

Diretivas e documentação de anotação de entidades mencionadas no corpus PetroNer

Maria Clara Castro, Tatiana Cavalcanti, Elvis de Souza

Supervisão: Cláudia Freitas

30 de setembro de 2022

Este documento apresenta o conjunto de etiquetas semânticas e decisões de anotação referentes ao projeto Petrolês. Embora o termo *entidade* remeta a substâncias, *propriedades* também estão incluídas nas classes de anotação utilizadas.

Sumário

Tagset	3
Formalização da anotação	3
• Coordenação de nomes	4
• Entidades descontínuas	4
Definições e decisões relativas à anotação semântica das entidades	5
BACIA	5
CAMPO.....	7
ESTRUTURA FÍSICA (do objeto geológico)	7
EVENTO DO SISTEMA PETROLÍFERO	9
FLUIDOS: FLUIDO E FLUIDO DA TERRA	12
POÇO	17
POÇO_T; POÇO_Q e POÇO_R: Reclassificação de poços	18
ROCHA.....	19
MATERIAL NÃO CONSOLIDADO.....	20
UNIDADE CRONOESTRATIGRÁFICA.....	22
UNIDADE LITOESTRATIGRÁFICA.....	23
TEXTURA (das rochas).....	24
TIPO POROSIDADE.....	25
MISC	26
Nomes de projetos com entidades	26

Tagset

Abaixo listamos o conjunto de etiquetas utilizados para a anotação de entidades no corpus PetroNer

BACIA

CAMPO

ESTRUTURA_FÍSICA: Estrutura física do objeto geológico

EVENTO_PETRO: Eventos do sistema petrolífero

FLUIDO

FLUIDODATERRA_i: Fluido da terra (de interesse da indústria do petróleo)

FLUIDODATERRA_o: Fluido da terra (outros)

NÃOCONSOLID: Material não consolidado

POÇO

POÇO_Q: Reclassificação de poços (qualidade)

POÇO_R: Reclassificação de poços (role - papel)

POÇO_T: Reclassificação de poços (tipo)

ROCHA

TEXTURA: Textura das rochas

TIPO_POROSIDADE: Porosidade das rochas

UNIDADE_CRONO: Unidade cronoestratigráfica

UNIDADE_LITO: Unidade litoestratigráfica

Formalização da anotação

A anotação de entidades é realizada na 9ª coluna do arquivo [CoNLL-U](#), originalmente dedicada às *enhanced dependencies*, mas inutilizada no PetroNer. Seguimos o formato de anotação **IOB2** (Inside—Out—Beginning), em que B indica o início de um pedaço de entidade, I indica um token no meio de uma entidade, e O indica a ausência de entidade.

Exemplo:

Explorou	O
a	O
Bacia	B=BACIA
de	I=BACIA
Santos	I=BACIA

Além disso, não deve haver anotações *duplas* no corpus PetroNer, de modo que qualquer caso em que o token seja anotado com mais de uma entidade (ex.: B=BACIA|B=CAMPO) é considerado uma ambiguidade a ser corrigida.

- **Coordenação de nomes**

Quando há a *coordenação* de nomes de entidades, por exemplo de **bacias** (ex.: **bacias de X, Y e Z**), o termo “bacia” recebe a etiqueta B=BACIA, a preposição “de” e o nome da primeira bacia são anotados como I=BACIA, e os demais (conj) como B=BACIA. Ou seja, “bacias de X” é a primeira entidade, “Y” é a segunda e “Z”, a terceira.

Exemplo: Chagas et al. (1993) apresentaram uma análise estratigráfica e estrutural da seção rift das **bacias de Sergipe e Alagoas** na parte emersa e na plataforma rasa.

bacias	B=BACIA
de	I=BACIA
Sergipe	I=BACIA
e	O
Alagoas	B=BACIA

- **Entidades descontínuas**

As **entidades descontínuas** têm apenas o núcleo anotado como B. Seu modificador descontínuo tem anotação O e, no campo MISC, a anotação D=NOME DA ENTIDADE (ex.: D=ESTRUTURA_FÍSICA). Se já houver outros itens no MISC, a

anotação será D=NOME DA ENTIDADE e estará o mais à direita possível, ou seja, no final do MISC.

contato	B=ESTRUTURA_FÍSICA	
basal	O	
abrupto	O	MISC: D=ESTRUTURA_FÍSICA

ou

contato	B=ESTRUTURA_FÍSICA	
basal	O	
abrupto	O	MISC: xxxxx_D=ESTRUTURA_FÍSICA

Definições e decisões relativas à anotação semântica das entidades

BACIA

A etiqueta BACIA contempla as seguintes situações:

- 1) O emprego do **termo genérico “bacia”** ou **“sub-bacia”**: o termo recebe a etiqueta **B=BACIA**.

Exemplo: As **bacias** sedimentares brasileiras têm tendência a apresentar grupos específicos de rocha-reservatório.

- 2) A identificação de **bacias específicas**: em casos como [Bb]acia + de + localização/nome da bacia (ex.: Bacia de Santos), a entidade deve ser anotada como B-I-I.

