Annotation Guidelines Version 09/22

KEEPHA team

 $July\ 25,\ 2024$

Contents

1 エンティティアノテーションのための一般規則 2 Drug	. 6 . 9 . 10 . 11 . 12 . 13 . 14 . 14
1 caused	. 18 . 19 . 19 . 20 . 20 . 21 . 21 . 22 . 22
2.2.2.2.2.2.2.2.2.2.3.3.3.3.3.3.3.3.3.3.3.3.3.3.3.3.3.3.	Entities 2.1 エンティティアノテーションのための一般規則 2.2 Drug 2.3 Disorder 2.4 Function 2.5 Anatomy 2.6 Test 2.7 Opinion 2.8 Measure 2.9 Time 2.10 Route 2.11 Doctor 2.11 E名化のためのエンティティ 2.13 Other 関係 3.1 caused 3.2 treatment_for 3.3 has_dosage 3.4 experienced_in 3.5 examined_with 3.6 has_result 3.7 refers_to 3.8 interacted_with 3.9 signals_change_of 3.10 has_time 3.11 has_route 3.12 is_opinion_about 3.13 misc

4	難し	い事例	2 2
	4.1	含意	23
	4.2	他のユーザからのアドバイス	23
	4.3	English	23
	4.4	French	23
	4.5	German	24
	4.6	Japanese	24

1 Introduction

以下では、開発した多言語コーパスKEEPHAのためのアノテーション方法を説明する. このコーパスのデータは、患者フォーラム、Twitter (X)などのソーシャルメディアなど、様々なソースから構成されている. Table ??では、現在アノテーションされている文書、その起源、およびその文書で使われている言語の概要を示している.

このアノテーション仕様はエンティティ、属性、および関係からなる。まず、エンティティと属性の定義について説明し(第2節)、必要に応じてアノテーションの難しい例を示す。*の付与された例は他言語から翻訳されたもの、**の付与された例は作例を表す。それ以外の例は全て実際に上記データセット内で見られたテキストである。その後、エンティティ同士を繋げる関係について説明する(第3節)。

section	entity	attributes
2.2	drug _{drug}	increase, decrease, stopped, started, unique_dose
2.2.1	change trigger _{change_trigger}	
2.3	disorder disorder	negated
2.4	$function_{function}$	negated
2.5	anatomyanatomy	
2.6	test _{test}	
2.7	opinion _{opinion}	positive, negative, neutral
2.8	measuremeasure	
2.9	time _{time}	frequency, duration, date, point
2.10	route _{route}	
2.11	doctor _{doctor}	
2.12	user _{user}	
2.12	urlurl	
2.12	${\tt personal information}_{{\tt perso_info}}$	
2.13	otherother	

Table 1: Overview of the entities and their attributes.

2 Entities

表1に、利用可能なすべてのエンティティとそれぞれの属性(Attribute)を示す。属性(Attribute)は、いくつかのエンティティアノテーションに、より正確な意味を追加するために使用される。

2.1 エンティティアノテーションのための一般規則

エンティティアノテーションのために一般的なルールを定義している.

スコープ

スコープについて、エンティティは修飾部(形容詞、副詞など)と共に名詞句や動詞句にアノテートするが、複雑な後修飾語は除外する。除外する例としては、「...である痛み」のような関係節や「昨日からの首の痛み」のような前置詞句による修飾語である。常に、核となる名詞句が最小になることが好ましい。この方針でうまくいかない場合、動詞句全体(あるいは句・文全体)の長いスパンが許容される。

```
(1) (en) I have absolutely no appetite<sub>disorder</sub> and constantly feel sick<sub>disorder</sub> to my stomach<sub>anatomy</sub>

(fr) (...) ça fait drolement mal<sub>disorder</sub> et très genant<sub>disorder</sub> .

[It hurts awfully<sub>disorder</sub> , and it is a great discomfort<sub>disorder</sub> ]

(de) (...) ich hatte in der Situation auch absolut keinen Hunger<sub>disorder</sub> .

[I was also absolutely not hungry<sub>disorder</sub> in this situation.]

(ja) ** 激しい嘔吐<sub>disorder</sub> に見舞われました

[I suffered severe vomiting<sub>disorder</sub> .]
```

メタファー(隠喩)と説明的言語

また、医学辞書に載っていない表現であっても、患者視点に基づく記述表現をアノテーションする. 常に患者視点に則るようにし、患者が「これ(症状や異常)はDisorderだ」と思ったものは、disorder とアノテートする.

```
(2) (en) (...) I felt like I would imagine someone that had MS<sub>disorder</sub> would feel like, swaying<sub>disorder</sub> , pulling myself up stairs<sub>disorder</sub> , (...)

(fr) (...) J'ai failli tomber dans les pommes<sub>disorder</sub> , (...)

[I almost passed out<sub>disorder</sub>]

(de) Die Beine<sub>anatomy</sub> fühlen sich nicht stark<sub>disorder</sub> man hat Gefühle des Ungleichgewichts<sub>disorder</sub> .

[The legs<sub>anatomy</sub> do not feel strong<sub>disorder</sub> , there are feelings of imbalance<sub>disorder</sub> ]

(ja) ずっと お腹<sub>anatomy</sub> が ゴロゴロいってる<sub>disorder</sub> [My stomach<sub>anatomy</sub> keeps saying grrrr<sub>disorder</sub>]
```

しばしば、患者は苦しみや、反対に薬物療法での良い経験に言及するために、メタファー(隠喩)を使用する. 隠喩も記述言語であるため、明示的な症状の記述のアノテーションと同じように、隠喩にも注釈を付ける. 後で、それらの "特殊な障害 "を再調査することができる.

```
(3) (en) I felt always like I had a veil over my head<sub>disorder</sub> and lead in my legs<sub>disorder</sub>.

(fr) (...) à chaque effort, mon cœur<sub>anatomy</sub> bat la chamade<sub>disorder</sub> et jai ne n'ai pas beaucoup de souffle<sub>disorder</sub>.

[With every effort, my heart<sub>anatomy</sub> is racing<sub>disorder</sub> and I am out of breath<sub>disorder</sub>]

(de) (...) nur 1 Woche später<sup>rel. point in time</sup> war ich wie im 7. Himmel<sup>positive</sup><sub>opinion</sub>.

[Just one week later<sup>rel. point in time</sup> I was on cloud nine<sup>positive</sup><sub>opinion</sub>.]

(ja) *マイスリー<sub>drug</sub> を飲むと 脳<sub>anatomy</sub> が強制終了<sub>disorder</sub> する

[Zolpidem<sub>drug</sub> makes my brain<sub>anatomy</sub> shut down<sub>disorder</sub>]
```

指示詞

指示詞(これ,あれ等)や指示代名詞(この,あの等)はエンティティに含めない.

