a morte de Siqueira Campos em desastre de avião no mês de maio de 1930, Ricardo Holl assumiu o comando revolucionário em São Paulo e destacou elementos do PD para o interior do estado, cabendo a Paulo Nogueira Filho o setor de Campinas. As principais bases de apoio político-militar dos revolucionários estavam localizadas no Rio Grande do Sul, Minas Gerais e Paraíba, que seriam os estados iniciadores do levante. Com a definição das linhas gerais do plano revolucionário, o PD designou Paulo Nogueira Filho e Aureliano Leite como seus representantes nas colunas gaúcha e mineira, respectivamente. Paulo Duarte, Carlos Morais Andrade, Carlos Prado, Aníbal Fagundes e outros líderes democráticos permaneceram na capital paulista coordenando núcleos de revoltosos que entrariam em ação mais tarde. Chegando ao Rio Grande do Sul, Paulo Nogueira Filho foi informado de que a eclosão do movimento havia sido adiada e resolveu seguir para Buenos Aires a fim de encontrar Prestes, que se recusou novamente a participar do movimento. Na viagem de volta, constatou a disposição de Assis Brasil e dos libertadores, favorável à luta armada. Paulo Nogueira Filho só chegou a Porto Alegre em 4 de outubro de 1930, no dia seguinte à deflagração do levante que obteve êxito fulminante em todo o estado. Com a consolidação da vitória dos revolucionários gaúchos e a organização da marcha para o norte, foi convidado a se juntar a Ildefonso Simões Lopes, Francisco Antunes Maciel, Joaquim Maurício Cardoso, Fernando de Abreu Ferreira e Anacleto Firpo no estado-maior civil de Getúlio Vargas, chefe supremo do movimento. Esses líderes partiram no dia 12 com o objetivo de alcançar no Paraná os destacamentos de vanguarda comandados por Miguel Costa, Alcides Etchegoyen e João Alberto Lins de Barros. No início da viagem, Nogueira Filho conversou com Vargas sobre a situação em São Paulo, ficando acertado que Francisco Morato, presidente do PD, assumiria o governo paulista depois da vitória. No decorrer da campanha militar, integrou, junto com Antunes Maciel, o primeiro escalão do destacamento comandado por Batista Luzardo, que chegou a Ponta Grossa (PR) em 18 de outubro. Quatro dias depois redigiu uma mensagem conclamando os paulistas a colaborarem com a revolução. Pressionada pelo avanço das tropas revolucionárias em todas as frentes, na alta hierarquia militar sediada no Rio de Janeiro o presidente Washington Luís no dia 24 de outubro, empossando uma junta que propôs a cessação dos combates. Os revolucionários exigiram então a rendição das tropas legalistas estacionadas na fronteira de São Paulo com o

Paraná, ao mesmo tempo em que o general Hastínfilo de Moura assumia interinamente a chefia do governo paulista em lugar de Júlio Prestes. No dia 3 de novembro, Getúlio Vargas foi empossado à frente do Governo Provisório. A reorganização política de São Paulo depois da vitória da revolução foi marcada por sucessivas crises resultantes do confronto entre os partidos tradicionais no estado e as correntes tenentistas apoiadas pelo novo governo federal. Paulo Nogueira Filho e seus companheiros esperavam a nomeação de Francisco Morato para a interventoria paulista e se opunham vigorosamente à entrega desse cargo ao tenente João Alberto. Durante o chamado governo dos 40 dias (que existiu em São Paulo entre 26 de outubro e 5 de dezembro de 1930, com o apoio do PD, sob a chefia de José Maria Whitaker, Plínio Barreto e João Alberto, sucessivamente) Nogueira Filho exerceu o cargo de subchefe de polícia, subordinado a Vicente Rao, a quem substituiu em diversas ocasiões. Numa delas, recebeu ordem de João Alberto para permitir a instalação de uma sede do Partido Comunista Brasileiro - então Partido Comunista do Brasil (PCB) - e garantir sua liberdade de atuação. Junto com Elias Machado, Joaquim Celidônio e os maiores Mário Barbosa de Oliveira e Heitor Lobato Vale, integrou também a Comissão Central de Sindicâncias, encarregada de coordenar o trabalho das dez juntas que atuavam no estado. Afastou-se nesse período, junto com José Marrey Júnior, da direção efetiva do Diário Nacional, continuando, mesmo assim, a figurar na sua diretoria. Paulo Nogueira Filho foi convidado para participar da direção da Legião Revolucionária, organização tenentista fundada em 12 de novembro de 1930 sob a liderança de Miguel Costa e João Alberto com o objetivo de apoiar o aprofundamento das reformas iniciadas com a vitória da revolução. Sua oposição às pretensões tenentistas em relação à interventoria paulista levaram-no porém a declinar do convite. Em 25 de novembro, João Alberto foi nomeado interventor, contrariando profundamente os desejos do PD.