Exemplo:

Bacia	B=BACIA
de	I=BACIA
Santos	I=BACIA

ATENÇÃO:

- A palavra **“Bacias”**, iniciada com letras *maiúsculas* e no plural, como em “Bacias de Sergipe e Alagoas”, é um PROPN que tem seu lema anotado como **“Bacias”**. Essa é uma decisão discutível e que foi tomada por praticidade, podendo ser alterada no futuro.

- Há situações em que o modificador da palavra bacia é uma outra entidade, como ocorre com as unidades cronoestratigráficas. Nessas situações, cada entidade deve receber sua respectiva etiqueta, B=BACIA e B=UNIDADE_CRONO.

Exemplo: Se nós temos **bacias cretáceas** produzindo, vamos empregar dinheiro nessas bacias cretáceas, mesmo que elas estejam debaixo d'água.

- Quando há alguma palavra entre o termo “bacia” e o nome próprio da bacia, como em “**bacia** sedimentar de **Curitiba**”, temos duas entidades a serem marcadas: bacia e Curitiba. Outros exemplos: bacia hidrográfica X, bacia marginal Y.

Exemplo:

bacia	B=BACIA
sedimentar	O
de	O
Curitiba	B=BACIA

- Em casos como “rifte [nome de bacia]”, “plataforma de [nome de bacia]” e “Zona de Fratura de [nome da bacia]”, anotamos o nome da bacia com a etiqueta BACIA.

Exemplos: rifte **Tucano/Jatobá**; plataforma de **Alagoas**; Zona de Fratura de **Sergipe**.

- Há casos em que a palavra “Bacia” é omitida e apenas a localização/nome da bacia está presente, de modo que é preciso observar o contexto. Na frase abaixo, as expressões em destaque são *pistas* para que anotemos “Barreirinhas” e “Pelotas” como BACIA:

- Até então a exploração de petróleo na plataforma continental brasileira tinha se caracterizado por algumas iniciativas que conduziram a trabalhos exploratórios entre 1957 e 1966, incluindo aquisição de dados de refração e reflexão sísmica e perfuração de uma série de poços estratigráficos ao longo da costa, de **Barreirinhas** a **Pelotas**.

- Bacias já foram **lagos**, e por isso o caso abaixo, embora seja precedido por "lago", refere-se a entidades do tipo BACIA. Nesta situação, apenas os nomes das bacias serão anotados.

- “As crescentes condições de aridez e evaporação associadas à subsidência tectônica promoveram a evolução de os lagos **Recôncavo, Tucano, Sergipe e Alagoas** à condição de lagos hidrologicamente fechados por quase todo o Andar Aratu, situação que durante as idades...”

CAMPO

A etiqueta CAMPO contempla as seguintes situações:

- 1) O emprego do **termo genérico “campo”**: o termo recebe a etiqueta **B=CAMPO**.
Exemplo: Os reservatórios aptianos estão situados em o bloco alto de falha principal de o **campo**.

- 2) A identificação de **campos específicos**: em casos como [Cc]ampo + de + localização/nome do campo (ex.: Campo de Lobato), a entidade deve ser anotada como B-I-I.

Exemplo:

Campo	B=CAMPO
de	I=CAMPO
Lobato	I=CAMPO

ATENÇÃO: Assim como as demais entidades, a palavra “campo” é anotada em contexto e portanto só recebe a etiqueta CAMPO quando se refere à entidade geológica. Desse modo, expressões como as seguintes não recebem qualquer tipo de anotação: *campo de distribuição; campo de tensão; campo magnético; campo da Geoquímica; trabalho de campo; guia de campo etc.*

ESTRUTURA FÍSICA (do objeto geológico)

ATENÇÃO:

- No caso de entidades que aparecem separadas no léxico, mas ocorrem juntas (ex.: “estratificação” e “convoluta”), deve-se olhar o contexto para definir qual é a entidade.

Exemplo: “Na base da seqüência S3, ritmitos espessos de arenito com *estratificação convoluta* e siltito com pseudonódulos (slurry), associados a arenitos com laminação cruzada clino-ascendente, constituem depósitos de suspensão de frente deltaica medial.” → Neste exemplo, apenas “**convoluta**” recebe a etiqueta B=ESTRUTURA_FÍSICA.

- Textura do tipo boudinage: apesar do uso da palavra “textura”, “**boudinage**” é ESTRUTURA FÍSICA.
- No contexto abaixo, “**laminação**” não recebe a etiqueta de ESTRUTURA:
“À Petrobras, pela liberação das amostras, pela autorização do uso dos laboratórios de testemunhos, **laminação** e petrografia.”
- A palavra “**dilatação**” pode ou não ser entidade, a depender do contexto.

Exemplo:

“O dique de diabásio NO está localizado na periferia deste alto gravimétrico, conforme se observa na figura 11b, indicando que esta posição corresponde a uma área de fraqueza na borda do alto estrutural e configura uma área de **dilatação** e percolação de fluidos..” (O; não é entidade)

“Há também diques de rochas básicas ao longo de falhas caracterizando zonas de **dilatação** (figs. 4 e 6b)..” (entidade do tipo **ESTRUTURA**)

Entidades ambíguas: ESTRUTURA FÍSICA ou POROSIDADE?

