

ru:corner

Разметка именованных сущностей и кореференции в brat

Анастасия Никифорова	Сергей Терновых
steysie@gmail.com	fostroll@gmail.com

Денис Киреев	Константин Ремизов
dkireev.71@gmail.com	mr.enslin@mail.ru

Январь 2021

Содержание

1	Общая информация о разметке в brat	1
1.1	Начало работы. Авторизация	1
1.2	Выбор текста из коллекции	2
1.3	Выделение и разметка текста	2
1.4	Связи между размеченными упоминаниями	3
1.5	Изменение и удаление меток	4
1.6	Нерелевантные тексты	4
1.7	Поиск незнакомых терминов в Google и Википедии	5
1.8	Комментарии к сущностям	5
1.9	Поиск по документу и коллекции	5
2	Разметка именованных сущностей в brat	6
2.1	Алиасы названий сущностей в интерфейсе	7
2.2	Вложенные и пересекающиеся сущности	7
2.3	Атрибуты типов сущностей	11
2.4	Особенности Time и Date	12
2.4.1	Различия времен Absolute, Relative, Period, Duration	12
2.4.2	Выражения с неявными временными рамками	12
2.5	Некоторые правила вложенности	13
2.5.1	Про Facility, Event и Service	13
2.5.2	Про Date в Facility и Event	14
2.5.3	Про Event в Date: Absolute и Period	14
2.5.4	Про Ordinal в Time/Date:Relative	14
2.5.5	Про сочетания Family с Person и Person Property	14
3	Разметка кореференции в brat	16
4	Отдельные случаи	18
4.1	Уточняющие прилагательные	19
4.2	Уточняющие именные группы	19
4.3	Продукт, серия, уникальное название – как быть?	19
4.4	Соцсети: Service vs. Service:Media vs. Org:Media	20
4.5	Предлоги	20
4.6	Знаки препинания	20
4.7	Кавычки	21
4.8	Одиночные именованные сущности и анафора	21

4.9	Numeric: Quantity vs. Cardinal в случае эллипсиса	22
4.10	Выделение сущности, разбитой на две строки	22
4.11	Возможные ошибки и предупреждения	23
Приложение A Changelog		26

1 Общая информация о разметке в brat

brat – инструмент для разметки именованных сущностей (и других текстовых интервалов) и связей между ними. В `ru:corner brat` используется для разметки именованных сущностей и кореференции.

Окно разметки в **brat** изображено на Рисунке 1.

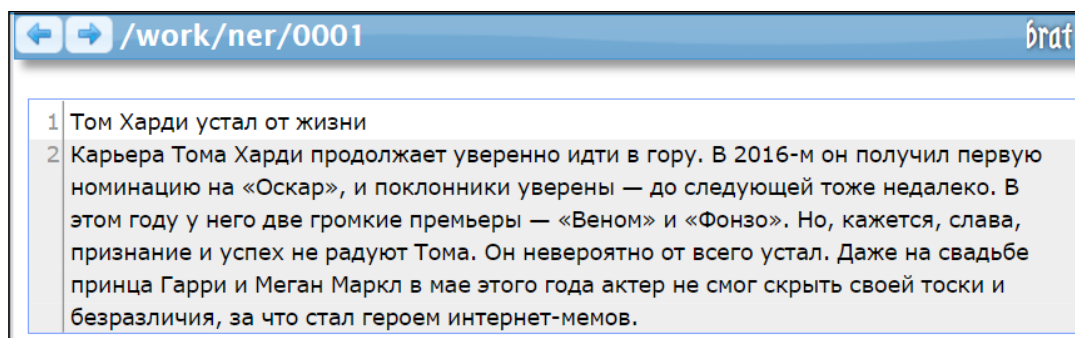


Рис. 1: Пример окна разметки в **brat**

Важно! Для разметки текстов в **brat** необходимо установить и использовать браузер **Mozilla Firefox**. Другие браузеры (Chrome, Opera и др.) склонны вызывать ошибки разметки, описанные в Разделе 4.11.

1.1 Начало работы. Авторизация

Разметка в **brat** доступна только для авторизованных пользователей. Чтобы авторизоваться, наведите курсор к шапке страницы, над текстом. В появившейся строке выберите **Login** (Рисунок 2).

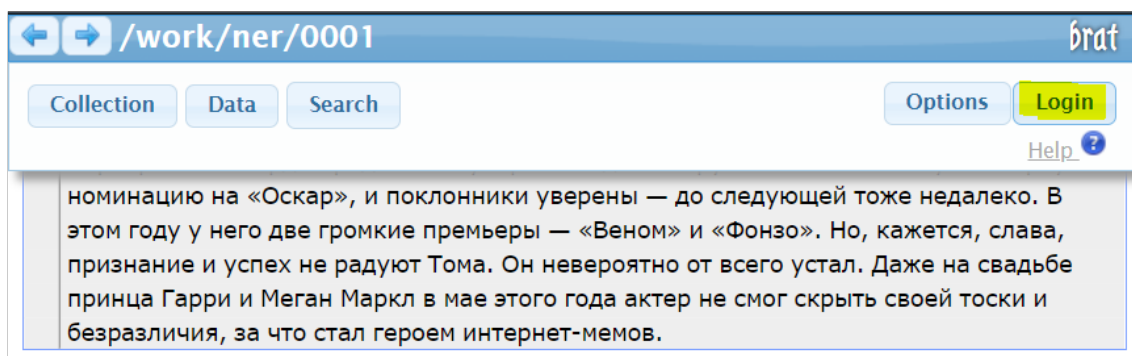


Рис. 2: Авторизация в **brat**

В появившемся окне введите логин и пароль и нажмите **ОК** (Рисунок 3). После успешной авторизации внизу страницы появится приветственное сообщение.

После входа в аккаунт можно выбрать текст из коллекции и приступить к разметке именованных сущностей или кореференции.

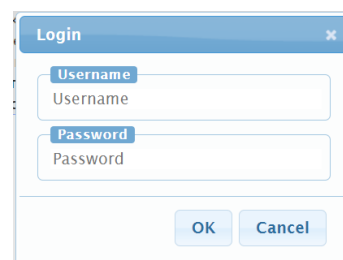


Рис. 3: Окно авторизации

1.2 Выбор текста из коллекции

Документы для разметки являются частью коллекции (Collections). Для выбора первого текста разметки, пройдите в нужную директорию и выберите файл двойным нажатием (Рисунок 4). Откроется окно с текстом как на Рисунке 1.

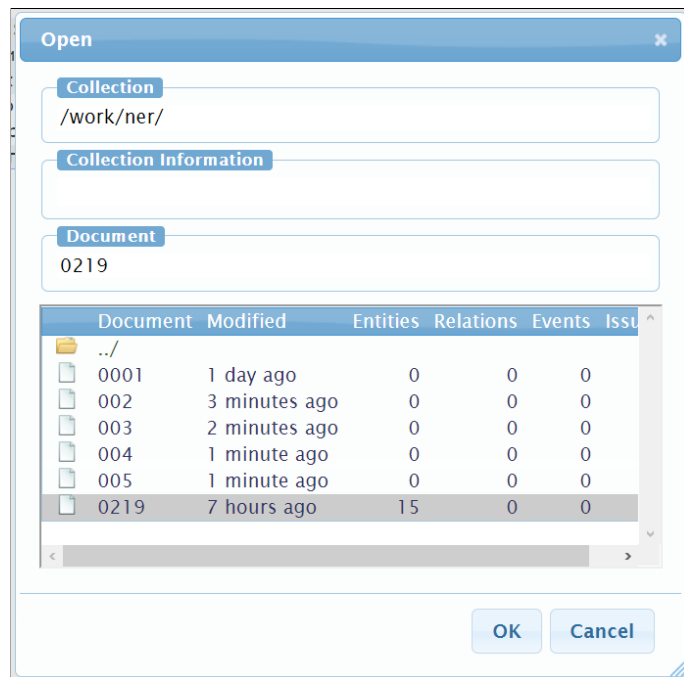


Рис. 4: Пример окна Collections

1.3 Выделение и разметка текста

Для того, чтобы разметить отрезок текста как именную сущность или упоминание, выделите курсором нужное слово или фразу от первой буквы первого слова, до последней буквы последнего слова. Единичные слова можно выделять, нажав на них дважды левой кнопкой мыши.