(4) (en) *I only really felt the burning sensation_disorder in the arms_anatomy and chest_anatomy when I ... (fr) (...) J'ai développé une très forte sinusite chronique_disorder ,(...)

[I have developed severe chronic sinusitis_disorder]

(de) Seit zwei Tagen duration ist meine Übelkeit negated wieder weg, doch da ...

[Since two days duration my naused negated disorder is gone again, but since ...]

(ja) **この 薬drug は 効かなかった negative opinion [This drug didn't work negative opinion]

列挙

カンマや接続詞で区切られた連続するエンティティや列挙は、分割して別々に注釈を付ける.



入れ子構造のエンティティ

入れ子構造のエンティティ (ネストされたエンティティ)は許可されず、常に最大のフレーズでアノテーションする. 例えば、「肺がん」の場合の場合、"肺"と "がん" に単語を分割しない. 一般的なルールとしては、文法構造に従う. つまり、基本的な名詞句(前置詞、助詞などを含まない)は、1つのエンティティとしてアノテートされる.

ただし,下記のように言語に依存する:

- German: do not split compound words, e.g. "Kopfschmerzen_{disorder}" [headache] is annotated as one entity.
- 日本語:
 - 複合名詞は分割しない
 - 助詞(て, に, を, は, の)で分割する
- French: basic noun phrase ("douleur thoracique_{disorder}") [chest pain] is annotated as one disorder, otherwise we separate the annotations, e.g. "douleur_{disorder}" au thorax_{anatomy}" [chest pain].

不連続なエンティティ

より簡単な解決策がない場合,すなわち,1つのエンティティとして注釈を付けることができない,意味をなさない場合,不連続なエンティティは認められる.

1. "it hurts in the left lunganatomy and right lunganatomy" (two entities)

- 2. "pain on the left lung, but no issues were found on the right" \rightarrow annotation using two separate, partially overlapping entities, one continuous and one discontinuous as shown below:
 - "pain_{disorder} on the left lung_{anatomy}, but no issues were found on the right" (continuous entity)
 - pain on the left lung_{anatomy}, but no issues negated were found on the right_{anatomy} (discontinuous entity)
- 3. Similarly as above, the phrase "nerves / muscle cramps" is annotated as two separate, partially overlapping entities:
 - nerves_{disorder} / muscle cramps_{disorder} (discontinuous entity)
 - "nerves / muscle cramps_{disorder}" (continuous entity)

句読点

句読点はアノテーションに含めない. (ハイフン, 略語等を除く)

誤字脱字/口語表現

特に医薬品名を書くときには誤字脱字が多いため、意味が容易に分かるものであれば正しく書かれているものとして扱う.

略語

さらに、略語表現もアノテーションの対象とする.略語は医薬品名によく使われるが疾患名や 医師の専門名などにもよく使われる.

- (6) (en) (...) I felt like I would imagine someone that had $MS_{ texttt{disorder}}$ would feel like (...)
 - (fr) Pour l'instant le gastro_{doctor} ne me parle pas de rémission_{other}!

[so far, my GI_{doctor} hasn't said anything about remission_{other} !]

- (de) Ich glaube, ich bleibe auch erst einmal bei dem Gyn_{doctor} der letzten Jahre, (...). [I think I'll stay with my gyn_{doctor} of the last years for the time being.]
- (ja) * ドセ_{drug} の 浮腫み_{disorder} も気になる [I'm concerned about swelling_{disorder} by doce(taxel)_{drug}]

2.2 Drug

drug drug Tンティティは、医薬品名、商品名、剤形の記載をアノテーションする. また栄養補助食品もアノテーション対象とする.

例外として、「化学療法」や「PUVA療法」(PUVA = Psoralen and Ultraviolet A)のような薬物を用いた治療も含みます.ただし、医薬品を使用しない治療法については、Otherエンティティとしてラベルする.

(7) (en) Lipidor_{drug} did the job_{opinion} on my cholesterol_{function} both LDL and HDL.

(fr) l'entyvio_{drug} c'est pas miracle mais bon ça aide positive opinion . bon courage à tous.

| entyvio_{drug} is no miracle, but well, it does help positive opinion . good luck to all.]

(de) Ich nehme Betablocker drug of Itake beta blockers drug of Itake beta blockers of Itake beta blockers of Itake of

drug の優先 route と drug が一まとまりの単語として出現した場合, drug を route よりも優先して単語全体に付与する.

- (8) (en) [H]ave the greatest respect for pills_{drug} in general, (...)
 - (fr) (...) que tu prenais un peu de crème_{drug} le matin.
 - (de) "Urogestgel_{drug}" or "Estradiol-Creme_{drug}" "
 [Urogest_{drug} gel / Estradiol_{drug} cream]
 - (ja) 吐き気を抑えるのに 飲み薬_{drug} と 胃薬_{drug} を使いました [I used an internal medicine_{drug} and a stomach medicine_{drug} to control the nausea.]

2.2.1 Triggers for Medication Status Changes

医薬品の服用量の変化を説明するフレーズにエンティティタグ change_trigger change_trigger を 付与する. 代表的な言葉は、例えば、 *increase*、 *reduce*、 *begin*、 *one dose of などである*.

どのような変化があったかを指定するため, start (服薬開始), stop (服薬中止), increase (用量増加), decrease (用量減少), and unique_dose (一回だけの服用) といった属性(Attribute)を言及する医薬品に付与する. (2.2.2節を見よ).



2.2.2 属性: Drug Changes

drug エンティティに関する属性(Attribute)は、薬剤が服用される状況についての追加的な詳細を提供する. started, stopped, increased, decreased および unique doseという属性(Attribute)を考慮する. 以下に例を示す.