As palavras *vesicular* e *bioturbação* estão presentes tanto no léxico de ESTRUTURA FÍSICA como no de POROSIDADE.

- Embora *vesicular* também se refira a um tipo de porosidade - podendo, portanto, receber a etiqueta B=TIPO_POROSIDADE -, a única ocorrência de *vesicular* no corpus se refere à **ESTRUTURA FÍSICA**: “À direita , lâmina petrográfica feita com fragmentos de hialoclastito **vesicular** , provenientes de amostras de calha de o poço C , prof. 3 147 m – 3 150 m.”
- O mesmo acontece com *bioturbação*. No corpus, no entanto, todas as ocorrências de bioturbação são casos de **ESTRUTURA_FÍSICA**.

Casos especiais: *fratura*

- Apesar de “fratura” estar na lista de ESTRUTURA FÍSICA, a expressão “fratura do grão/intergranular”, por se tratar de uma escala de grão, não configura ESTRUTURA e sim TEXTURA

EVENTO DO SISTEMA PETROLÍFERO

A entidade EVENTO foi anotada a partir da lista de eventos do sistema petrolífero segundo a Carta de Eventos da ANP. Os eventos são listados desde a formação das rochas que vão compor o sistema petrolífero, passando pela migração de petróleo para o reservatório, até que se tenha o sistema completo. São eles:

ATENÇÃO: casos ambíguos

Alguns dos termos da lista podem ser ambíguos entre diferentes tipos de entidades, e a desambiguação pelo contexto é essencial para que determinemos quais deverão ser anotados.

- **geração**: configura-se como EVENTO se for acompanhada das palavras “*óleo*”, “*hidrocarboneto*”, “*gás*” ou “*petróleo*” (mas apenas *geração* será anotada).

“...a **geração** de **óleo** na bacia ocorre nos sedimentos mais profundos...”
(EVENTO)

“...permite elaborar mapas muito elucidativos de as áreas mais favoráveis para a **geração** de **hidrocarbonetos**.” (EVENTO)

“A técnica de regressão múltipla para a **geração** de **curvas** sintéticas tem sido amplamente utilizada” (O)

“Dentre os processos diagenéticos que contribuíram para a **geração** da **porosidade**...” (O)

- **gerador**:

“Foram então reconhecidos sedimentos de o cretáceo inferior com provável **potencial gerador** que , em o entanto , se encontravam em estado de maturação senil.” (EVENTO)

“Para o estudo de séries naturais de esta **seção geradora** , foram selecionadas amostras de poços...” (EVENTO)

- **migração**: configura-se como evento se estiver fazendo referência à migração de “óleo”, “hidrocarboneto”, “fluido”, “petróleo”, “gás”, “fluxo” ou “betume”. É importante ressaltar que esses termos nem sempre aparecem na frase, de modo que é necessário observar o contexto.

“Tal relação estratigráfica é responsável pela drenagem dos blocos sotopostos e conseqüente **migração** de **óleo** e **gás** para os turbiditos da seção marinha regressiva.” (EVENTO)

“Com base nesta premissa, a utilização dos valores de S, como indicadores da **migração** **primária** resultaria sempre em cálculos superestimados da eficiência de expulsão.” (EVENTO)

“O segundo tipo de reservatório de águas profundas, os complexos de canais amalgamados (fig. 1), representa o caso em que o sistema é dominado por canais relativamente rasos, caracterizados por freqüente **migração** **lateral**.” (O)

“A **migração** **lateral** ocorre em locais de valores baixos a moderados de intensidade e densidade de estruturas.” (EVENTO)

“A subida do seu nível , e a conseqüente **migração** dos **sistemas** fluviais para montante...” (EVENTO)

- **momento crítico**: não apareceu como entidade no corpus.

“Mas também, certamente, essa história de sucesso se deve a posturas adequadas de gestão; decisões tomadas em **momentos críticos** do processo exploratório definiram trajetórias, assumiram riscos, buscaram alternativas, e o resultado global está aí...” (O)

- **preservação**: configura-se como entidade se se estiver fazendo referência a termos como “matéria orgânica”, “organolitas”, “hidrocarboneto”, “petróleo” ou outros, a depender do contexto. Por exemplo, casos como “mecanismos de preservação” ou “condições de preservação” podem ou não indicar a presença de entidade.

“Anomalias mais expressivas podem representar , além da redução da taxa de sedimentação, **condições** mais favoráveis de **preservação** associadas a águas estratificadas e condições anóxicas de fundo.” (EVENTO)

“A impressão de que as assembléias florísticas apresentam variação em a quantidade de as formas e níveis de extinção , dependendo da bacia que está sendo analisada, é compreensível, visto que as distâncias envolvidas são continentais e as **condições** de **preservação** distintas.” (O)

- **reservatório**: todas as ocorrências da palavra, com apenas *uma* exceção, configuraram entidade.