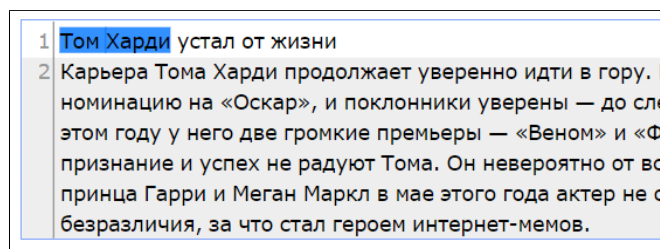


Рис. 5: Пример выделенного отрезка текста

Как только текст выделен, появится окно выбора типа сущности (Рисунок 6). Подробнее о видах меток отдельно для разметки именованных сущностей и кореференции можно прочитать в разделах 2 и 3.

Рис. 6: Пример окна выбора метки именованной сущности

В окне выбора типов сущностей выберите подходящую метку и нажмите **Enter** или щелкните курсором по **OK**.

Когда весь текст размечен, нажмите клавишу \rightarrow или кликните на кнопку \Rightarrow в левом верхнем углу **brat**.

1.4 Связи между размеченными упоминаниями

При разметке кореференции необходимо попарно связать размеченные упоминания стрелкой и выбрать тип связи.

Для образования связи, нажмите на метку одного из упоминаний (часто – зависимый член), как на Рисунке 7 и перетащите появившуюся стрелку на второе упоминание, как на Рисунке 8.

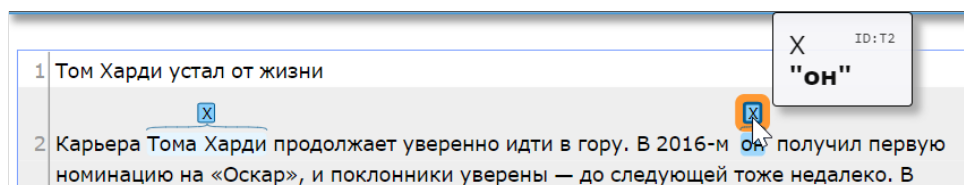


Рис. 7: Шаг 1. Выбор метки одного из упоминаний

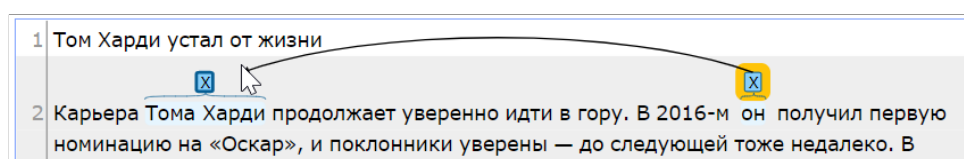


Рис. 8: Шаг 2. Связывание кореферентной группы

Когда упоминания связаны, появится окно выбора типа связи. Подробнее о видах кореферентных связей можно прочитать в разделе 3.

После выбора тип связи отобразится на стрелке (Рисунок 9).

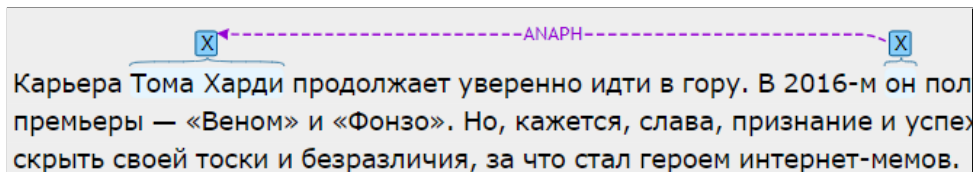


Рис. 9: Шаг 3. Выбор связи

1.5 Изменение и удаление меток

Чтобы изменить или удалить метку сущностей (в случае ошибочного выбора метки и т. п.), дважды щелкните на название метки. Появится окно выбора типа сущностей (Рисунок 10).

Для изменения типа сущности, выберите другую метку из списка.

Для изменения выделенного диапазона, нажмите **Move** и заново выделите сущность.

Для удаления метки, нажмите **Delete**.

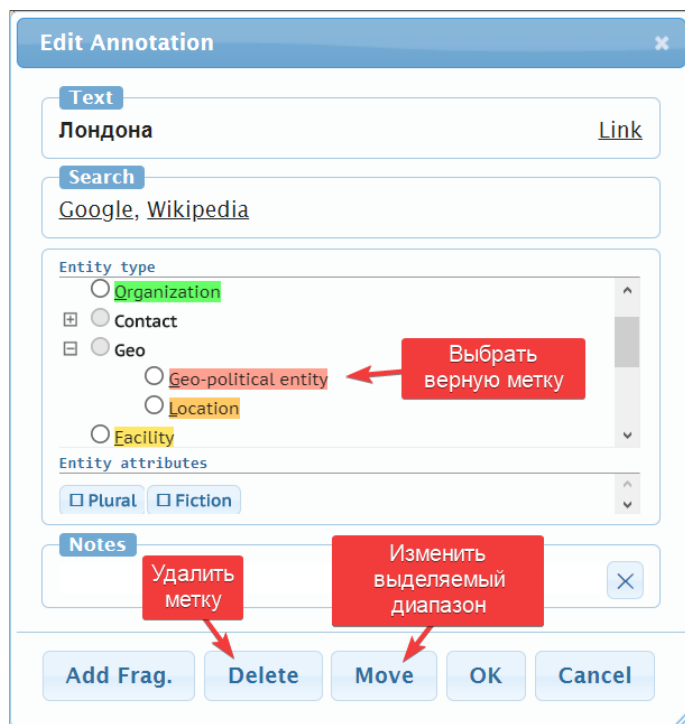


Рис. 10: Изменение и удаление меток именованных сущностей

1.6 Нерелевантные тексты

В коллекции могут попадаться тексты, в которых преобладают другие языки, которые также используют кириллицу – например, украинский или белорусский. Также могут встречаться тексты, преимущественно состоящие из стихов.

В таких текстах необходимо отметить только первое слово специальным тегом `!!! INVALID DOCUMENT !!!`. Весь остальной текст следует оставить неразмеченным, как показано на Рисунке 11.

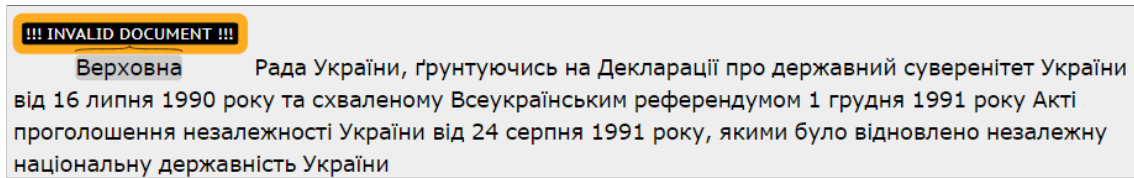


Рис. 11: Пример разметки нерелевантного текста

Если в тексте преобладает русский язык, но встречаются фразы на других языках, такой текст размечается как обычно.

