

Краткое руководство по разметке тестового корпуса. Задача «Medicine light»

Дата: 09.10.2014

Версия: 1.6.*.

Выполнено: ИСА РАН и НЦЗД.

Содержание

1. Краткий справочник аннотаций	3
2. Disease (disorder) (Заболевание (нарушение))	3
3. Body_location (Область тела).....	6
4. Severity (Тяжесть)	8
5. Course (Стадия)	9
6. Modifier (Характеристика)	10
7. Treatment (Процедуры лечения)	10
8. Drug (лекарство).....	11
9. Symptom (симптом).....	12

1. Краткий справочник аннотаций

- Disease (disorder) – заболевание (нарушение). Свойства: Negation {True, False}; Conditional {True, False}; NotPatient {True, False}. Аннотация может быть связана с Body_location, Severity, Course, Modifier.
- Body_location – область тела, к которой относится, выделенное заболевание). Аннотация должна быть связана с Disease.
- Severity – степень заболевания. Свойства: Degree {light, medium, hard}, по умолчанию medium. Аннотация должна быть связана с Disease.
- Course – течение заболевания. Свойства: Direction {unmarked, changed, improved, worsened, resolved.}. Аннотация должна быть связана с Disease.
- Modifier – модификатор или характеристика заболевания. Аннотация должна быть связана с Disease.
- Treatment – процедуры, направленные на лечение заболевания. Свойства: Effect {unknown, positive, negative, no_effect}, по умолчанию unknown; NotPatient {True, False}.
- Drug – лекарство.
- Symptom – симптом. Свойства: Negation {True, False}; SType {disorder, normal}.

2. Disease (disorder) (Заболевание (нарушение))

2.1 Эта аннотация должна назначаться словам и словосочетаниям, которые обозначают заболевания, а также различные нарушения. Предлагается размечать атомарные устойчивые словосочетания и выражения, обозначающие любые заболевания или нарушения. Атомарность здесь означает, что словосочетание обозначает только одну болезнь / нарушение. Устойчивость означает, что все словосочетание является одним термином, который достаточно часто встречается и употребляется экспертами в области медицины.

- 1) Пример: «В 6 лет (август 2008 г.) перенес отек Квинке при контакте с пони»

Здесь словосочетание «отек Квинке» необходимо обозначить как Disease.

- 2) Пример: «При погрешности в диете (молочные продукты, шоколад, куриный белок, рыба) – отек гортани»

Здесь в качестве Disease необходимо выделить слово «отек» (слово «гортань» нужно выделить как Body_location см. раздел 3).

2.2 Выделяется только непосредственное упоминание заболевания в тексте.

1) Пример: «у матери **псориаз**.»

2) Пример: «у бабушки по л/о – **поллиноз**»

Здесь выделяется только упоминание болезни, а не все словосочетание целиком.

3) Пример: «частые **респираторные инфекции**.»

Аналогично, здесь слово «частые» не должно входить в аннотацию Disease, поскольку обозначает не саму болезнь, а ее характер.

2.3 Модификаторы, которые указывают только на тяжесть протекающего заболевания, течение или стадию также не входят в аннотацию Disease.

1) Пример: «**Бронхиальная астма, атопическая**, легкое интермиттирующее течение, неконтролируемая.»

Здесь словосочетание «Легкое интермиттирующее течение» относится к категории «Severity» (тяжесть заболевания).

2.4 На Ваше усмотрение остается включать или не включать область тела («Body_location») в аннотацию заболевания.

1) Пример: «**Вторичные изменения поджелудочной железы**»

В размеченном Вами примере словосочетание «поджелудочной железы» входит в аннотацию заболевания.

2) Пример: «**дакриоцистит** правого глаза.»

В размеченном Вами примере словосочетание, обозначающее область тела «правого глаза» не входит в аннотацию заболевания.

2.5 В аннотацию заболевания не входят языковые конструкции, указывающие на отрицание.

1) Пример: «**хрипов** нет»

Здесь в аннотацию заболевания (нарушения) не входит слово «нет», поскольку оно указывает на отсутствие заболевания (но у аннотации

«хрипов» обязательно необходимо выставить свойство «Negation», см. п. 2.7).

2.6 Для заболевания (нарушения) можно указать 4 свойства: Negation, NotPatient, Conditional.

2.7 Свойство Negation выставляется аннотации Disease, если в тексте указывается, что заболевание, указанное в Disease отсутствует.

1) Пример: «**воспаление**^(Negation=True) – отсутствует»

Здесь необходимо в аннотации к слову «воспаление» выставить галку напротив свойства Negation.

2) Пример: «**дизурии**^(Negation=True) нет.»

2.8 Свойство NotPatient выставляется аннотации Disease, если из текста ясно, что выделенное заболевание относится не к пациенту, а, например, к его родственнику.