“Entre outros grandes produtores em **reservatórios carbonáticos**, estão os campos gigantes Asmari...” (EVENTO)

“Na figura 13 apresenta-se a distribuição de diversos **reservatórios deltaicos** da **Formação** Coqueiro Seco ao longo da Falha 2...” (EVENTO)

“A conexão entre o *reservatório auxiliar* e o canal é normalmente realizada utilizando uma tubulação e um difusor transversal ao canal, auxiliando a distribuição do material no canal.” (O)

- **selo**: não apresentou ambiguidade e todas as ocorrências da palavra configuraram entidade.

“Esta seção argilosa encontra-se justaposta aos reservatórios ao longo do plano de falha, assim como ao longo dos flancos erodidos do bloco basculado, provendo o **selo** necessário para a formação da trapa combinada.” (EVENTO)

- **soterramento**: não apresentou ambiguidade e todas as ocorrências da palavra configuraram entidade.

“Com relação à presença de sal ou de lama, cujos contrastes de densidade com a encaixante são mais sutis, a modelagem gravimétrica é inconclusiva devido à grande profundidade de **soterramento** da estrutura.” (EVENTO)

- **trapa**: não apresentou ambiguidade e todas as ocorrências da palavra configuraram entidade.

“Como resultado arenitos depositados em ambiente lacustre são recobertos por lençóis de pelitos que isolam os mesmos de modo a fornecer importantes **trapas** estratigráficas.” (EVENTO)

FLUIDOS: FLUIDO E FLUIDO DA TERRA

FLUIDO

A etiqueta FLUIDO é utilizada para anotar fluidos que **não são** *fluidos da terra*, ou seja, não se originam de processos geológicos. Exemplos: *água desmineralizada*; *água destilada*; *fluido de perfuração*; *óleo vegetal* etc.

Em óleo + modificador (ex.: óleo vegetal, óleo do cárter), a entidade deve ser anotada como B-I:

óleo	B=FLUIDO
vegetal	I=FLUIDO

Não é FLUIDO: “inclusão fluida”

FLUIDO DA TERRA: de Interesse da Indústria do Petróleo ou Outros

A entidade **Fluido da Terra** faz parte da classe mais genérica (que não anotamos) *Materiais da Terra*, e diz respeito aos fluidos oriundos de processos geológicos (naturais). Por isso, “óleo diesel”, “óleo lubrificante”, “óleo de perfuração” e “águas fluviais” são apenas FLUIDOS, e não FLUIDOS da TERRA.

Fluido da Terra é dividido em **Fluido da Terra de Interesse da Indústria do Petróleo** (anotado como FLUIDODATERRA_i) e **Fluido da Terra Outros** (anotado como FLUIDODATERRA_o). Na classe de *Interesse da Indústria*, temos as ocorrências que acreditamos serem relevantes para quem faz uma pesquisa a partir dessa entidade fluido; já em *Fluido da Terra Outros* temos as ocorrências mais relacionadas à história geológica ou a contextos mais abertos.

São exemplos de **FLUIDODATERRA_i** as seguintes ocorrências: água de formação, óleo cru; óleo RFT; óleo leve; óleo continental; óleo autóctone; óleo bruto; óleo lacustre; óleo marinho; óleo evaporítico; betume e asfalto. A palavra

hidrocarboneto, aqui um sinônimo de petróleo, também é anotada como **FLUIDODATERRA_i**.

ATENÇÃO: petróleo, água e gás

As entidades *água* e *gás* podem ser ambíguas entre FLUIDODATERRA_i e FLUIDODATERRA_o, mas alguns modificadores e contextos dão pistas sobre a etiqueta a ser utilizada.

Pistas para que *água* seja anotada como FLUIDODATERRA_i. Em alguns casos, apenas a palavra *água* recebe etiqueta; em outros, todo do sintagma *água* + modificador recebe a etiqueta. Os elementos que deverão ser anotados estão em *itálico* em cada caso.

- a. *Água de/da formação*;
- b. *Água* dos poros;
- c. Corte de *água*
- d. Exploração de *água*
- e. Influxo de *água*
- f. Injeção de *água*
- g. Isoprodução de *água*
- h. Percolação de *água*
- i. Produção de *água*
- j. Saturação de *água*

Pistas e contextos para que *água* seja anotada como FLUIDODATERRA_o. Em alguns casos, apenas a palavra *água* recebe etiqueta; em outros, todo do sintagma *água* + modificador recebe a etiqueta. Os elementos que deverão ser anotados estão em *itálico* em cada caso.