Na Revolução de 1932

No início de 1931, Paulo Nogueira Filho foi credenciado como delegado do PD junto ao Governo Provisório e, em fevereiro desse ano, recebeu a incumbência de organizar a Comissão Central de Compras do governo federal. Nesse período, a crise política em São Paulo se agravou, levando o PD a divulgar no dia 5 de abril um manifesto rompendo com João Alberto, acusado de preterir paulistas ilustres em benefício de pessoas estranhas ao estado, manter no governo elementos inconvenientes à tranquilidade geral e entregar diversas prefeituras a militares de sua confiança. Paulo Nogueira Filho foi signatário desse documento e participou ativamente da articulação da Frente Única Paulista (FUP), coligação do PD com o PRP lançada oficialmente em 16 de fevereiro de 1932 em torno de um programa que defendia a imediata reconstitucionalização do país e a restauração da autonomia estadual. A FUP passou a discutir a deflagração de levante armado, cujos preparativos foram intensificados depois das grandes manifestações populares hostis à presença de Osvaldo Aranha, ministro da Fazenda, em São Paulo. Nesse contexto, Paulo Nogueira Filho promoveu a formação da Guarda Paulista, mais tarde transformada no MMDC, organização que se destacou na preparação e no apoio à luta armada. Pouco depois, foi substituído por Aureliano Leite na direção da organização. Paulo Nogueira Filho cooperou diretamente com a chefia do movimento revolucionário. No dia 8 de julho, quando a tensão chegava ao ápice, viajou junto com João Neves da Fontoura para o Rio de Janeiro a fim de realizar uma última tentativa de evitar a luta armada. Não chegou a cumprir essa missão em virtude da deflagração da Revolução Constitucionalista no dia seguinte. No Rio, embarcou clandestinamente com um grupo de oficiais do Exército em uma frágil embarcação pesqueira que rumou para o litoral paulista, e aí chegando se integrou às forças revolucionárias estacionadas no sul do estado sob o comando do coronel Basílio Taborda. O desenvolvimento da guerra civil foi desfavorável aos constitucionalistas que, isolados do resto do país e incapazes de avançar

para além das fronteiras de seu estado, assumiram uma posição fundamentalmente defensiva durante o conflito. Apesar de todo o esforço de guerra, marcado pela conversão de fábricas à produção bélica e pela intensa mobilização popular, as dificuldades dos revoltosos foram se avolumando diante de um inimigo que em todas as frentes se mostrava muito mais numeroso e melhor provido tecnicamente. Nesse contexto, Nogueira Filho recebeu do governo paulista a importante missão de chefiar uma delegação à Argentina, onde conseguiu levantar mil libras esterlinas e 60 mil dólares correspondentes ao futuro envio de 250 mil sacas de café.

Com a derrota da revolução no início de outubro de 1932, Paulo Nogueira Filho partiu para o exílio, de onde retornou no ano seguinte. Em outubro de 1934, foi eleito deputado federal por São Paulo na legenda do Partido Constitucionalista, passando a apoiar a candidatura afinal vitoriosa de Armando de Sales Oliveira ao governo estadual. Em maio de 1935 iniciou seu mandato na Câmara dos Deputados. Favorável à rigorosa Lei de Segurança Nacional, apoiou em julho o fechamento da Aliança Nacional Libertadora (ANL), frente política contra o fascismo, o imperialismo e o latifúndio.

Em 10 de novembro de 1937, o presidente Getúlio Vargas liderou um golpe militar que implantou o Estado Novo, fechando todos os órgãos legislativos do país. Contrário ao novo regime, Paulo Nogueira Filho foi preso e deportado junto com Armando de Sales Oliveira e outros líderes políticos. Seu novo período de exílio durou cerca de sete anos, durante os quais continuou combatendo o Estado Novo. Junto com outros exilados negressos do Partido Constitucionalista, assinou o manifesto enviado em 7 de janeiro de 1939 ao presidente dos Estados Unidos, Franklin Roosevelt, contra a ditadura de Vargas, e participou de organizações internacionais como o Mundo Livre e a União Republicana Latina, da qual foi vice-presidente.