(en) I now have increased change_trigger my intake of Vitamin C_{drug} to 16,000 mg / day_{measure} (10), and we'll see how this works out. (fr) Je suis sous sertraline_{drug} (25_{measure} puis_{change_trigger} en dosage) $50_{\mathtt{measure}}$ [I've been taking sertraline_{drug} (dosage 25_{measure} $then_{change_trigger}$ $50_{measure}$)] (de) Ich hatte vor einer Woche time die Betablockerderug $reduziert_{change_trigger}$ aber nur von einer 23mg_{measure} auf die Hälfte $_{\mathtt{change_trigger}}$. 辞めました_{change_trigger} (ja) * 副作用_{disorder} が強くて、 レクサプロ $\frac{stopp}{drug}$ $II stopped_{change_trigger}$ Lexapro stop because a side effect_{disorder} was strong

患者が薬を忘れたと記述した場合,文脈から完全な停止が示唆される場合(この場合はstopを付与する)を除き,decreaseという属性(Attribute)を追加する.逆に例えば,患者が2倍の用量を服用したと記述した場合,これはincreaseを付与する.

医薬品名の略称

また、医薬品の略称や俗称もアノテートする.

- (11) (en) *In my desperation, I also tried an AD_{drug}.
 - (fr) (...) je passe une radiotest des 2 mains_{anatomy} mercredi matin^{rel. point in time} avant ma perf_{drug}, (...)

 [I'm having a CT_{test} for both hands_{anatomy} on Wednesday morning^{rel. point in time} before my IV_{drug}]

 (de) Nun habe ich also mit Candesartan stopped und Opi stopped (...) aufgehört change_trigger .

 [So now I have stopped change_trigger using Candesartan stopped and Opi stopped of the stopped of th
 - (ja) なんか $SSRI_{drug}$ 増やしてほしい。 [I need more $SSRI_{drug}$]

医薬品への言及

さらに、医薬品に言及しているが医薬品名を明示していないものもアノテートする (例 13 (英語)を参照).

(12) (en) The pill_{drug} does its job_{opinion}, (...)

(fr) (...), j'ai pris deux_{measure} pilules_{drug} par jour_{measure}, (...)

[I took two_{measure} pills_{drug} a day_{measure}]

(de) Ich nehme die Tabletten_{drug} seit 2 Tagen_{time}

[I take those pill_{drug} for two days_{time}]

(ja) 漢方薬_{drug} を購入して飲み始めた_{change_trigger} 所、(...)

2.3 Disorder

disorder disorder アノテーションは患者の健康に関連する精神的な問題を含むあらゆる疾患,徴候,症状を示す. disorder は値とパラメータと同時に表現されることもある.例えば,高LDL $_{\rm disorder}$ (パラメータ=LDL, 値=高) 1 . 値が正常範囲外の場合,これは disorder を記述しているとみなす.. disorder は,かなり大雑把な場合もあり,例えば,患者が単に「気分が悪い」と言った場合,このような表現も disorder として扱われる.

(13) (en) I triedchange_trigger the advertised Arthritis medicines stopped with severe side-effects_disorder and only tried this one because the doctor_doctor had samples.

(fr) J'ai une maladie de crohn_disorder depuis 36 ans duration (...)

[I've had crohn's disease_disorder for 36 years duration diagnostiziert.

[After a long, agonizing 11 Monaten duration of diagnosed with shingles without a rash_disorder diagnostiziert.

[After a long, agonizing 11 months duration of duration of diagnosed with shingles without a rash_disorder opinion またに 気持ち悪い disorder diagnosed with shingles without a rash_disorder opinion またに 気持ち悪い disorder diagnosed with shingles without a rash_disorder opinion またに 気持ち悪い disorder diagnosed with shingles without a rash_disorder opinion またに 気持ち悪い disorder diagnosed with shingles without a rash_disorder opinion またに 気持ち悪い disorder diagnosed with shingles without a rash_disorder opinion またに 気持ち悪い disorder diagnosed with side effects_disorder are so bad_egative |

Disorder と Function の区別

生物学的に有害な過程をdisorder,中立/良好な過程をfunctionとしてアノテーションする. すなわち,一般に,否定的なfunctionはdisorderとする. (逆は成り立たない) disorder とfunctionを区別するため,以下のルールを適用する.

- 1. 問題のエンティティが名詞句で,不調を示す場合, negated function negated function (negated は 属性, functionはエンティティ)ではなく, disorder disorder と ラベル付けする(例: 食欲不振disorder です).
- 2. しかし、連続した名詞句を越えて不調が表現される場合は、 negated function negated confunction と がいする (例: 食欲negated function がなくなった).

2.3.1 Disorderの否定

現在, negated 属性は disorder と function エンティティにのみ適用している. disorder に関して,完全に消失しない症状 (e.g. "慢性的な頭痛 $_{ ext{disorder}}$ は ほとんどなくなった $_{ ext{opinion}}^{ ext{positive}}$ ") は わずかながらも存在し続けるため negated を付与しないよう注意する. このようなほとんどの 場合, disorder の状態をより豊かに表現する opinion エンティティを伴うことになる.

(14) (en) There is no abnormality negated abnormality disorder

¹病変となる主たる名詞に修飾語が直接隣接しない場合,それぞれに適切なエンティティを付与する:例) "LDL_{function} が 高い_{measure}"

- (fr) Par contre j'étais pas plus fatigué disorder que la normale.

 [however, I wasn't more tired negated disorder than usual.]
- (de) Folgende Beschwerden sind schon weg: Nackenschmerzen negated isorder , ...

 [The following ailments are already gone: neck pain negated isorder ,]
- $(ja) * 1777 = \frac{negated}{disorder}$ $absorber{c}$ $absorber{c}$ abso

2.3.2 症状の部分否定

部分否定によって含意される症状は対象外とする.

- (15) (en) [M]y hair anatomy has not completely fallen out.
 - (fr) [L]es attaques de panique_{disorder} ont presque disparu.

 [The panic attacks_{disorder} are almost gone.]
 - (de) Die Brust_{anatomy} tut schon viel weniger weh.

 [The chest_{anatomy} hurts much less already.]
 - (ja)
 髪anatomy
 が完全に抜けたわけではありません

 [My
 hairanatomy
 has not completely fallen out.]