1) Пример: «у бабушки по м/л **бронхиальная астма**^(NotPatient=True)»

В аннотации к словосочетанию «бронхиальная астма» необходимо выставить галку напротив свойства NotPatient, поскольку заболевание относится не к пациенту, а к его родственнику.

2) Пример: «**туберкулезный**^(NotPatient=True) диспансер»

Здесь слово «туберкулезный» указывает на заболевание, однако оно не относится к пациенту, оно именуется медицинское учреждение, поэтому аннотации выставляется свойство NotPatient.

2.9 Свойство Conditional выставляется аннотации Disease, если из текста следует, что заболевание, обозначаемое аннотацией Disease, проявляется не всегда, а при каких-то условиях, или в тексте указано, что заболевание является условием для каких-то действий.

1) Пример: «**отек Квинке**^(Conditional=True) при контакте с пони»

Здесь заболевание «отек Квинке» проявляется при условии «контакт с пони», поэтому аннотации необходимо выставить свойство «Conditional».

2) Пример: «при возникновении **отека**^(Conditional=True), необходимо обратиться в медицинское учреждение»

В этом примере указано, что нарушение «отек» является условием для других действий «обратиться в медицинское учреждение», поэтому аннотации устанавливается свойство «Conditional».

2.10 Заболевание (нарушение) можно связать с 4 другими аннотациями: Severity, Body_location, Course, Modifier. См. разделы 3, 4, 5, 6, соответствующие этим аннотациям.

3. Body_location (Область тела)

3.1 Аннотация Body_location должна назначаться словам (существительным и прилагательным) и словосочетаниям, обозначающим области тела, с которыми связаны выделенные заболевания (аннотация Disease). Слова и словосочетания, обозначающие области тела, которые не указывают на место заболевания (нарушения) не должны быть аннотированы. От аннотации Body_location к аннотации Disease всегда должна быть проведена связь.

1) Пример: «**дакриоцистит** правого глаза»

Здесь выделена аннотация Body_location (серым цветом) и аннотация Disease «дакриоцистит» (зеленым цветом). От аннотации Body_location словосочетания «правого глаза» к аннотации Disease слова «дакриоцистит» должна быть проведена связь.

2) Пример: «**Вторичные изменения** поджелудочной железы.»

Здесь конструкция «Вторичные изменения поджелудочной железы» целиком выделена как заболевание (нарушение) Disease, словосочетание «поджелудочной железы» выделено как область тела, при этом от всей конструкции Disease должна быть проведена связь к словосочетанию с аннотацией Body_location.

3) Пример: «Мочеполовая система сформирована по мужскому типу»

Хотя словосочетание «мочеполовая система» обозначает область тела, в этом примере она не аннотируется, поскольку не указывает на место какого-либо заболевания.

4) Пример: «Носовое дыхание свободное»

Этот пример аналогичен предыдущему. Слово «носовое» не аннотируется, поскольку не указывает на место какого-либо заболевания.

3.2 Нужно назначать аннотацию `Body_location`, словам, которые непосредственно именуют область тела. Не нужно выделять как `Body_location` заболевания, указывающие на область тела.

1) Пример: «Аллергический ринит»

Здесь выделено заболевание «аллергический ринит», однако заболевание «ринит» не выделено как `Body_location`, поскольку непосредственно не именует область тела, хотя и связано с носом.

3.3 Необходимо обозначать как `Body_location` словосочетание, которое наиболее точно характеризует область тела.

1) Пример: «Перелом левой руки»

Здесь выделено все словосочетание «левой руки», а не только слово «руки», поскольку словосочетание наиболее точно характеризует область тела.

2) Пример: «аномалия подчлвчичного аппарата.»

Здесь выделено все словосочетание «подчлвчичного аппарата», а не только слово «подчлвчичного».

3.4 Аннотация `Body_location` должна включать в себя предлог или наречие, если они указывают на место некоторого заболевания.

1) Пример: «Боли в животе»

Здесь предлог «в» входит в аннотацию `Body_location`, поскольку указывает на место связанного с ним заболевания (нарушения) «боль».

2) Пример: «Опухоль была найдена снаружи поджелудочной железы»

Здесь предлог «снаружи» уточняет область тела, с которой связано нарушение «опухоль», поэтому оно входит в аннотацию `Body_location`.

4. Severity (Тяжесть)

4.1 Аннотация Severity назначается словам и словосочетаниям, которые указывают на тяжесть заболевания, обозначенного выделенной аннотацией Disease. От аннотации Severity к аннотации Disease всегда должна быть проведена связь.

- 1) Пример: «**Бронхиальная астма, atopическая форма**, **среднетяжелое течение**, контролируемая».