1.7 Поиск незнакомых терминов в Google и Википедии

Чтобы найти значение незнакомых слов в Google и в Википедии, выделите нужный отрезок текста. В появившемся окне в разделе **Search** нажмите на **Google** или **Wikipedia** (Рисунок 12).

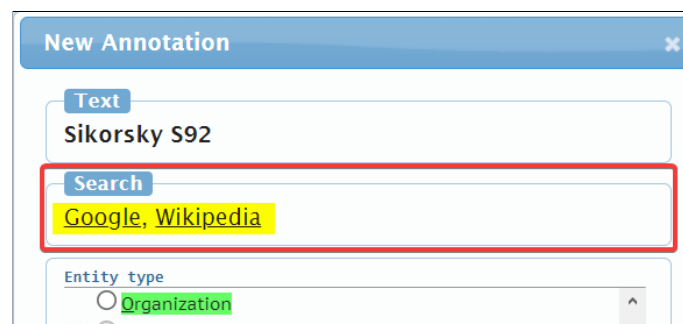


Рис. 12: Поиск незнакомых терминов в Google и Wikipedia

1.8 Комментарии к сущностям

Если вы не уверены, правильно ли выделена сущность, можно оставить короткий комментарий в окне выбора метки в разделе **Notes**. Чтобы удалить комментарий, нажмите на крестик справа в поле **Notes**.

1.9 Поиск по документу и коллекции

В brat предусмотрен удобный поиск по тексту и коллекции. Для вызова окна поиска, нажмите **Ctrl+F** или **Search** в левом верхнем углу интерфейса. На Рисунке 13 изображено окно поиска с описанием расширенных параметров поиска (появляется при нажатии **Show Advanced**). Для поиска введите в текстовое поле слово, часть слова или регулярное выражение, выберите необходимые параметры и нажмите **ОК**. В появившемся окне выберите первый документ -

искмое слово в нем будет подсвечиваться оранжевым. Для прохода по всем найденным случаям, нажимайте на клавиатуре кнопку вправо → (обратно - влево ←), пока не дойдете до последнего элемента.

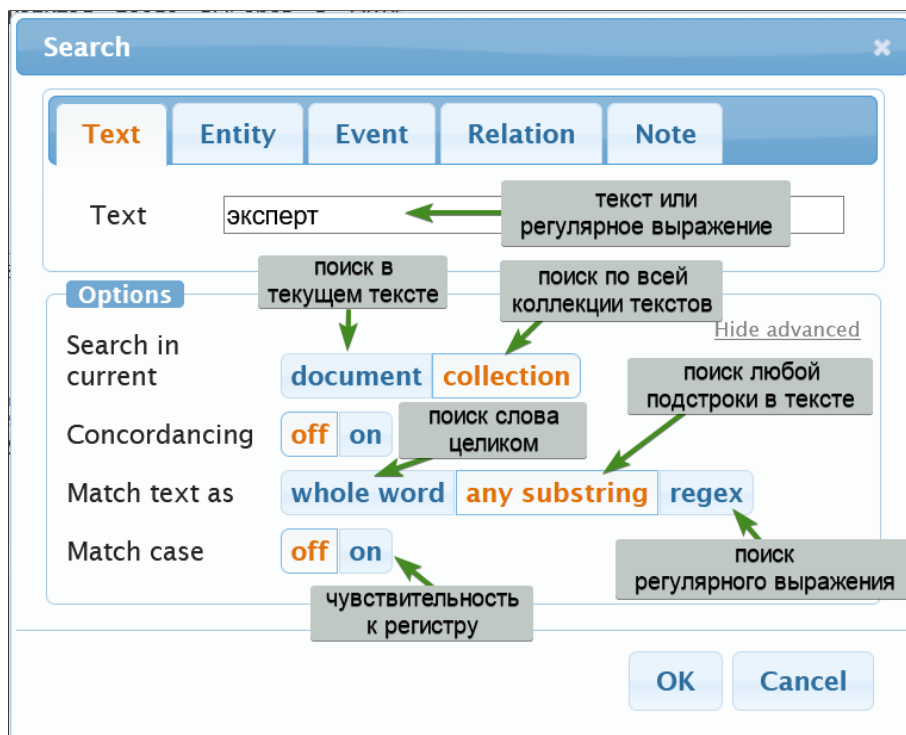


Рис. 13: Окно поиска

Пока работает поиск, коллекция отфильтрована по документам, в которых есть искомые слова. Для выхода из режима поиска, нажмите на крестик на кнопке **Search** в левом верхнем углу.

После разметки всей коллекции рекомендуется воспользоваться поиском для проверки возможных пропусков и ошибок.

2 Разметка именованных сущностей в brat

Для разметки именованных сущностей в `gui:corner` предусмотрено 44 различных типа сущностей. Подробное описание сущностей с примерами приведено ниже в Таблице 1.

Некоторые сущности могут иметь дополнительные атрибуты, один или несколько из следующих категорий: *Plural*, *Fiction*, *Unconscious*, *Department*, *Media*, *Citizenship*, *Project*, *Trademark*. Подробнее об атрибутах – в Разделе 2.3.

Для наиболее частотных типов сущностей предусмотрены горячие клавиши (колонка **Hotkey** в Таблице 1). При открытом окне выбора типа именованных сущностей нажмите подходящую горячую клавишу и нажмите **Enter**. Над выделенным упоминанием появится нужный тип сущности.

2.1 Алиасы названий сущностей в интерфейсе

В зависимости от длины выделенной фразы, в интерфейсе могут использоваться укороченные варианты одного и того же типа сущности, как показано на Рисунке 14.

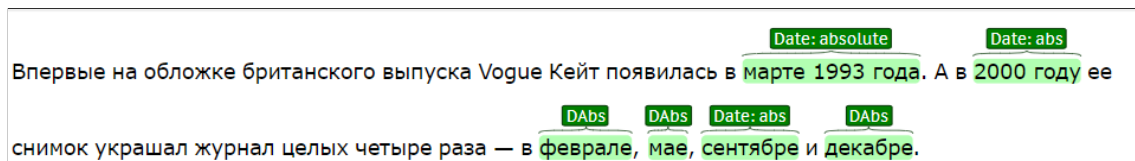


Рис. 14: Примеры алиасов сущности Date: absolute

2.2 Вложенные и пересекающиеся сущности

Именованные сущности могут быть составными, то есть могут включать в себя вложенные сущности (а-ля “матрёшка”). Примеры разметки с учетом вложенности на Рисунке 15.