Logo depois da derrubada do Estado Novo em 29 de outubro de 1945, o Supremo Tribunal Federal (STF) concedeu habeas-corpus impetrado por quinhentos advogados para que Armando de Sales Oliveira, Otávio Mangabeira e Paulo Nogueira Filho pudessem retornar ao país. De volta ao Brasil, Paulo Nogueira ingressou na União Democrática Nacional (UDN) passando a fazer parte do secretariado nacional dessa agremiação, presidido por Virgílio de Melo Franco. Dirigiu no período seguinte a campanha financeira do partido e criou o seu departamento de ação social.

No pleito de 2 de dezembro de 1945, obteve a maior votação entre os candidatos udenistas que concorreram em São Paulo à Assembleia Nacional Constituinte, reunida a partir de 5 de fevereiro de 1946. Teve destacada atuação no decorrer dos trabalhos de elaboração da nova Carta, apresentando, junto com seu colega Jurandir de Castro Pires Ferreira, 155 emendas ao projeto original elaborado pela Comissão Constitucional. Com a promulgação da Constituição (18/9/1946), passou a exercer o mandato ordinário, ao mesmo tempo em que exerceu interinamente a presidência da agremiação durante a licença de Otávio Mangabeira, que concorria ao governo da Bahia. Nesse período publicou também diversos trabalhos relativos à sua atividade parlamentar como: Discursos sobre a organização do estado (1946), Emendas ao projeto de Constituição (1946), Em defesa da autonomia de São Paulo (1947), Em defesa das instituições, Fundamentos do regime representativo (1947).

Em 1948, juntamente com outros membros da UDN divergiu da direção do partido e participou da formação da Ação Popular Renovadora, ingressando em seguida no Partido Social Progressista (PSP). Tornou-se depois secretário-geral dessa agremiação, afastando-se do cargo durante a campanha eleitoral de 1950 para a presidência da República. O PSP, aliado ao Partido Trabalhista Brasileiro (PTB), lançou a candidatura de Getúlio Vargas, que venceu as eleições por ampla margem de votos.

Deixando a Câmara dos Deputados em janeiro de 1951, Paulo Nogueira Filho foi nomeado por Vargas, em fevereiro seguinte, conselheiro político da delegação brasileira, chefiada por João Neves da Fontoura, que compareceu à

IV Conferência de Consulta dos Chanceleres Americanos, realizada em Washington. Durante os trabalhos, apresentou aos outros membros da delegação a proposta de criação da Assembléia dos Povos Americanos, que seria constituída no âmbito da Organização dos Estados Americanos (OEA) por parlamentares eleitos nos diversos países. Embora a idéia contasse com a aprovação de João Neves, não pôde ser levada a plenário por fugir à pauta da conferência. De volta ao Brasil, foi nomeado representante do governo na delegação do Instituto de Organização Racional do Trabalho (IDORT) à Conferência da Organização Internacional do Trabalho, realizada em Bruxelas, na Bélgica, ainda em 1951. Paulo Nogueira Filho exerceu a direção do Serviço de Assistência ao Menor (SAM) durante o governo de João Café Filho, iniciado logo após o suicídio de Vargas em 24 de agosto de 1954 e encerrado com o movimento militar de 11 de novembro de 1955. Solidário ao presidente, Nogueira Filho se demitiu no dia 22 seguinte. Sua última participação política foi em apoio ao movimento político-militar que derrubou o presidente João Goulart em 31 de março de 1964. Em 18 de setembro de 1968, tomou posse na Academia Paulista de Letras. Paulo Nogueira Filho morreu em São Paulo no dia 29 de outubro de 1969. Além de discursos e conferências, escreveu diversas obras sobre história política. Reuniu sob o título de Idéias e lutas de um burguês progressista uma série de livros descrevendo sua biografia política e os principais acontecimentos de sua época, mas a sequência ficou incompleta em virtude de sua morte. Sob essa denominação geral, foram publicados dois volumes intitulados O Partido Democrático e a Revolução de 1930 (2ª ed., 1965) e quatro volumes denominados genericamente A guerra cívica de 1932 (1965-1967). Escreveu ainda Regime de liberdade social (1951), Sangue, corrupção e vergonha (1956), Conjugação das opressões capitalista e comunista (1965) e Autogestão (1969).