2.4 Function

function,は、体のすべての機能や過程を表す.体の機能はバイオマーカー(「HDL」や「白血球数」など)としてよく表現される. function には精神的な機能も含む.

- (16) (en) Lipidor_{drug} did the job_{opinion} on my cholesterol_{function} both LDL and HDL.
 - (fr) La seule chose qu'on m'a donné pour essayer de dormir_{function} c'est du Xanax_{drug}

 | Xanax_{drug} is the only thing I was given to help me | sleep_{function} |
 - (de) Opripramol_{drug} hatte_{change_trigger} ich ja nur zwei Abende lang_{time} zum Schlafen_{function}
 je eine halbe_{measure} genommen, also nur eine winzige Dosis_{measure}.

 [I had_{change_trigger} only taken half a dose_{measure} of Opripramol_{drug} for two evenings_{time} or

 sleeping_{function}, so only a tiny dose_{measure}.]
 - (ja) ** <mark>白血球_{function}</mark> は問題なかった。 [no problem with leukocyte_{function}]

2.4.1 Function の否定

Function, については,不調のさいには否定属性(negated)を付与する.これは不在や中止(例:食欲 $\frac{negated}{function}$ がなくなる),限らず,異常,すなわち,もともとの身体の過程が何らかの変化をした場合(例 白血球 $\frac{negated}{function}$ が減少している)も含む.ただし不調を表す連続的な名詞句・動詞句は negated functionではなく disorder とラベル付けすることに注意する(例: 食欲不振 $\frac{disorder}{disorder}$ もみられた).これは不調が disorder とみなされるためである (2.3節を参照).

2.5 Anatomy

anatomy によって,すべての臓器や解剖学的部位にアノテーションが施される. 通常,部分的な組織や血液細胞のような小さな部位はanatomyではなく functionとしてアノテーションする. しかし,ある文章が細胞に見られる disorderを記述している場合,その細胞は anatomy とラベル付けする.

(17) (en) Had numerous odd aches_{disorder}, especially in the leg area_{anatomy}.

(fr) (...) mais elle a fait une réaction_{disorder} au pancréas_{anatomy}, (...).

[but her pancreas_{anatomy} was affected_{disorder},]

(de) Ich besitze nur noch eine Niere_{anatomy}.

[I only have one kidney_{anatomy} left.]

(ja) ** お腹_{anatomy} かず 対すように痛い_{disorder}

[I feel a stabbing pain_{disorder} in my stomach_{anatomy}]

Anatomy エンティティは、 disorder や testエンティティのような、より大きなエンティティの中に含まれる場合は、アノテーションしない。 つまりこういった場合、 disorder や testを優先する.



2.6 Test

testエンティティは医学的検査、問診、診察、その他医学的診断に用いる結果をもたらす手順すべてを指す.



2.7 Opinion

Opinionエンティティは薬物治療(drug), 健康状態 (disorder) または生物学的プロセス(function)に対する,個人的な評価や意見を示す.このエンティティは,精神的なFunctionやdisorderの詳細な説明であることがある(例: 気分function は 安定しているfunction で). 通常,これらの評価は口語的であり,適切なスパンを見つけることは困難である.そこで,ここでも「できるだけ短いスパンで」という原則に従う.

Opinion エンティティには常に、positive、negative、neutralのいずれかのSentiment(感情)属性が付与される. Positive な評価は、多くの場合、病気の改善や、薬物療法による患者の良い経験と関連している。対応する disorder には negated 属性は付与しない. なお患者に限らず全ての評価・意見をエンティティとしてアノテーションする。これは他のエンティティのアノテーション方針と揃えることで一貫させるためである。患者に関係したエンティティは、のちの関係付与によって識別する.



絵文字 / **顔文字** 絵文字や顔文字だけでも十分に評価や感想を表現している場合, アノテーションの対象とする.

2.7.1 Sentiment属性

患者が現在の健康状態を説明するとき、あるいは服用した薬の結果を評価するとき、感情的な言葉を使うことが多い. そのため、Opinionにアノテーションされたフレーズには、該当する場合、positive, negative, neutral のいずれかの感情マーカーをSentiment属性として追加する.

(ja) レクサプロ_{drug} は 強迫_{disorder} には 効かなかったですnegative opinion

[Lexapro_{drug} didn't helpnegative with obsessive thoughts_{disorder} .]

2.8 Measure

Measure,エンティティは、薬剤の投与量や検査結果など、臨床に関連する測定値をマークする.この表現は、通常、測定単位に付随する数値である.

(22) (en) Started 2 years ago time with 10 mg measure then 6 mos later duration upped to 20 measure .

(fr) pentasa drug, cortancy drug (10 mg measure), pentasa drug, cortancy drug (10 mg measure), (de) Das Utrogest drug sind Weichkapseln route mit 100 mg measure naturid. Progesteron drug und sie sind verschreibungspflichtig.

[Utrogest drug are soft capsules route with 100 mg measure of nature id. progesterone drug and they are prescription only.]

(ja) 生検test の結果 Ki67 function が 46% measure だったため、通院にて治療中です [The biopsytest results showed a Ki67 function of 46% measure], and I am in the hospital for treatment.

testと一緒に出てくるフレーズは, opinion なのか measure なのか区別がつかない場合がある. このような場合は, measureを優先する.

経時的測定

「月に5回」のような経時的な測定表現は、薬剤の量/投与量を示す場合を除き、 time エンティティとしてアノテートする必要があることに注意する $(2.9\mathfrak{m})$. ある drug に、後述するdosage関係を付与できる場合には、その経時的測定表現は measureにできることが多い.

- 1. **I took three days' worth_{measure} of medicines_{drug} at once today date
- 2. **Mom should take this pill_{drug} three times a day_{measure} .

2.9 Time

経時的なマーカーとして、frequency (頻度)、duration (期間)、date (日時)、relative points in time (相対的時点)に関するすべての言及を Timeエンティティとしてアノテーションする。同時に、上記の細かい分類(frequency, duration, ...)をTemporality属性として付与する。ただし上記のどれにも該当しない場合はTemporality属性に値を付与しない。なお、助詞的表現(「に」、「で」、「から」、「まで」など)も時間的関係を意味するので、エンティティに含める.