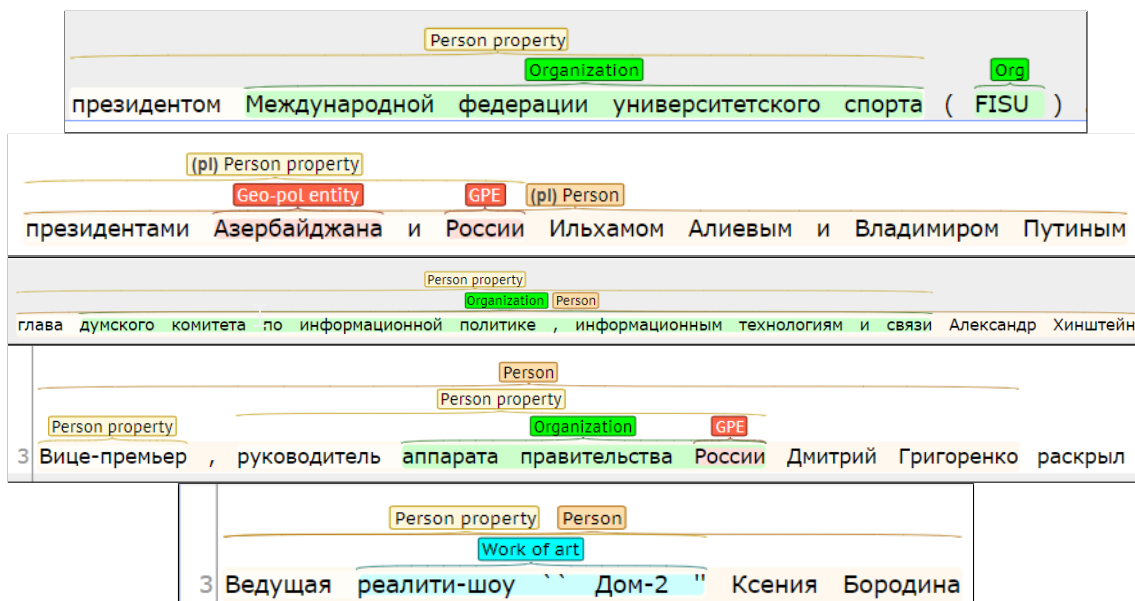


Рис. 15: Примеры пересекающихся и вложенных сущностей

Любые типы сущностей могут быть составными. Чаще других составными оказываются сущности Person, которые часто включают в себя сущности Person Property, GPE и т. д.

Примечание. Пересекающимися могут быть только сущности из категории Quantity (количество). Например, “{шестерых [членов] экипажа}”

Таблица 1. Типы сущностей, их описания и примеры

Категория	Тип сущности	Атрибуты	Hotkey	Описание типа сущности	Пример
!!! INVALID DOCUMENT !!!				Любые стихи и тексты, написанные преимущественно не на русском языке (украинский, белорусский и др.)	Отметьте первое слово в нерелевантном тексте
Person		<input type="checkbox"/> Plural <input type="checkbox"/> Fiction <input type="checkbox"/> Unconscious [default]	P	Имя или псевдоним реального человека. В качестве Person могут употребляться не только имена собственные, но и другие устойчивые выражения, зачастую имя+свойство, если это выражение часто используется в СМИ, кино или литературе для идентификации этого известного человека или персонажа.	Смирнов Иван Петрович, Сан Саныч, Машуня, Андриях, Петр I, Бейонсе, Роберт Дауни Младший, Шакил О'Нил, Шекспир, Пушкин, Моцарт, 50 cent, Моргенштерн, королева Елизавета, Президент Путин, товарищ Сталин, дорогой Леонид Ильич, президент Борис Ельцин
		<input checked="" type="checkbox"/> Plural		Имена нескольких людей в одной сущности	Павел и Марина Смирновы, Дэвид и Виктория Бэкхам
		<input checked="" type="checkbox"/> Fiction		Имена или псевдонимы вымышленных персонажей, богов (кино, литература, комиксы, религия, мифология и т. п.)	Евгений Онегин, Губка Боб, попугай Кеша, Энакин Скайуокер, Иисус Христос, Будда, кот Матроскин, Мальчик который выжил, Тот-чь-имя-нельзя-называть
		<input checked="" type="checkbox"/> Unconscious		Любые имена, которыми люди называют неразумные уникальные объекты: животные, техника, корабли и т. п. (включая тип объекта, если он указан, напр., "корабль")	Барсик, Белка, Стрелка, Несси (лох-несское чудовище), броненосец "Потемкин", корабль "Сметливый", баркас "Fishizzle", "Союз МС-16", эсминец Джон Маккейн, корабль Орион, ледокол "Арктика"
Person property		<input type="checkbox"/> Plural <input type="checkbox"/> Fiction [default]	R	Чин, звание, титул, должность, профессия и т. п.	генерал-майор, королева, канцлер, сварщик, менеджер, президент, ученые, разработчики, депутаты, чиновник, папа римский, специалист, эксперт, исследователи, срочник, заключенные, задержанные
		<input checked="" type="checkbox"/> Plural		Несколько Property в одной сущности	"главы России, Азейбарджана и Армении"
		<input checked="" type="checkbox"/> Fiction		Вымышленные титулы/профессии и т. п.	профессор защиты от темных искусств, штурмовик Первого Ордена
Organization		<input type="checkbox"/> Plural <input type="checkbox"/> Fiction <input type="checkbox"/> Department <input type="checkbox"/> Media [default]	O	Компании, агентства, радиостанции, институты, политические партии, армии стран и т. п., имеющие орг. структуру	Ростелеком, Amazon, ПАО "Газпром", партия "Единая Россия", РПЦ, компания Орифлейм, Белорусская оппозиция, турецкая армия
		<input checked="" type="checkbox"/> Plural		Несколько организаций или отделов в рамках одной сущности	Норильский и Алтайский горнодобывающие заводы, финансовый и юридический отделы, Московская и Екатеринбургская епархии
		<input checked="" type="checkbox"/> Fiction		Вымышленные организации (кино, литература и т. п.)	галактический сенат, Stark Industries, Снектр, школа волшебства Хогвартс
		<input checked="" type="checkbox"/> Department		Отделы внутри организаций	IT-департамент, отдел кадров, совет директоров, топ-менеджмент
		<input checked="" type="checkbox"/> Media		СМИ: издательства, журналы, газеты и т.п., включая печатные издания	издательство ЭКСМО, Эхо Москвы, Лента.ру, Cosmopolitan, АиФ, Первый канал
Contact	Address		A	Адрес местоположения, зачастую - город, улица, дом, квартира, индекс	123456, Москва, Тверская улица, дом 10, строение 2; адрес: ул. Ленина, д. 5
	Phone		H	Номер телефона	+7 (123) 456-78-90, тел. 81234567890, телефон 44-22-33
	Email		M	Адрес электронной почты	name@wsite.com, email: user@edu.site.org
	Web address			Ссылка на веб-страницу	https://website.com/info/, google.com
	Other-contact			Имя пользователя, профиль в instagram и т. п.	@someusername, telegram: @bestname, tg username, instagram: followme
Event		<input type="checkbox"/> Plural [default]	E	Ураганы, сражения, войны, спортивные состязания, праздники и т. п. (не одиночные слова типа "концерт", "церемония" и т. д.)	Новый год, ураган Катрина, Чемпионат мира по футболу, Брусиловский прорыв, пандемия коронавируса, эпидемия, предизидентские выборы в США, бой Хабиба и Коннора, концерт Баскова и Киркорова (один концерт)
		<input checked="" type="checkbox"/> Plural		Несколько Event в рамках одной сущности	ураганы Катрина и Рита, концерты Баскова и Киркорова (разные мероприятия)
Geo	Geo-political entity (GPE)	<input type="checkbox"/> Plural <input type="checkbox"/> Fiction [default]	G	Географическая зона, имеющая политическую структуру + космические станции (GPE)	Ростов-на-Дону, Швейцария, Ближний Восток, г. Москва, СНГ, СССР, Катманду, деревня Кунцево, Алтуфьевский район, столица (если понятно, что Москва), МКС
		<input checked="" type="checkbox"/> Plural		Несколько GPE	Московская и Ленинградская области
		<input checked="" type="checkbox"/> Fiction		Вымышленная географическая зона	Атлантида, Хогвартс, Нарния, Вестерос, Средиземье, Готэм-сити, Асгард
	Location	<input type="checkbox"/> Plural <input type="checkbox"/> Fiction [default]	L	Места, природные: горные цепи, водоемы + планеты, галактики, созвездия, кометы и т. п.	озеро Байкал, Волга, р. Нева, оз. Чад, Гималаи, Эверест, Земля, Луна, Марс, Астероид 501647, Млечный путь, естественный спутник Земли, космос
		<input checked="" type="checkbox"/> Plural		Несколько Location	реки Тигр и Евфрат
		<input checked="" type="checkbox"/> Fiction		Вымышленные места	Мглистые горы, Андуин, река Яруга