'title': 'NOGUEIRA FILHO, Paulo'},

'Paulo Nogueira <b class="match term0">Filho nasceu na cidade <b class="match term1">de São Paulo no dia 16 <b class="match term1">de novembro <b class="match term1">de 1898, <b class="match term2">filho <b class="match term1">de Paulo <b class="match term1">de Almeida Nogueira e <b class="match term1">de Ester Nogueira. Seu avô... <b class="match term0">Filho publicou artigo no Diário <b class="match term1">de São Paulo <b class="match term1">de 19 <b class="match term1">de abril <b class="match term1">de 1928 intitulado O pensamento... <b class="match term0">Filho morreu em São Paulo no dia 29 <b class="match term1">de outubro <b class="match term1">de 1969. Além <b class="match term1">de discursos e conferências')

Agora já temos os ingredientes necessários para treinar um modelo de entidades usando a biblioteca spacy.

```
In [13]: import random
         from pathlib import Path
         import spacy
         from spacy.util import minibatch, compounding
```

Primeiro precisamos criar o conjunto de treinamento do modelo. e deve ter a forma de uma lista como a descrita abaixo.

```
TRAIN_DATA = [
    ("nasceu em Itaporanga ( PB ) no dia 29 de dezembro de 1944 , filho de Argemiro Abílio de Sousa
     {"entities": [(10, 20, "LOC"), (69, 93, "PERSON")]})
],
```

```
In [14]: def gera_dados_treinamento(dicionário, tag):
         pass
```

[14]: 69

```

In [ ]: def main(model=pt, output_dir=dhbb_nlp, n_iter=100):
    """Load the model, set up the pipeline and train the entity recognizer."""
    if model is not None:
        nlp = spacy.load(model) # load existing spaCy model
        print("Loaded model '%s'" % model)
    else:
        nlp = spacy.blank("pt") # create blank Language class
        print("Created blank 'pt' model")

    # create the built-in pipeline components and add them to the pipeline
    # nlp.create_pipe works for built-ins that are registered with spaCy
    if "ner" not in nlp.pipe_names:
        ner = nlp.create_pipe("ner")
        nlp.add_pipe(ner, last=True)
    # otherwise, get it so we can add labels
    else:
        ner = nlp.get_pipe("ner")

    # add labels
    for _, annotations in TRAIN_DATA:
        for ent in annotations.get("entities"):
            ner.add_label(ent[2])

    # get names of other pipes to disable them during training
    other_pipes = [pipe for pipe in nlp.pipe_names if pipe != "ner"]
    with nlp.disable_pipes(*other_pipes): # only train NER
        # reset and initialize the weights randomly but only if we're
        # training a new model
        if model is None:
            nlp.begin_training()
        for itn in range(n_iter):
            random.shuffle(TRAIN_DATA)
            losses = {}
            # batch up the examples using spaCy's minibatch
            batches = minibatch(TRAIN_DATA, size=compounding(4.0, 32.0, 1.001))
            for batch in batches:
                texts, annotations = zip(*batch)
                nlp.update(
                    texts, # batch of texts
                    annotations, # batch of annotations
                    drop=0.5, # dropout - make it harder to memorise data
                    losses=losses,
                )
            print("Losses", losses)

    # test the trained model
    for text, _ in TRAIN_DATA:
        doc = nlp(text)
        print("Entities", [(ent.text, ent.label_) for ent in doc.ents])
        print("Tokens", [(t.text, t.ent_type_, t.ent_iob) for t in doc])

    # save model to output directory
    if output_dir is not None:
        output_dir = Path(output_dir)

```

```
if not output_dir.exists():
    output_dir.mkdir()
nlp.to_disk(output_dir)
print("Saved model to", output_dir)

# test the saved model
print("Loading from", output_dir)
nlp2 = spacy.load(output_dir)
for text, _ in TRAIN_DATA:
    doc = nlp2(text)
    print("Entities", [(ent.text, ent.label_) for ent in doc.ents])
    print("Tokens", [(t.text, t.ent_type_, t.ent_iob) for t in doc])
```