Здесь словосочетанию «среднетяжелое течение» назначена аннотация Severity (выделено фиолетовым цветом), которая указывает на тяжесть заболевания, обозначенного словосочетанием «Бронхиальная астма, atopическая форма», которому назначена аннотация Disease (выделено зеленым цветом). При этом от аннотации «среднетяжелое течение» к аннотации «Disease» проведена связь.

- 2) Пример: «**умеренные** **нарушения бронхиальной проходимости**»

Аналогично в этом примере словосочетанию «нарушения бронхиальной проходимости» назначена аннотация Disease, к которой проведена связь от аннотации Severity, назначенной слову «умеренные».

4.2 Для аннотации Severity должно быть указано свойство Degree. Это свойство обозначает на сколько аннотация Severity изменяет тяжесть связанного заболевания. Свойство Degree имеет три значения: {light, medium, hard}. Light – облегченное заболевание; medium – без «утяжелений» и без «облегчений» (значение по умолчанию); hard – утяжеленное заболевание.

- 1) Пример: «**Бронхиальная астма, atopическая форма**, **среднетяжелое течение**^(Degree=medium), контролируемая»

Здесь словосочетанию «среднетяжелое течение» выставлено свойство Degree=«Medium», поскольку это означает, что заболевание «бронхиальная астма» проходит без «утяжелений» и без «облегчений».

- 2) Пример: «**Бронхиальная астма, atopическая форма**, **легкое персистирующее течение**^(Degree=light)».

Здесь словосочетанию «легкое персистирующее течение» назначена аннотация Severity, которой выставлено свойство Degree=«Light», поскольку словосочетание указывает на облегченную форму заболевания «Бронхиальная астма, atopическая форма».

- 3) Пример: «**Бронхиальная астма**, **тяжелое персистирующее течение** ^(Degree=hard), **атопическая форма**, период неустойчивой ремиссии.»

Здесь словосочетанию «тяжелое персистирующее течение» назначена аннотация Severity, которой выставлено свойство Degree=«Hard», поскольку оно указывает на утяжеленную форму заболевания «Бронхиальная астма, атопическая форма».

- 4) Пример: «Беспокоит **сильный зуд**.»

5. Course (Стадия)

5.1 Аннотация Course указывает на стадию протекания выделенного заболевания (нарушения), обозначенного аннотацией Disease: «улучшение», «ухудшение», «изменение», «выздоровление» и т.д. От аннотации Course к аннотации Disease всегда должна быть проведена связь.

- 1) Пример: «**Аллергический ринит**, персистирующая форма, **ремиссия**»

Здесь словосочетанию «аллергический ринит» назначена аннотация Disease, а слову «ремиссия» - аннотация Course, поскольку это слово указывает на стадию протекания заболевания. От аннотации Course к аннотации Disease проведена связь.

- 2) Пример: «**Аллергический ринит**, **период обострения**»

Аналогично в этом примере словосочетанию «период обострения» назначена аннотация Course и от аннотации Course к аннотации Disease проведена связь.

5.2 Для аннотации Course должно быть указано свойство CType. Это свойство обозначает стадию протекания заболевания (нарушения). Свойство CType может иметь следующие значения: {unknown, improved, worsened, changed, resolved}, значение по умолчанию unknown. Improved – ремиссия, улучшение состояния пациента; worsened – обострение, ухудшение состояния пациента; changed – состояние пациента изменилось, но в тексте не указано каким образом; resolved – заболевание было вылечено.

- 1) Пример: «**Хронический гастродуоденит**, **неполная ремиссия** ^(CType=improved)»

Здесь словосочетанию «хронические гастроуденит» назначена аннотация Disease, а словосочетанию «неполная ремиссия» аннотация Course, которой выставлено свойство CType=improved.

6. Modifier (Характеристика)

6.1 Аннотацию Modifier мы предлагаем назначать словам и словосочетаниям, которые каким-либо значимым образом характеризуют заболевание или нарушение, обозначенное в тексте аннотацией Disease, однако эти слова и словосочетания не входят непосредственно в аннотацию Disease или в какие-либо другие аннотации. От аннотации Modifier к аннотации Disease всегда должна быть проведена связь.

1) Пример: «Круглогодичный аллергический ринит»

7. Treatment (Процедуры лечения)

7.1 Мы не задаем жестких ограничений на эту аннотацию. В качестве Treatment мы предлагаем выделять устойчивые словосочетания, которые на Ваш взгляд представляют «атомарные» процедуры, направленные на лечение заболеваний.

1) Пример: «получал терапию в реанимации»

Здесь словосочетание «терапия в реанимации» выделена как процедура, направленная на лечение. Слово «получал» не выделено, поскольку никак не конкретизирует эту процедуру.

2) Пример: «купировавшиеся диетой, антигистаминными препаратами, местными средствами»

Здесь отдельно выделены как Treatment слова и словосочетания: «диетой», «антигистаминными препаратами», «местными средствами».