Таблица 1. Типы сущностей, их описания и примеры

Категория	Тип сущности	Атрибуты	Hotkey	Описание типа сущности	Пример
Facility		<input type="checkbox"/> Plural <input type="checkbox"/> Fiction [default]	F	Достопримечательности, здания, аэропорты, шоссе, парки, мосты, улицы, площади, переулки и т. п.	памятник Ильичу, Внуково, Трасса М4, ТРЦ "Авиапарк", Дворцовый мост, Арбат, Трехсвятская улица, проспект Ленина
		<input checked="" type="checkbox"/> Plural		Несколько Facility	трассы М4 и М5
		<input checked="" type="checkbox"/> Fiction		Вымышленные Facility (кино, литература и т. п.)	Замок Саурона, Замок Дарта Вейдера, Галактический рынок
Language		<input type="checkbox"/> Plural <input type="checkbox"/> Fiction [default]		Языки естественные или искусственные (не из кино/литературы)	русский язык, итальянский, иврит, хинди, язык йоруба, урду, эсперанто, северные диалекты русского языка
		<input checked="" type="checkbox"/> Plural		Несколько объектов Language в рамках одной сущности	южные и восточные диалекты, тюркские и финноугорские языки, говоры и наречия русского языка
		<input checked="" type="checkbox"/> Fiction		Вымышленные языки (кино, литература)	клингон, на'ви, дотракийский, новояз
Social-group	Nationality	<input type="checkbox"/> Citizenship <input type="checkbox"/> Plural [default]		Принадлежность к нации по происхождению, рождению или иным образом	русские, итальянец, гречанка, папуасы, финны, афроамериканцы, кореец, индус, армянского происхождения
		<input checked="" type="checkbox"/> Citizenship		Гражданство какой-либо страны	граждане РФ, гражданка Эстонии, жители Китая, американцы
		<input checked="" type="checkbox"/> Plural		Несколько разных объектов Nationality в одной сущности	Граждане Грузии и Армении, жители Северной и Южной Кореи
	Family		Y	Обозначения родственных связей	мать, брат, старший сын, сводная сестра, двоюродная тетя, брат жены, семья, род, сожители, пара, молодожены
	Religious group	<input type="checkbox"/> Plural [default]		Принадлежность к определенной религии	православные, католик, старообрядцы, амиши, шииты, буддист, пастафарианцы
		<input checked="" type="checkbox"/> Plural		Несколько разных религиозных групп в одной сущности	православные, католические и протестантские христиане
	Political group	<input type="checkbox"/> Plural [default]		Принадлежность к политической группе	республиканцы, члены партии "Единая Россия", национал-демократы, коммунисты, члены партии «Свобода»
		<input checked="" type="checkbox"/> Plural		Несколько политических групп в одной сущности	члены КПРФ и ЛДПР, члены партий Единая Россия и Яблоко
	Other group	<input type="checkbox"/> Plural <input type="checkbox"/> Fiction [default]		Другие социальные группы, члены которых имеют что-то общее	безработные, веганы, гангстеры, новые русские, готы, геи, бомжи, преступник, избиратели, бюджетники, поклонники
		<input checked="" type="checkbox"/> Plural		Несколько разных групп в одной сущности	представители готов и эмо, участники клубов рукоделия и гончарства
		<input checked="" type="checkbox"/> Fiction		Вымышленные социальные группы	хоббиты, эльфы, орки, джедаи, члены банды "Железные рукава"
Product		<input type="checkbox"/> Plural <input type="checkbox"/> Project <input type="checkbox"/> Trademark [default]	U	Электроника, автомобили, оружие, продукты питания, одежда, техника, серии водных/космических судов и т. п. (конкретные продукты с указанием серии/производителя)	пылесос Dyson, диваны Ikea, Range Rover, Айфон, Орион чокопай, пистолет Макарова, АК-47, косметика Орифлейм, платье бренда Виктория Бекхам, танк Т34, космические корабли "Союз", Катюша (оружие)
		<input checked="" type="checkbox"/> Plural		Несколько продуктовых сущностей в рамках одной	диваны Ikea и Hoff, пылесосы Dyson и Xiaomi, косметика Эйвон и Орифлейм
		<input checked="" type="checkbox"/> Project		Проекты, программы (государственные, научные, космические и т. п.)	мегасаенс-проект NICA, нацпроект "Наука", программа обмена студентами Erasmus +, программа «Вояджер»
		<input checked="" type="checkbox"/> Trademark		Бренд, запатентованная технология, торговая марка продукта (не в значении "компания/организация")	Dyson, Apple, Орифлейм, Виктория Бекхам, Шанель, Tesla, технология ProMotion
Service		<input type="checkbox"/> Plural <input type="checkbox"/> Media [default]	S	Различные предприятия, приложения и технологии предоставления услуг, онлайн-платформы	чистка обуви, ремонт ноутбуков, доставка еды, скорая помощь, Яндекс.Еда, Telegram, TikTok, YouTube, Facebook
		<input checked="" type="checkbox"/> Plural		Несколько сервисов в одной сущности	услуги прачечной и химчистки, услуги телефонии и интернета
		<input checked="" type="checkbox"/> Media		Аккаунты, каналы, страницы, сайты, сообщества и т. п. (главный признак - они могут публиковать контент и имеют уникальное имя)	Телеграмм-канал "IT юмор", страница Ольги Бузовой, сайт ООН, твиттер-аккаунт Трампа, в т. ч. одиночные "пост", "страница" и др., если есть уточнение платформы
Work of art				Названия книг, песен, картин, фильмов, ТВ шоу, сериалов и т. п.	Властелин колец, Мона Лиза, "Белые розы", Черный квадрат Малевича
Law		<input type="checkbox"/> Plural [default]		Нормативно-правовые акты: законы, статьи (без названий) и т. п.	Конституция, статья 20.6.1 КоАП, Закон о защите прав потребителей
		<input checked="" type="checkbox"/> Plural		Несколько Law в одной сущности	п. 2 статьи 228 и пп. 3 и 4 статьи 230 УК РФ