- a. *Água* fluvial
- b. *Água* meteórica
- c. *Água* superficial
- d. *Águas* continentais
- e. *Águas* frias/quentes
- f. *Águas* oceânicas
- g. *Águas* tranquilas

- h. Evaporação das *águas*
- i. Incursões de *águas*
- j. Ingresso de *águas*
- k. Massa de *água*
- l. Trânsito de *águas*

Há casos em que não é possível generalizar a partir dos modificadores e uma análise caso a caso é necessária. Abaixo, modificadores de *água* que **não permitem generalização sobre serem fluido da terra de interesse ou fluido da terra outros:**

- a. Água do mar/ água marinha
 - FLUIDODATERRA_i: “Essa substância oleosa estaria misturada com água de o mar , que restou em esses sedimentos ; separar se de a *água do mar* por forças de capilaridade , porque , sendo mais leve , flutuaria.”
 - FLUIDODATERRA_o: “Entre os objetivos comuns , hoje , a todos os cruzeiros científicos voltados para a pesquisa marinha tem lugar o estudo de o material em suspensão encontrado em a *água do mar*.”
- b. Água doce
 - FLUIDODATERRA_i: “Os óleos de origem continental possuem teor percentual de enxofre muito baixo , sempre inferior a 0,05 % , em decorrência de a pouca solubilidade de o enxofre em *água doce* (baixo Eh) , a o contrário de o que ocorre em águas salgadas (Eh elevado) , onde a solubilidade é acentuada (fig. 2).”
 - FLUIDODATERRA_o: “É representado por folhelhos depositados em ambiente lacustre de *água doce*/ salobra de idade neocomiana.”
 - NÃO DEVE RECEBER ETIQUETA: “A influência continental ainda continua bastante intensa , notadamente em a parte inferior de o Santoniano (fig. 18) em a área de o 1-MAS-16 , com ricas associações de algas de *água doce* e foraminíferos aglutinantes de o tipo *Psammosphaera* (2).”
- c. Água salgada/água salina
 - FLUIDODATERRA_i: “Os óleos de origem continental possuem teor percentual de enxofre muito baixo , sempre inferior a 0,05 % , em decorrência de a pouca solubilidade de o enxofre em água doce (baixo Eh

), a o contrário de o que ocorre em *águas salgadas* (Eh elevado) , onde a solubilidade é acentuada (fig. 2).”

- FLUIDODATERRA_o: “O modelo conceitual mais aceito para a acumulação de evaporitos em águas rasas em as bacias de Sergipe e Alagoas advoga fluxo de *água salgada* proveniente de um mar aberto - (posicionado mais a leste) para lagunas restritas localizadas em a região mais continental (...).”

d. Água, óleo e gás (sem pista aparente)

- FLUIDODATERRA_i: “ o *óleo* vai subindo (...) fica retido logo abaixo de o *gás* e acima de a *água salgada*.”;
- FLUIDODATERRA_i: “fluidos a a base de *óleo*, *gás* e *água* podem ser aprisionados em os contatos de grão (...)

Pistas e contextos para que *gás* seja anotado como FLUIDODATERRA_i. Em alguns casos, apenas a palavra *gás* recebe etiqueta; em outros, todo do sintagma *gás* + modificador recebe a etiqueta. Os elementos que deverão ser anotados estão em itálico em cada caso.

- a. Escape de gás
- b. Fase gás
- c. Gás associado
- d. Gás biogênico
- e. *Gás condensado*
- f. Gás em solução
- g. Gás escapado
- h. Gás injetado
- i. Gás leve
- j. Gás metano
- k. Gás natural
- l. Gás original
- m. Gás pesado
- n. Gás seco
- o. Gás úmido
- p. Hidrato de gás
- q. Indício de gás

- r. Injeção de gás
- s. Jazida de gás
- t. Migração de gás
- u. Reservatório de gás

Pistas e contextos para que ***gás*** seja anotado como FLUIDODATERRA_o:

- a. *Gases nobres*
- b. Gás carbônico

- Instâncias em que o termo “**petróleo**” **não é entidade**:
 - Exploração de petróleo; indústria de petróleo; companhia de petróleo; geologia de petróleo; pesquisa de petróleo; geólogo de petróleo; geoquímica de petróleo; geografia de petróleo; monopólio de petróleo; setor de petróleo; comércio de petróleo; independência de/em petróleo.
- Expressões com **água** que **não devem ser anotadas** como entidade (referem-se a localidades, e não ao fluido da terra em si):
 - Água(s) rasa(s); água(s) profunda(s); água(s) ultraprofunda(s)/ultraprofunda(s)
- Em contextos laboratoriais **água e petróleo** também **não devem ser anotados**:

“A rotina operacional é a mesma supradescrita , com a ressalva de que se terá de filtrar mais *água*.”

“Em um diário de bordo anotar , com exatidão , o número de o par de filtros , a hora , o lugar , as condições de o mar , temperatura , salinidade e índice de Forel , e , especial mente , a quantidade de *água* filtrada.”

“Quando inclusões fluidas de *petróleo* e *água* são contemporâneas , as intersecções de as isócoras (curvas de iso-volume) , calculadas a partir de suas temperaturas de homogeneização , podem fornecer as temperaturas e pressões reais de aprisionamento de a inclusão de *petróleo*.”

Pistas e contextos em que ***gás*** NÃO DEVERÁ SER ANOTADO:

- a. Gás de iluminação

- b. Espécies enriquecidas em gás

Todas as ocorrências de *óleo* e *petróleo* são FLUIDODATERRA_i, exceto nos seguintes contextos, em que *óleo* NÃO DEVERÁ SER ANOTADO

- a. Óleo de imersão
- b. Óleo de baleia
- c. Óleo de lamparina

POÇO

Segundo a Basic Formal Ontology (BFO), um poço é um objeto caracterizado por um orifício perfurado no solo, adequadamente equipado para atingir a profundidade do reservatório ou formação de interesse.