7.2 В аннотацию Treatment могут попадать аннотации других категорий, например, лекарства.

1) Пример: «В терапии кромогексал, лекролин, зиртек с временным небольшим временным положительным эффектом»

Здесь вся конструкция «В терапии кромогексал, лекролин, зиртек» выделена как Treatment, а слова «кромогексал», «лекролин», «зиртек» выделены еще и как лекарства.

7.3 В аннотацию Treatment не следует включать выражения, указывающие на то, какой эффект дала процедура (это фиксируется отдельным свойством Effect).

1) Пример: «в терапии назначались сальбутамол, эуфиллин, муколитики с положительным эффектом»

Здесь слова «с положительным эффектом» не входят в аннотацию Treatment, поскольку они указывают на эффект от процедуры.

7.4 Аннотации Treatment может быть установлено свойство Effect. Свойство Effect указывает на то, какой эффект выделенная процедура производит на пациента {unknown, positive, negative, no_effect}. По умолчанию значение свойства Effect – «unknown».

1) Пример: «в терапии назначались сальбутамол, эуфиллин, муколитики^(Effect=positive) с положительным эффектом»

2) Пример: «В терапии местные средства^(Effect=positive), антигистаминные препараты^(Effect=positive) с положительным эффектом.»

7.5 Свойство NotPatient указывает, на то, что процедура не направлена на лечение заболевания пациента, а относится, например, к его родственнику.

1) Пример: «срочных родов путем кесарева сечения»

Из текста ясно, что процедура «кесарево сечение» относится к матери пациента, поэтому аннотации Treatment надо установить свойство NotPatient.

7.6 Аннотацию Treatment не следует связывать ни с какими другими аннотациями.

8. Drug (лекарство)

8.1 Аннотация Drug назначается словам и словосочетаниям, которые являются названиями препаратов, присутствующих в международной номенклатуре лекарственных средств.

Пример: «купируются ингаляциями с беродуалом, беротеком, вентолином»

Здесь аннотацию Drug получили слова, обозначающие названия лекарственных препаратов.

8.2 Не следует выделять словосочетания, которые, хотя и обозначают лекарственные препараты, не входят в номенклатуру лекарственных средств.

1) Пример: «курс АСИТ с пыльцевыми аллергенами сублингвальным способом введения»

Здесь словосочетание «пыльцевыми аллергенами» не выделено в качестве лекарственного препарата, поскольку оно не входит в номенклатуру лекарственных средств (однако все словосочетание должно быть отнесено к категории Treatment, см. раздел 7.)

8.3 Лекарствам не следует назначать никакие свойства.

8.4 Лекарства не следует связывать ни с какими другими аннотациями.

9. Symptom (симптом)

9.1 Мы не задаем жестких ограничений на аннотацию Symptom. В качестве симптома мы предлагаем выделять устойчивые словосочетания, которые на Ваш взгляд представляют собой «атомарные» симптомы.

1) Пример: «отек гортани, приступы затруднения дыхания»

В размеченном Вами примере отдельно выделяются словосочетания «отек гортани» и «приступы затруднения дыхания».

2) Пример: «Сердечные тоны звучные, ритмичные.»

Здесь можно выделить все предложение, кроме знака пунктуации в конце.

9.2 В аннотации симптомов могут попадать аннотации других категорий, например, заболевания (нарушения).

1) Пример: «отек гортани»

Здесь слово «отек» выделено как заболевание (нарушение), а все словосочетание – как симптом.

9.3 Аннотации Symptom можно указать два свойства: Negation, SType.

9.4 Свойство Negation показывает, относится ли фрагмент текста с аннотацией Symptom к пациенту или нет. Свойство Negation может иметь два значения {True, False} (по умолчанию False). Если Symptom не относится к пациенту свойству Negation нужно выставить значение True – иначе оставить False.

1) Пример: «Менингеальных знаков нет»

Здесь словосочетанию «менингеальных знаков» назначена аннотация Symptom. При этом аннотации выставлено свойство Negation=True, поскольку слово «нет» указывает на отсутствие симптома.

9.5 Свойство SType определяет является ли аннотация Symptom нарушением или нормальным состоянием. Свойство SType может иметь два значения {Disorder, Normal} (по умолчанию Disorder). Если симптом является нарушением (патологией), то этому свойству необходимо указать значение Disorder. Если симптом не является нарушением, то этому свойству необходимо указать значение Normal.

1) Пример: «Сердечные тоны громкие, ритмичные^(SType=Normal), небольшой систолический шум^(SType=Disorder).»

В этом примере выделено две аннотации Symptom: «громкие, ритмичные» и «небольшой систолический шум». Первой аннотации Symptom указано свойство SType=Normal, а второй аннотации указано свойство SType=Disorder.

9.6 Симптомы не следует связывать ни с какими другими аннотациями.