Time expressions include e.g. "night", "afternoon", "for one week" (mostly duration), but also expressions like "last Monday", "in two weeks" (relative points in time), "every morning", "after lunch" (frequency) or "11.07.2022" (date).

(23) (en) Instead of taking the $pill_{drug}^{decreased}$ 2X daily_{measure}, as prescribed I take it (...) usually about 3-4 times per week_{measure}, but often skipping a week or two at a time frequency

(fr) Bonjour, Je suis sous entyvio_{drug} depuis juillet 2019 duration et ce traitement fonctionne fort bien positive opinion au niveau de la colite disorder .

[Hi, I've been taking entyvio_{drug} since july 2019 duration and the drug has been working very well positive for my colitis disorder .]

(de) Hallo liebe < user > , vor 2 Monaten el point in time habe ich meine Frauenärztin doctor gefragt, da ich zwischendurch einen Hormontest wollte.

[Hello dear < user > , 2 months agorel point in time asked my gynecologist doctor because I wanted a hormone test in the meantime.]

(ja) 3年前の2月ごろに date 乳がんが発覚

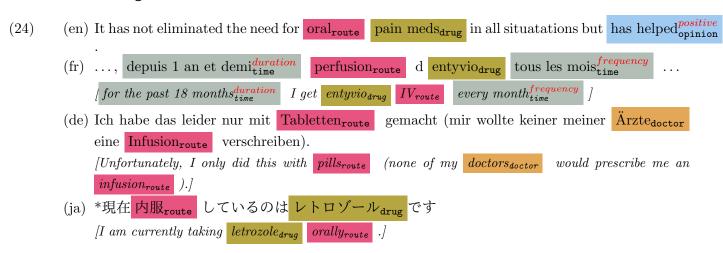
[Breast cancer disorder was discovered around February 3 years agodate]

2.9.1 Temporality属性

Timeのアノテーションでは、時間表現の属性(attribute)として frequency, duration, date, relative point in timeを考慮した. この属性はアノテーション付き時間表現をより詳細に指定するのに役立つはずである(2.9節の例を参照). どの属性値も当てはまらない場合は空欄とする.

2.10 Route

Route エンティティは錠剤や注射など、薬物治療の投与経路をアノテーションする. したがって、「飲む」「注射する」など医薬品の服用を示す動詞をアノテーションし、一般的すぎる動詞(「取る」など)は除外する. もし「錠剤」のような言及が(摂取手段ではなく)医薬品を指している場合は drug とアノテーションする必要がある.



2.11 Doctor

医師,歯科医師,看護師,(精神)療法士などの医療系職種についてもアノテーションを付ける.ただしこのエンティティタグは,医師の氏名ではなく,職業名に適用される.

(25) (en) I took lipitor $\frac{increased}{drug}$ 10mg $_{measure}$ for only about two months $\frac{duration}{time}$ then $\frac{cardiologist_{doctor}}{cardiologist_{doctor}}$ increased $_{change_trigger}$ it to $\frac{20mg_{measure}}{cardiologist_{doctor}}$.

- (fr) (peut etre de l'arthrose_{disorder} dixit mon docteur traitant_{doctor}) ... [(could be arthritis_{disorder} according to my GP_{doctor})]
- (de) (...), hatte ich gestern $_{\texttt{time}}^{\textit{rel. point in time}}$ einen Termin bei meinem Internisten $_{\texttt{doctor}}$. [..., I had an appointment with my internist $_{\texttt{doctor}}$ yesterday $_{\texttt{time}}^{\textit{rel. point in time}}$
- (ja) **病院の 先生_{doctor} がこの薬を 飲め_{route} って言うんです。でも信用できません [The doctor_{doctor} at the hospital says I should drink_{route} this medicine. But I don't trust him.]

2.12 匿名化のためのエンティティ

ユーザー名, URL, メールアドレスなどの個人情報が, たとえ前処理を施したとしても最終的なコーパスに残らないように, これらの情報にもそれぞれのラベルを付けて, 後で最終的な後処理を施すようにする. これらのエンティティを対応するマーカー, 例えば<user>に置き換えても, ドキュメントの内容が変わることはないはずである.

User

ソーシャルメディアユーザーはクリエイティブな名前を使うため、前処理ツールで見つけるのは必ずしも容易ではない. そこで、「残った」ユーザー名にもマークを付け、必要に応じて後から個人を特定できないようにする. なお以下の例はランダムに生成したもので、実際の既存データから取ったものではない.

- (26) (en) *Dear Leopard Footballer_{user}, thank you for your message.
 - (fr) Bonjour Mûre Vive_{user} pour ma part ils n'ont pas été efficace $_{opinion}^{negative}$ (...)

 [Hi Mûre Vive_{user}, for me they were not effective $_{opinion}^{negative}$]
 - (de) Liebe Broccoli Klarinette $_{user}$ danke für deinen Bericht. [Dear Broccoli Klarinette $_{user}$, thanks for your report.]
 - (ja) ゆーりん_{user} さん、ありがとうございます! [Thank you, Yurin_{user} !]

URL

匿名化後もURLや電子メールアドレスとみなせる表現が残っている場合,後で削除できるようにラベル付けします.

- (27) (en) Go to www.xxxxx.xx.xx_{url} for more info.

 - (ja) ここに詳しく載ってます www. xxx. xx_{url}

Personal information

データ公開前に削除する必要がある他のすべての個人情報(仕事,都市,医師/病院名...)にこのエンティティをアノテーションする.匿名化処理でこれらの表現が見落とされた場合に備えてのものである.以下はランダムに生成された例である.

(28) (en) **My name is Emanuel Streich_{perso_info} and I am 16 years old_{perso_info}.

(fr) J'ai appris que j'avais crohn_{disorder} il y a cinq ans^{rel. point in time} J' avais 20 ans_{perso_info}.

[I was diagnosed with crohn_{disorder} five years ago^{rel. point in time} I was 20 years old_{perso_info}.]

(de) *Ich wohne in Musterhausen_{perso_info}, das kennst du sicher.

[I live in Musterhausen_{perso_info}, you probably know it.]

(ja) ** 伊藤晴美_{perso_info}, 大学2年生_{perso_info} です。

[I am Harumi Ito_{perso_info}, a second-year college student_{perso_info}.]