Таблица 1. Типы сущностей, их описания и примеры

Категория	Тип сущности	Атрибуты	Hotkey	Описание типа сущности	Пример
Date	Date: absolute		D	Явная дата (когда? какой день/месяц/год?), включая упоминания праздников	20 октября 2000 г., пятое августа, январь 2021 года, 2020, в марте, "в среду, 2 марта", после майских праздников
	Date: relative		Z	Относительная дата, включая конечные предлоги (когда относительно другой даты?)	три дня назад, прошлый год, по прошествии двух лет, на следующий день после, на 5-й день, сегодня, завтра, вчера
	Date: period		X	Период, измеряемый в днях, неделях, месяцах и т. д., включая упоминания праздников	с 5 января по 10 февраля, Мезозойская Эра, Эпоха Возрождения, за сентябрь, за последние шесть лет, на майские праздники, с 15 ноября до Нового года, холодное время года, лето, зима, весна, осень
	Date: duration			Продолжительность, измеряемая в днях, неделях, годах и т. д. (как долго?)	в течение двух недель, три года, пять месяцев и два дня, за день, полдня
Time	Time: absolute		T	Явное время, измеряемое в секундах, минутах, часах или сутках (когда? во сколько?)	15:30, полтретьего, пятнадцать минут одиннадцатого, три часа
	Time: relative			Относительное время, включая конечные предлоги (когда относительно другого времени?)	два часа назад, через минуту, за два часа до, по прошествии двух часов, в течение трех часов (напр., будет закончена работа, выполнен заказ и т.п.), на 56-й минуте, через сутки
	Time: period			Временной период, измеряемый в секундах, минутах, часах или сутках	с двух до пяти, с 12-ти до 23-х, за последние два с половиной часа, темное время суток, днем, ночью, вечером, утром
	Time: duration			Продолжительность времени, измеряемая в секундах, минутах, часах или сутках (как долго по времени?)	в течение трех часов, пару минут, три часа, за полчаса, за сутки
Numeric	Ordinal		3	Числа и слова для подсчета предметов (какой по счету?)	первый, второй, тысячный, следующий, предыдущий, (пред)последний, повторно, первичный, впервые, несколько
	Money		4	Денежная сумма, включая название валюты, в т. ч. "от ... до ..." (сколько денег?)	\$500, два евро, 100 рублей, один тенге, 300 динаров, 55 франков, 45 рублей 22 копейки, от 20 до 30 тыс. руб., от \$200 до \$300, более 2 млн рублей, почти \$100, около 50 евро
	Percent		5	Проценты (включая "%") и дроби, в т. ч. слово "половина"	десять процентов, полпроцента, 99.9%, 147%, на 30.1% меньше, половина, треть, две трети, 1/3, 25-я часть
	Age		6	Возраст человека или предмета	25 лет, 20-летний, годовалый, новорожденный
	Quantity		8	Измерения, количество чего-либо с единицей измерения (сколько?)	два кило, два стакана, 3 ч. л., 670 км, полтора метра, 6.4 Вт, метровый, более трех человек, около пяти видов, ни разу, один из, дважды, вчетверо реже, однажды (в значении "один раз"), пара кексов, тройка лошадей, единственный, топ-10
	Cardinal		0	Числа, не относящиеся к категориям выше (нельзя восстановить единицу измерения)	один, два, десять, миллион, 532, 12.331, 52 828, 5:0, пара (оценка), пятерка
Other term	<input type="checkbox"/> Plural <input type="checkbox"/> Fiction [default]			Термины, не подпадающие под категории выше, включая научные термины, болезни, технологии и т. д.	4G, гигабайт, рибонуклеотид, минорные актиноиды, премия "Оскар", COVID-19, коронавирус, инсульт, ишемия миокарда
	<input checked="" type="checkbox"/> Plural			Несколько разных терминов, объединенных в одну сущность	гепатит А и В, азотная и соляная кислоты
	<input checked="" type="checkbox"/> Fiction			Вымышленные термины	Татуин, Пандора, Скайнет, криптонит

2.3 Атрибуты типов сущностей

Некоторые сущности могут иметь один или несколько атрибутов из следующих:

- **Plural** – два или более разных объекта, объединенных в одну именованную сущность.
- **Fiction** – вымышленная сущность, как правило, относящаяся к литературе, кино, сериалам и др.
- **Unconscious** – имена, которыми люди называют неразумные объекты. Доступно только для типа сущности **Person**.
- **Department** – отдел внутри организации. Доступно только для типа сущности **Organization**.
- **Media** – любые медиа-источники: издательства, социальные сети, телевизионные каналы и т. п. Доступно только для типов сущности **Organization** и **Service**.
- **Citizenship** – гражданство какой-либо страны. Доступно только для типа сущности **Nationality**.
- **Project** – проекты, программы (государственные, научные, космические и т. п.). Доступно только для типа сущности **Product**.
- **Trademark** – бренд, торговая марка продукта (не в значении *Organization*). Доступно только для типа сущности **Product**.

The screenshot shows a web interface for configuring an entity. At the top, there is a 'Text' input field containing 'издание Daily Mail' and a 'Link' button. Below this is a 'Search' input field containing 'Google, Wikipedia'. The 'Entity type' section is a scrollable list with radio buttons for 'Person', 'Person property', 'Organization' (which is selected and highlighted in green), 'Contact', 'Event', 'Geo', and 'Facility'. At the bottom, the 'Entity attributes' section is highlighted with a red box and contains four checkboxes: 'Plural', 'Fiction', 'Department', and 'Media' (which is checked).

Рис. 16: Выбор атрибутов

По умолчанию ни один атрибут не выбран. Чтобы выбрать атрибут, после выбора типа сущности нажмите на один из атрибутов в разделе **Entity attributes**. В соответствующем поле появится флажок, и атрибут подсветится оранжевым (Рисунок 16).

2.4 Особенности Time и Date

Time – сущность, измеряемая в сутках, часах, минутах, секундах и т. д.

Date – сущность, измеряемая в днях, неделях, месяцах, годах и т. д.

Примечание. “Два дня” – Date, а “Двое суток” – Time.

2.4.1 Различия времен Absolute, Relative, Period, Duration

На Рисунке 17 наглядно изображены различия подтипов Date и Time.

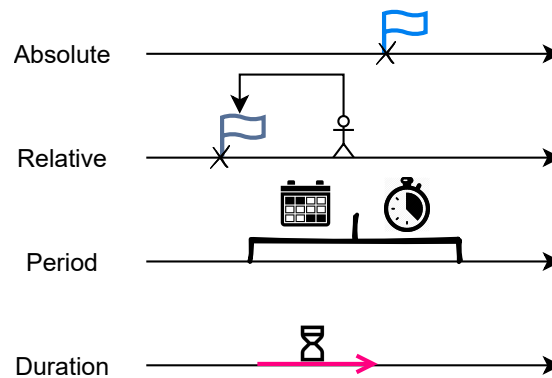


Рис. 17: Различия подтипов Time и Date

Absolute и **Relative** – всегда точечное время. **Period** – период времени, который можно ограничить началом и концом. **Duration** – измеряемая продолжительность времени.

2.4.2 Выражения с неявными временными рамками

Временные выражения размечаются, несмотря на отсутствие явно обозначенных временных рамок.

Время относительно другого времени (часто – времени написания текста) → **Date:Relative**: *сейчас, сегодня, завтра, вчера, в этом/прошлом/следующем году, два года назад, через два года*.

Дни недели (одиночные упоминания без даты) → **Date:Relative**, например: *понедельник, вторник, среда, четверг, пятница, суббота, воскресенье*.

Время суток → **Time:Period**, например: *утром, вечер, ночь, днём, темное время суток*.

Времена года → **Date:Period**, например: *лето, зима, весна, осень, холодное время года*.

Слова “сейчас, сегодня” и их синонимы “в настоящее время”, “в текущий момент”, “во время которого”, “в это время”, “в то же время”, “на данный момент”, “теперь” и т. п., “в начале/конце...”, “к началу/концу...” а также одиночные слова “ранее/раньше”, “позднее/позже”, “впоследствии”, “далее”, “затем”, “потом” → Date:Relative.

Промежуток времени, начинающийся со слов “в течение”, “во время”, “в период”, “в ходе”, “со времен”, “до”, “после”, и т. п. → Date:Period. Например: “в течение урока”, “в ходе интервью”, “во время приема пищи”, “со времен войны”, “в течение лета”, “до урока”, “после окончания университета”.

2.5 Некоторые правила вложенности

Сущности могут содержать в себе вложенные сущности, с которыми они связаны управлением¹ и их не разделяют предлоги или другие слова (см. “Вложенность” в Таблице 2).

Если сущности разделены предлогом или другими словами, они выделяются раздельно, без вложенности (см. “Раздельные сущности” в Таблице 2).

В Таблице 2 приведены различия цельных, вложенных и раздельных сущностей.