A etiqueta POÇO contempla as seguintes situações:

- 1) O emprego do **termo genérico “poço”**: o termo recebe a etiqueta **B=POÇO**.

Exemplo: O **poço** situava-se em profundidade de água de 293 m.

- 2) A identificação alfanumérica de **poços específicos**:

- a. em poço + identificação/nome do poço (ex.: poço 1-RSS-3), a entidade deve ser anotada como B-I.

poço	B=POÇO
1-RSS-3	I=POÇO

- b. quando a identificação do poço NÃO é precedida pela palavra “poço”, ela mesma recebe a etiqueta B=POÇO.

Exemplo: No **PP-10** (B=POÇO), a espessura é maior (19 m) e os arenitos são finos.

ATENÇÃO: Em “geólogo de **poço**”, a palavra “poço” não recebe etiqueta. Isso acontece porque recuperar essa instância de “poço” dificilmente seria útil tendo em vista os objetivos das buscas semânticas que serão feitas no material.

POÇO_T; POÇO_Q e POÇO_R: Reclassificação de poços

- **POÇO_T: Tipos de poços**

A etiqueta POÇO_T refere-se à classificação de **tipos** (T) de poços. Essa classificação inclui termos como “pioneiro”, “estratigráfico”, “especial” etc. É importante frisar que esses modificadores só recebem a etiqueta quando associados à palavra “poço”. Assim, em “poço pioneiro”, temos duas entidades distintas: o token “poço” é anotado como B=POÇO, e “pioneiro” é anotado como B=POÇO_T.

Como a anotação é feita em contexto, não é necessária a presença explícita da palavra poço. No exemplo abaixo, “pioneiro” recebe a etiqueta B=POÇO_T. A partir do contexto, percebe-se que há menção a dois poços, Testa Branca 2 e Floriano – para se referir ao primeiro, usam "pioneiro", e para falar do segundo, utilizam "poço".

“Dois resultados interessantes de poços perfurados em essa etapa de trabalhos são os de Testa Branca 2 - MA , um **pioneiro** apoiado por geologia de superfície e sísmica analógica que testemunhou um intervalo de a Formação Cabeças impregnado de óleo biodegradado, e o de Floriano - MA, poço locado por geologia de superfície e onde foi testado um intervalo de arenitos de a Formação Ipu, com queima de gás.”

- **POÇO_Q: Qualidades dos poços**

A etiqueta POÇO_Q refere-se à classificação de **qualidades** (Q) dos poços. Essa classificação inclui termos como “descobridor”, “comercial”, “seco” etc. É importante frisar que esses modificadores só recebem a etiqueta quando associados à palavra “poço”. Assim, em “poço seco” temos duas entidades distintas: o token “poço” é anotado como B=POÇO, e “seco” é anotado como B=POÇO_Q.

- **POÇO_R: Funções dos poços**

A etiqueta POÇO_R refere-se à classificação das **funções** (*roles* – R) dos poços. Essa classificação inclui termos como “produtor de gás|óleo|condensado”, “injetor de óleo|CO2|água”, dentre outros. É importante frisar que esses modificadores só recebem a etiqueta quando associados à palavra “poço”. Assim, em “poço injetor de óleo”, o token “poço” é anotado como B=POÇO, “injetor” é anotado como B=POÇO_R, e “de” e “óleo” recebem a anotação I=POÇO_R.

poço B=POÇO
injetor B=POÇO_R
de I=POÇO_R
óleo I=POÇO_R

ROCHA

A entidade ROCHA faz parte da classe mais genérica (que não anotamos) *Materiais da Terra*, e refere-se a *materiais rochosos consolidados*.

ATENÇÃO:

- Ambiguidade: Os termos “*aglomerado vulcânico*”, “*subvulcânico*” e “*aglomerado*” integram tanto a lista de ROCHA quanto a lista de MATERIAL NÃO CONSOLIDADO, e por isso sua classificação dependerá do seu estado no momento da descrição do texto (consolidado ou não).
- Os termos “*siliciclásticos*” e “*terrígenos*” podem ser sinônimos de *rocha siliciclástica*. Nesses casos, recebem a etiqueta B=ROCHA.
- O termo “**carbonato**” pode ser sinônimo de *rocha carbonática*, e recebe a etiqueta B=ROCHA. No entanto, o termo “**carbonato de cálcio**” **não é ROCHA**. Dependendo do contexto, “**carbonato**” pode estar se referindo ao *teor de carbonatos* e por isso não deve ser anotado como ROCHA

Exemplos em que **carbonato** não será anotado:

“Para eliminação dos **carbonatos**, aproximadamente 0,250 g de amostra foi submetida a ácido clorídrico a quente por cerca de uma hora.”