2.13 Other

Otherエンティティは、治療や医療上のイベント(妊娠、外傷など)の種類に言及しているが、カテゴリーがない残りのエンティティをすべてアノテーションする。例えば、薬物ではない治療法(「認知行動療法」など)のエンティティは本仕様では定義されていないため、Otherとラベル付けする。他に考えられる例としては、治療器具(例:注射器)、医療機器(例:ウィッグ、歯科用インプラント、ペースメーカー)などがある。

- (29) (en) Just physical therapy_{other} and pain medication_{drug}.
 (fr)
 (de) Ohne BH fallen die Implantate_{other} runter.
 - [Without bra the implantsother fall off].

 (ja) * 抗がん剤_{drug} の 副作用_{disorder} で 髪_{anatomy} が 抜け_{disorder} て、 ウィッグ_{other} を着用していました。
 [Wore a wigother because of hair loss_{disorder} due to side effects_{disorder} of anti-cancer drugs_{drug}]

3 関係

以下では,エンティティ間の関係をどのようにアノテーションするかについて,詳しく説明する.より良い視覚化のために,現在のセクションの関係とエンティティのみを強調する.また,各例が1行に収まるように,原文の一部を切り取っている.表3は,すべての関係とその引数の概要を示す.例外的に, $other_{other}$ は実用上の理由から,すべての関係を取ることができる.

以下のいずれかに該当する場合は、関係アノテーションを行わない.

- 上記のエンティティが関係しない
- 書き手による仮説や推測に基づくのみである
- 質問の形式になっている

section	relation	argument 1	argument 2
3.1	caused	{drug , disorder }	{disorder , function}
3.2	${\tt treatment_for}$	{drug }	$\{ exttt{disorder} \ , \ exttt{function} \}$
3.3	has_dosage	{drug }	$\{measure\}$
3.4	experienced_in	{disorder }	$\{anatomy\}$
3.5	$examined_with$	{disorder, anatomy, function}	{test}
3.6	has_result	{test}	{measure, disorder, function}
3.7	refers_to	{disorder }	$\{ ext{disorder} \; , \; ext{function}^{negated} \}$
3.7	refers_to	{drug }	{drug }
3.7	refers_to	{anatomy}	$\{anatomy\}$
3.7	refers_to	{function}	$\{function\}$
3.8	$interacted_with$	{drug }	{drug }
3.9	signals_change_of	{change_trigger}	{drug }
3.10	has_time	{drug , disorder }	$\{ exttt{time}\}$
3.11	has_route	{drug }	{route}
3.12	is_opinion_about	{opinion}	{drug, disorder, function}
3.13	misc	{ANY}	{ANY}
	$not\ tracked$	doctor, user, URL, personal	
		info, other	

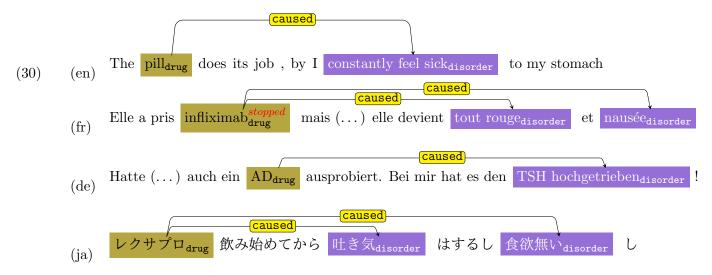
Table 2: Overview of available relations and the entities they associate.

3.1 caused

caused 関係は、言語に依存せず、言語知識が少なくても使用できるよう、下記のように定義する. drug, disorder, function などのエンティティが関係する場合のみ、caused関係をアノテートする. そのために、以下の点を考慮する.

- アノテータがはっきり判断できる範囲での書き手の意図
- 常識(例:いくら書き手が真実だと主張しても、ふつう足の骨折は歯磨きのせいで生じない).
- メタデータから得られる文脈上の手がかり(例えば、メタデータには必ずしも本文中で言及されていない薬剤名などが含まれている可能性がある).

- 次のような言語表現で明示的にわかる原因・結果の関係
 - 「だから」「なので」「のせいで」などの原因を示す表現(「(それ)で」などの弱い接続表現も対象)
 - 条件構文:「~のとき,~なった」、「~して,~なった」など
 - 制約表現:「~したときだけ,~なる」
- 後述する refer_to 関係で繋がった原因や結果を示すエンティティ
- 書き手本人に起こった現象, および親や親戚, 友人など書き手の周囲の人に起こった現象
- 原因が先で、結果が後であるという時間的関係を満たす
- 書き手の理解や仮説を裏切るような事象の発生
 - ~があるから~はならないと思っていたけど、違った!
 - ~が~になるなんて思わなかった



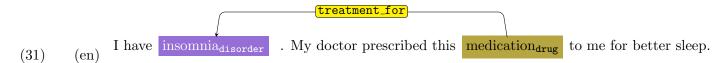
3.1.1 特別な場合: 医薬品の服用を止めたことによる影響

抗精神薬など、依存性のある医薬品を止めたことによる悪影響(離脱症状)は公的に副作用とみなされている。 したがって、薬品名 drug と、それを止めたことによる症状 disorder には caused 関係を付与する.

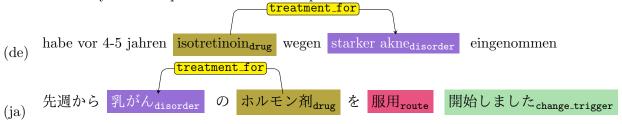
ただし、医薬品の服用を止めたことによって、その医薬品によって治療しようとしていた症状が復活することがある。アノテーションを簡単にするために、stopped drug とその結果となる disorder との間には常に caused 関係を付与する。とはいえ治療目的の症状が復活した場合には、おそらく treatment_for が別の文脈で付与されている可能性が高い。

3.2 treatment_for

この関係はdrug と対象のdisorder を結びつけ、患者の症状を治療するために使用された薬物療法を記述する.