Таблица 2: Примеры цельных, вложенных и раздельных сущностей

Одна сущность	Вложенность	Раздельные сущности
[воронежский аэропорт]	[аэропорт [Воронежа]]	[аэропорт] в [Воронеже]
[немецкий канцлер]	[канцлер [Германии]]	[канцлер] в [Германии]
***	[Бурзянский район [Республики Башкортостан]]	[Бурзянский район] в [Республике Башкортостан]
***	[село Старосубхангулово [Бурзянского района [Республики Башкортостан]]]	[село Старосубхангулово] в [Бурзянском районе] в [Республике Башкортостан]
***	[ЦК [партии “Коммунисты России”]]	[ЦК] в [партии “Коммунисты России”]
***	[бой [Головкина] против [Альвареса]]	между [Головкиным] и [Альваресом] состоялся [бой]

2.5.1 Про Facility, Event и Service

Одиночные сущности, относящиеся к категориям Facility, Event и Service (например, “аэропорт” и “бой” в Таблице 2, и упоминания сервисов “сайт”, “форум” и т.п.), могут выделяться как именованная сущность, только если **в этом же абзаце в пределах трех предложений справа и слева, включая центральное**, есть уточнение (местоположение, год и т. п.), конкретизирующее эту

¹**Управление** – вид подчинительной связи, при которой зависимое слово употребляется в том косвенном падеже, которого требует главное слово. В нашем случае – это, как правило, связь *существительное+существительное* (а точнее, *именная группа+именная группа*), например “[президент [Российской Федерации]]”

одинокую сущность (“в Воронеже”, “между Головкиным и Альваресом” в Таблице 2, для сервисов – упоминание организации, которой принадлежит сервис).

Примечание. Это правило не действует на одиночные упоминания, если в предыдущем контексте есть их полное название (например, [Международный аэропорт Воронеж] или [сайт Всемирной организации здравоохранения]) или если в этом абзаце в пределах трех предложений справа и слева, включая центральное предложение, нет уточнений/конкретизаторов этого упоминания.

2.5.2 Про Date в Facility и Event

Дата может быть вложенной в сущности **Facility** и **Event**, только если она является частью названия. Сравните:

- (1) Олимпиада 1980 года
- (2) Бой Коннора и Пуарье_{Event} 24 января 2021 года_{Date:Absolute}

В первом примере дата является частью названия, такой сущность можно увидеть на афишах и в СМИ. Во втором случае, дата – добавочная информация.

2.5.3 Про Event в Date: Absolute и Period

Сущности **Date: Absolute** и **Period** могут включать в себя сущности **Event**, обозначающие праздники, например:

- после 9 мая_{Event} _{Date:Absolute}
- с 15 ноября до Нового года_{Event} _{Date:Period}

Примечание. Если происходящее событие или действие не имеет явно выраженных временных рамок, например “в период войны”, “во время приема пищи”, оно всё еще выделяется как **Date:Period** (Подробнее в разделе 2.4.2).

2.5.4 Про Ordinal в Time/Date:Relative

В **Relative** датах или времени типа на 9-й день или на 5-й минуте число размечается как вложенный **Ordinal**.

2.5.5 Про сочетания Family с Person и Person Property

В тексте будут встречаться разные комбинации сущности **Family** с **Person Property** и двумя вариантами **Person**². На Рисунках 19-20 изображена правильная разметка разных сочетаний с сущностью **Family**.

²Именительный и родительный падежи **Person** при именительном падеже **Family**. Сравните: “сын Иван” vs. “сын Ивана”.

При разметке нужно опираться на следующие правила:

1. **Family** – атомарная сущность, т.е. в нее входит только термин, обозначающий родство (см. примеры в Таблице 1 и на Рис. 19-20).
2. **Family + Person** (род. падеж) дополнительно объединяются в сущность **Person**, только если это выражение является подлежащим, действующим лицом в контексте (сравните варианты в примерах 1, 4, 5 на Рис. 19-20).
3. **Person** (им. падеж) может включать сущность **Family**, только если между ними есть **Person** и/или **Person Property** в родительном падеже, поясняющие сущность **Family** (сравните примеры 2 vs. 3, 6, 7 на Рис. 19-20).

В противном случае, **Person** (им. падеж) не включает в себя **Family** (пример 2 на Рис. 19-20).

Примечание. Конструкции с приложениями, как на Рисунке 18, не выделяются целиком как **Person**.

Рис. 18: Разметка **Family** в конструкции приложения.

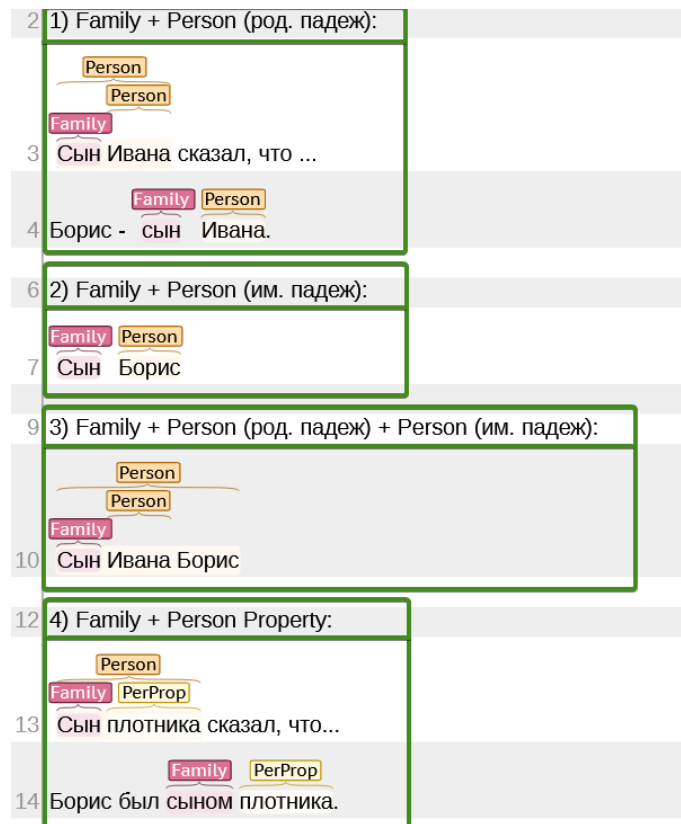


Рис. 19: Часть 1. Разметка сочетаний **Family** с **Person Property** и двумя вариантами **Person**

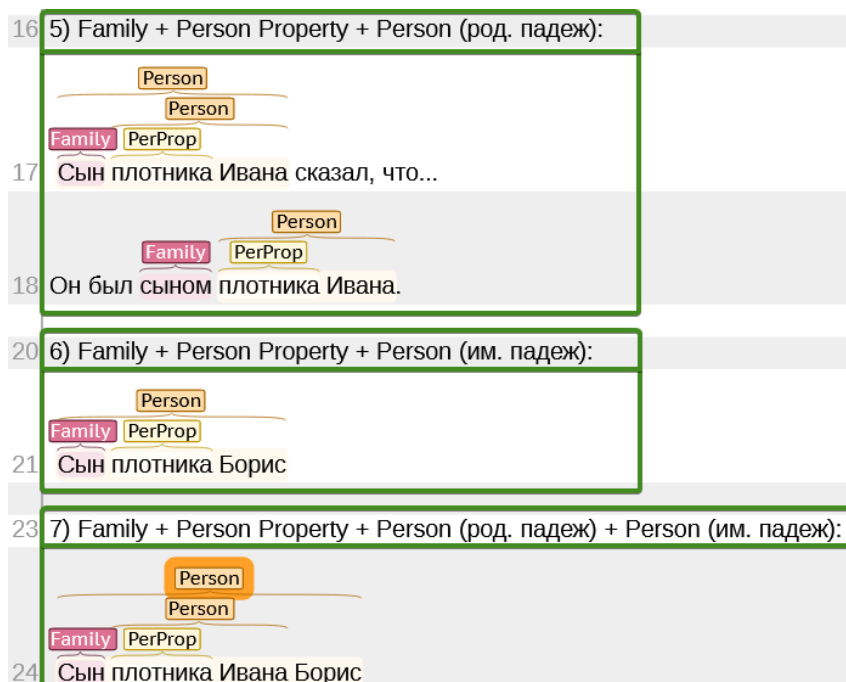


Рис. 20: Часть 2. Разметка сочетаний Family с Person Property и двумя вариантами Person

3 Разметка кореференции в brat

Все упоминания в `brat` отмечаются меткой `[X]`. На Рисунке 21 показаны выделенные упоминания в тексте.