“Neste caso, foram as amostras com maior teor semiquantitativo de argilominerais expansivos bem como ausência de **carbonatos**, micas e feldspatos potássicos.”

- Suíte **basáltica**: A expressão significa um conjunto de *rochas basálticas*, de modo que “basáltica” recebe a etiqueta B=ROCHA.

- Por analogia a “metassedimentar”, presente na lista de ROCHA, o termo “**metavulcânica**” foi adicionado na lista. Ambas são rochas metamórficas: metassedimentar é a rocha *sedimentar* que foi metamorfizada, e metavulcânica é a rocha *vulcânica* que foi metamorfizada.

MATERIAL NÃO CONSOLIDADO

A entidade MATERIAL NÃO CONSOLIDADO faz parte da classe mais genérica (que não anotamos) *Materiais da Terra*. Refere-se a materiais ainda não consolidados como tipos de *sedimentos* (*areia, pelito, argila* ou *silte*, por exemplo), *fragmentos* e *partículas*.

Os termos “**aglomerado vulcânico**”, “**subvulcânico**” e “**aglomerado**” se encontram tanto na lista de MATERIAL NÃO CONSOLIDADO quanto na lista de ROCHA, e por isso sua classificação dependerá do seu estado no momento da descrição do texto (consolidado ou não).

ATENÇÃO:

- **Vulcânico:** A palavra “vulcânico/vulcânica” **não é anotada** quando é modificada por:
 - *cone, alto, atividade, manifestação, natureza, erupção, evento, edifício, acreção, arquipélago, conduto, cunha, emanção, extrusão, fonte, ilha, margem, monte, muralha, passado, pluma, processo, plugs, plug, estrutura, construção, fase, elemento e região.*
- *As seguintes palavras se antecederem **vulcânicas/vulcânicos** devem, a princípio, caracterizar um Material da Terra (não anotado). Dessas, apenas **material, fragmento, lava** e **partícula** são mais prováveis de caracterizarem MATERIAL NÃO CONSOLIDADO.*
 - *intrusivo, **material**, sequência, composição, vidro, corpo, extrusivo, fragmento, lava, nível, **partícula**, sequência, sucessão, suíte e tufo.*

“Segundo Fisher (1997), o termo “vulcanoclástico” inclui todas as **partículas vulcânicas**, independentemente de suas origens, que podem estar relacionadas a múltiplos processos.”

partículas	B=NÃOCONSOLID
vulcânicas	I=NÃOCONSOLID

“No caso de **grãos de origem vulcânica**, esse autor também apresenta uma técnica analítica para separá los em três diferentes categorias , a saber : • grãos derivados de a erosão de antigas *seqüências vulcânicas* **[B=ROCHA]**... e são chamados de **grãos paleovulcânicos**; • grãos gerados por vulcanismo ativo em a área-fonte ou atividade submarina em a própria bacia... constituem os chamados **grãos neovulcânicos**; interações entre magma e água

grãos	B=NÃOCONSOLID
de	I=NÃOCONSOLID
origem	I=NÃOCONSOLID
vulcânica	I=NÃOCONSOLID

seqüências	B=ROCHA
vulcânicas	I=ROCHA

grãos	B=NÃOCONSOLID
neo/paleovulcânicos	I=NÃOCONSOLID

OBS.: Embora *vulcânicas*, em geral, remeta a *material não consolidado*, o sintagma *seqüências vulcânicas* do exemplo acima (em vermelho) se refere ao conjunto de *rochas* de origem vulcânica que são a fonte dos grãos citados originalmente, e por isso é entendido - e anotado - como uma referência a várias rochas.

→ “A composição do **vidro vulcânico [B=ROCHA]** constituinte dos **vulcanoclastos [B=NÃOCONSOLID]** originais seria equivalente aos traquitos...”

OBS.: *Vidro vulcânico* é um material rochoso, *consolidado*. Ele foi erodido e passou a ser um material não consolidado vulcanoclástico. Embora seja uma fonte de vulcanoclastos, o vidro citado no exemplo é uma ROCHA.

Casos especiais

- **Filmes de argila:** Apesar de “argila” ser um *material não consolidado*, o segmento “*filmes de argila*” constitui uma entidade do tipo **ESTRUTURA FÍSICA**, de modo que a anotação prioriza a leitura de estrutura. Assim, “argila” **NÃO** receberá anotação dupla (anotações duplas são consideradas erros). Um caso semelhante se encontra em “diques de arenito”.

filmes	B=ESTRUTURA_FÍSICA
de	I=ESTRUTURA_FÍSICA
argila	I=ESTRUTURA_FÍSICA

diques	B=ESTRUTURA_FÍSICA
de	I=ESTRUTURA_FÍSICA
arenito	I=ESTRUTURA_FÍSICA

UNIDADE CRONOESTRATIGRÁFICA

A etiqueta UNIDADE_CRONO corresponde a:

- 1) Intervalos de tempo geológico.

Exemplo: O **Santoniano** retoma o desenvolvimento iniciado em o **Turoniano** , transgredindo sobre os remanescentes neríticos...

- 2) Objetos geológicos identificados com base em um período (Idade, andar, série, sistema, época + NOME).