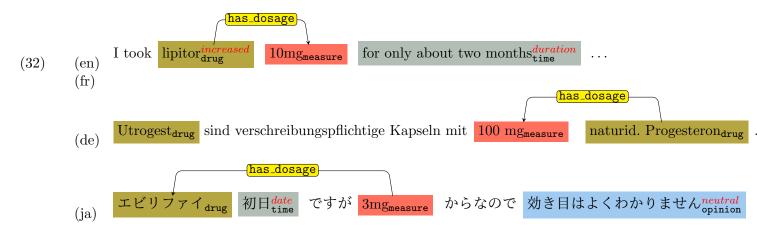


(fr) atteinte de Crohn depuis 2017, actuellement sous perfusion entyvio depuis début 2019, le remsima ayant eu trop d'effets secondaires pour moi a été abandonné.



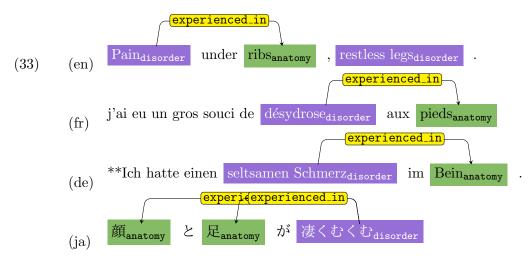
3.3 has_dosage

錠剤の数(measure) などの測定値と、薬の名前(drug)を関連付ける.



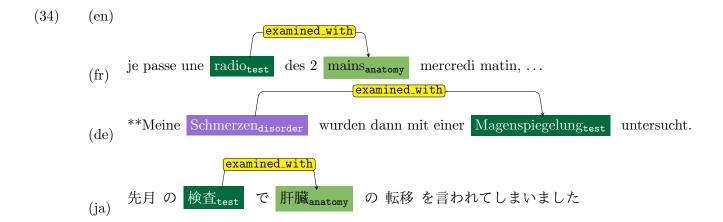
3.4 experienced_in

この関係は、disorder を、それが感じられる/経験される場所、すなわち身体の一部(anatomy)と関連付けるものである.



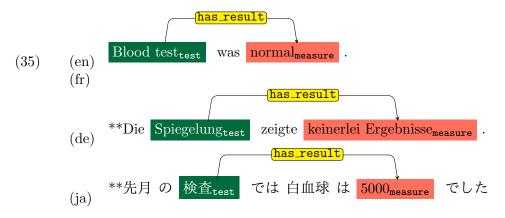
3.5 examined_with

この関係は、disorder やfunctionをtestと関連付ける. また、testは、disorder を調べるために使用され、Anatomyエンティティもまたtestで調べることができるため、examined_withで関連付ける.



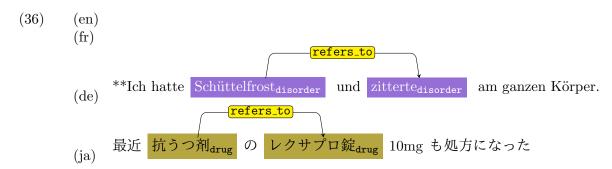
3.6 has_result

上記の考え方に沿って、has_result 関係は、test とその結果(通常measureを関連付ける.この関係は、causedが当てはまる場合は適用しない.



3.7 refers to

refers_to 関係は、おおよそ同じものを指す異なる表現を関連付けるために使われ、drug、disorder、anatomy、and function の各エンティティ間に対して一方向のリンクとして付与される。故にdisorder と function regated との間にも適用される。ただし、表面的に全く同一の表現をいちいち接続する必要はない。このrefers_to 関係を通じて、接続されたすべてのエンティティに付与された他の関係を伝播させることができる。この関係は、対応するエンティティ間で連鎖させて良いので、必ずしも最初のエンティティから関係を張る必要はない。



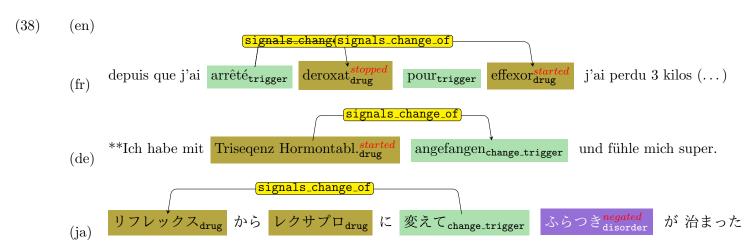
3.8 interacted_with

場合によっては、drug-drug 相互作用の結果、disorder.が発生することがある.interacted_with関係ははまさにこの目的のために定義されている.

(gr) (fr) (fr) (fr) (interacted_with (de) **Die Kombination aus Grapefruitsaftdrug und Aspirindrug hat meine Schmerzen nur verschlimme (ja) * レクサプロdrug と ビールdrug 飲んだらroute 嘔吐disorder した

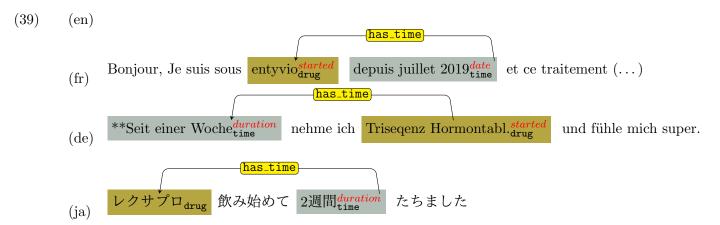
3.9 signals_change_of

change triggers と影響を受ける drug を結びつけるために signals_change_of の関係を用いる.



3.10 has_time

time と, drug かdisorderのいずれかを接続する関係である.



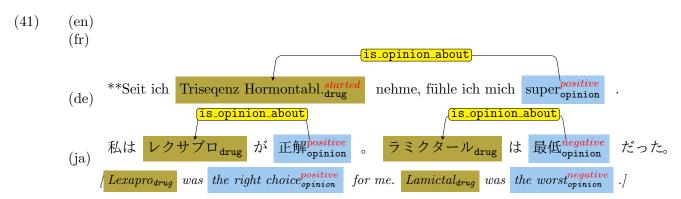
3.11 has_route

drug を使用する方法、すなわち投与経路と関連付けるものである



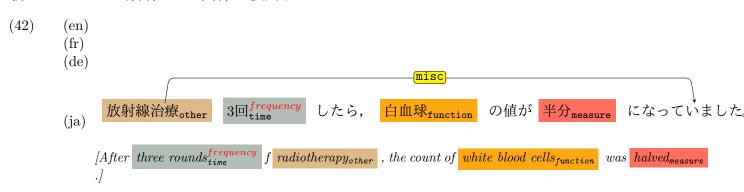
3.12 is_opinion_about

The is_opinion_about関係は,任意のエンティティと opinionエンティティを結ぶために使われる. たいていの場合,これらはdrug と function のエンティティである.接続される2つの明示的なエンティティが見つからないこともあることに注意する.