Exemplo: A maioria dos reservatórios do Membro Mucuri são de **idade Alagoas**, e, devido a sua melhor amostragem, estes reservatórios foram descritos mais detalhadamente.

idade	B=UNIDADE_CRONO
Alagoas	I=UNIDADE_CRONO

ATENÇÃO:

- Qualificadores como *inferior*, *superior*, *médio* e *terminal* também fazem parte da entidade.

Exemplo: O **Andar Aptiano Inferior** (porção terminal de o **Andar Jiquiá** e **Andar Alagoas Inferior**) é caracterizado por uma sedimentação predominantemente continental com incursões marinhas que ocorreram durante o final da fase rifte.

- Nos casos em que há abreviação de uma entidade (“Eopermiano” se torna “Eo”), ela recebe a etiqueta normalmente.

Exemplo: “Entre o **Eo** e o **Neopermiano** ocorreu na Bacia do Paraná, em resposta a três ciclos de quarta ordem de variação relativa do nível do mar, a aco-modação de três seqüências deposicionais (fig. 1) constituídas exclusivamente de tratos de sistemas transgressivos e de mar alto (Araújo, 2001).”

UNIDADE LITOESTRATIGRÁFICA

A etiqueta UNIDADE_LITO é utilizada para classificar:

- 1) Conjuntos de rochas (Membro, Formação, Grupo/Subgrupo) + NOME

Exemplo: Na base da **Formação Coqueiro Seco** ocorre uma camada de arenito com cerca de 100 m.

Formação	B=UNIDADE_LITO
Coqueiro	I=UNIDADE_LITO
Seco	I=UNIDADE_LITO

ATENÇÃO:

- Anotamos “Camada Ponta do Tubarão” como entidade, uma vez que é sinônima de “Membro Ponta do Tubarão”. Isso porque antes de ser formalizada como

“membro”, documentos antigos referiam-se a essa unidade como “Camadas P do T”.

TEXTURA (das rochas)

Em casos de “textura + modificador”, apenas o modificador poderá vir a ser anotado como entidade.

Exemplo: Seu estudo microscópico indicou *textura ofítica / subofítica* e composição essencial de plagioclásio, clinopiroxênio, opacos e matriz

textura	O
ofítica	B=TEXTURA
/	O
subofítica	B=TEXTURA

É importante notar que a entidade TEXTURA pode existir sem ser antecedida pela palavra "textura". Do mesmo modo, nem todos os casos de “textura + modificador” configuram entidades TEXTURA. No exemplo a seguir, a entidade TEXTURA está associada a “faneríticos”, que recebe a etiqueta B=TEXTURA. Enquanto isso, nem “textura” nem “gabróica” são anotadas.

Exemplo: Nos locais de estudo, os diques são **faneríticos** (às vezes com *textura gabróica*), com índices de coloração de melano a mesocrática...

ATENÇÃO

- **Cristalino:** A palavra “*cristalino/cristalina*” só deve ser anotada quando for antecedida pela palavra “*embasamento*” (que por sua vez é uma outra unidade, anotada como UNIDADE LITO). Outras ocorrências de *cristalino* no corpus não se integram a classe TEXTURA.

Exemplos:

“Desta forma, foi possível inferir que, na maioria destas bacias, os rios dissecam rochas metamórficas e ígneas ferro-magnesianas do *embasamento cristalino*, fornecendo grande aporte de lama para a plataforma.” (TEXTURA)

Os limites a sul, leste e oeste são dados por falhamentos com rochas do *embasamento cristalino*

embasamento B=UNIDADE_LITO

cristalino B=TEXTURA

No exemplo abaixo, cristalino **não** é entidade;

“Assim , eliminamos em um mapa geológico todas as áreas de *terrenos cristalinos* ou metamórficos.” (O)

TIPO POROSIDADE

A entidade TIPO POROSIDADE contemplou as seguintes ocorrências no corpus:

- caverna
- contração (shrinkage)
- intragranular
- microporosidade
- móldico
- vug/vugue

ATENÇÃO: Quando se fala em porosidade “primária” ou “secundária”, temos uma informação que NÃO diz respeito a um tipo de porosidade, mas sim a uma assinalação que indica se a porosidade foi formada durante a formação da rocha (primária) ou depois dela (secundária). Portanto, ela não é anotada. No caso abaixo, “contração” e “shrinkage” recebem a etiqueta B=TIPO_POROSIDADE.

Exemplo: A transformação de esmectita em ilita envolve desidratação e contração, o que promoveu o desenvolvimento de porosidade **secundária** de **contração** (**shrinkage**), tipo predominante em alguns intervalos da Formação Serai.

MISC

Nomes de projetos com entidades

Em casos de nomes de projetos contendo entidades, como por exemplo “**Projeto** **Bacia Pernambuco-Paraíba**” ou “**Projeto** **Andar Dom João**”, não realizamos nenhuma anotação. Apesar de até poder se revelar útil para quem busca informações, esta poderia ser considerada uma anotação falsa uma vez que não traz exatamente o significado da entidade.