3.13 misc

misc関係は、other エンティティと同様の様々な関連を示す関係である。有害事象という観点から、2つのエンティティの間に意味ある関係があることが明らかであるが、(まだ)関係が定義されていない場合、misc関係を使用することができる。



4 難しい事例

以下では、アノテーションが困難と思われる例を紹介する. これらの例は曖昧であったり、一般に患者が何を意味したのかが明確でないものである.

4.1 含意

症状や機能について暗黙に示唆する表現は今後の課題とし,対象外とする.

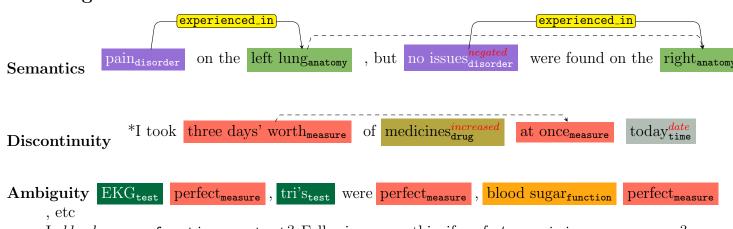
- (43) (en) My eyebrows remain halfway. (meaning the other half is gone)
 - (fr)
 - (de)
 - (ja) まつ毛が半分だけ残った(=もう半分は 脱毛_{disorder} したことを意味する)
- (44) (en) I only eat bread. (meaning the person cannot eat anything else)
 - (fr)
 - (de)
 - (ja)

4.2 他のユーザからのアドバイス

QAサイトやフォーラムなど、質問を投げかけたあるユーザ(患者)の症状を、回答者である別ユーザが回答の中で繰り返し言及することがある。アノテーションは文書単位で実施しているので、回答文については質問者の視点は導入しない。つまり、質問者が経験した副作用の関係は質問文だけでアノテーションし、回答文の中で繰り返されている部分については(回答者が経験したことではないので)関係を付与せず、エンティティのみとする。

- (45) (en) $\underline{\text{Your}}$ fatigue_{disorder} is certainly due to the heavy bleeding_{disorder}.
 - (fr)
 - (de) Gegen die Hitzewellen_{disorder} würde ich Johanniskraut_{drug} nehmen.
 - (ja)

4.3 English



Is blood sugar a function or a test? Following up on this, if perfect an opinion or a measure?

... Lipidor_{drug} did the job_{opinion} on my cholesterol_{function} both LDL and HDL. Should LDL and HDL be annotated as function?

4.4 French

... ..., des boutons_{disorder} de la tête au pied_{anatomy} , ...

Function or not? Suite à ma deuxième grossesse_{other} j'ai souhaité passé sur un traitement à faire à domicile, (...)

[After my second pregnancyother] I wanted to go ahead with home care, (...)] Depending on the context, grossesse (pregnancy) might be a function. UMLS is not clear about that.

4.5 German

4.6 Japanese

4.6.1 感情表現におけるDisorder, Function, Opinionの区別

「怖い」といった感情表現は精神疾患の兆候を表すこともあれば (disorder), 心理的に正常な反応であることもある(function). また, 医薬品に対する一種の評価と取れることもある(opinion). これらは文脈に応じて判断されなければならない.

レクサプロ_{drug}。 副作用_{disorder} が <mark>怖い_{function}。</mark> [Lexapro_{drug} . Afraid_{function} of side effects_{disorder} .] 今日からduration 漢方のみにされるらしい、<mark>こわfunction</mark> [They are going to make me take only Chinese medicine_{drug} from today^{duration} scary_{function} !] も酷い<u>negative</u>。 の 症状_{disorder} 幸い_{disorder}。 不眠disorder [I also have terrible negative opinion] of insomnia_{disorder} . Painful_{disorder} .] $symptoms_{disorder}$ 最も辛いnegative opinion 髪の毛_{anatomy} がなくなる_{disorder} ということは女性にとって、 ことです。 of hair anatomy is the most painful negative most painful negative $|The| loss_{disorder}|$ thing for a woman.

4.6.2 一般的な表現を使った症状

- 体力がガクンと落ちて_{disorder} しまった
 [My physical strength has gone through the roof_{disorder} .]
 "体力"だけでは症状エンティティになり得ないが、フレーズ全体で症状を表す。
- 何をするにも 体_{anatomy} が 重い_{disorder} 、 だるい_{disorder} 、 きつい_{disorder} [My body_{anatomy} feels heavy_{disorder} , sluggish_{disorder} , and tired_{disorder} in any way] "重い" 単体では症状ではないが、"体"と組み合わせると症状になる.
- 地獄negative のように 体調が悪いdisorder
 [My body condition goes baddisorder as hellopinion]

"体調が悪い"は、この表現全体で初めて不調を表す表現となる。フレーズになっているけれども、この表現は頻出なので、例外的に全体を disorder とする。これは「体調悪い」などの形式で形容詞的に用いられることもあるからである:"体調悪いdisorder".

4.6.3 重量表現を伴った体重の増減

● 突然の 嘔吐_{disorder} と 下痢_{disorder} で、 1日で<u>duration</u> 体重2キロ 減った_{disorder}

[Due to sudden vomiting_{disorder} and diarrhea_{disorder} , my weight_{disorder} dropped_{disorder} 2 kilograms

in one day <u>furation</u>]

"2キロ"という重量表現が"体重"と"減った"の間に挿入されており、discontinuousアノテーションとなっている.

• 手術other 前にくらべ 体重disorder が8キロ 増加disorder しました
[My weightdisorder] has increaseddisorder by 8 kg compared to before the surgeryother .]

"8キロ"という重量表現が"体重が増加"の間に挿入されていることで、 discontinuous アノテーションとなっている.