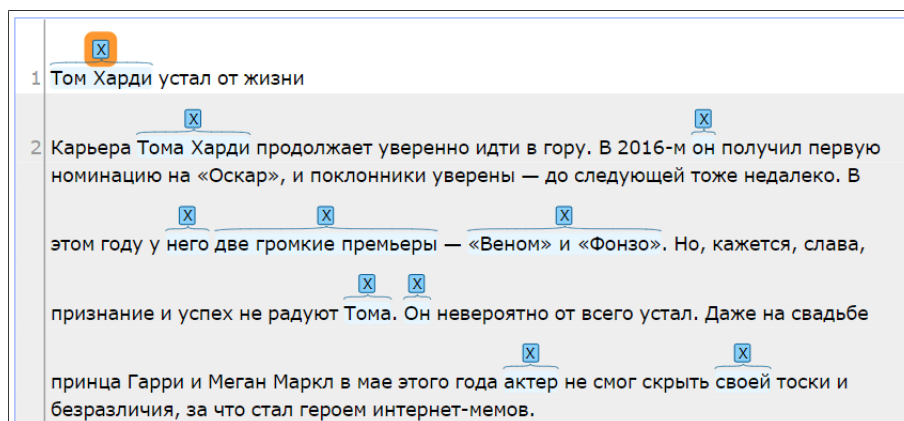


Рис. 21: Пример выделения упоминаний в `brat`

Все выделенные упоминания нужно попарно связать кореферентными связями. Всего в `ru:corner` четыре вида кореферентных связей: `ANAPH`, `CATAPH`, `IDENT` и `APPOS`. Подробную информацию об этих связях можно найти в Руководстве по разметке кореференции в `ru:corner`.

На Рисунке 22 показан пример текста с размеченными связями между упоминаниями.

Если в тексте нет референта, и все упоминания выражены местоимениями

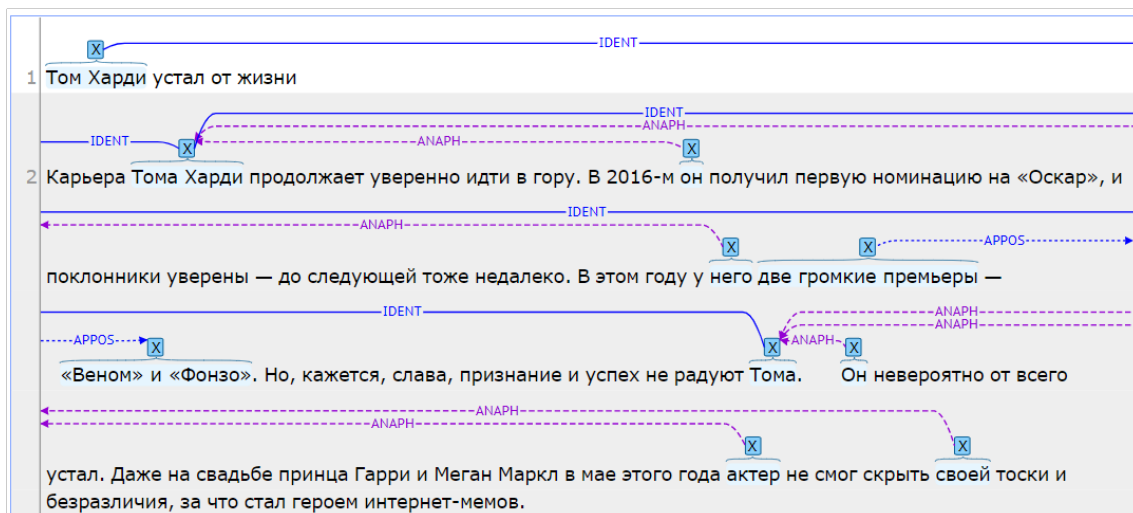


Рис. 22: Пример кореферентной разметки в brat

или неререферентными именными группами, следует отметить все упоминания такой кореферентной цепочки атрибутом **NoRef** (Рисунок 23).

Если у кореферентной группы указан референт в тексте, атрибут **NoRef** выбирать не нужно.

Рис. 23: Выбор атрибута **NoRef** для неререферентных упоминаний

Связь **IDENT** ненаправленная, связь **ANAPH** всегда направлена влево, связь **APPOS** всегда направлена вправо, а связь **APPOS** может быть направлена и влево, и вправо.

Для выбора другой метки связи, выберите её из списка.

Примечание. Связь **IDENT** ненаправленная, ее нельзя переименовать в одну из направленных связей.

Чтобы поменять направление связи между упоминаниями, дважды кликните на нужную стрелку и в появившемся окне выбора связи нажмите **Reverse**.

Чтобы изменить упоминание-цель (упоминание, на которое указывает стрел-

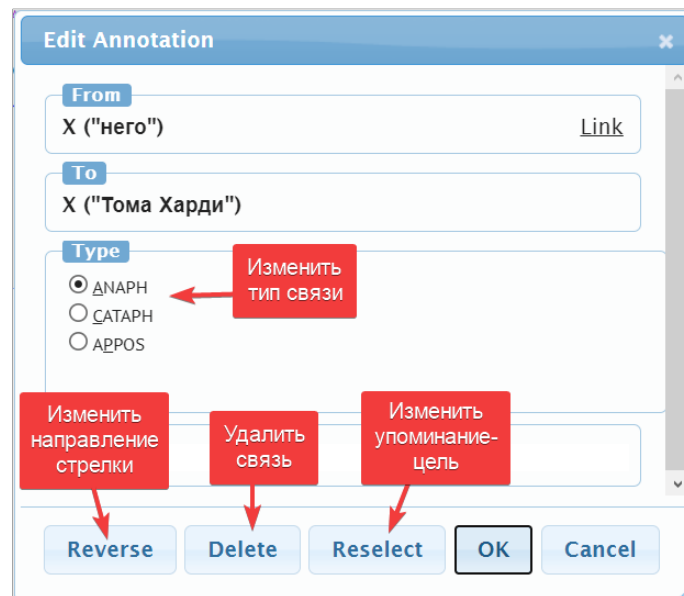


Рис. 24: Изменение или удаление связи между упоминаниями

ка), дважды кликните на стрелку, нажмите **Reselect** и переместите освободившуюся стрелку на новое слово.

Для удаления связи, дважды кликните на стрелку и нажмите **Delete** в окне выбора типа связей.

4 Отдельные случаи

В интерфейсе **brat** между всеми токенами³ для удобства разметки – три пробела. Каждое предложение отображается на новой строке, а между каждым абзацем есть пустая строка. Пример текста изображен на Рисунке 25.

1	Бывший муж Памелы Андерсон Джон Питерс снова помолвлен
3	Джон Питерс снова собрался жениться .
4	Об этом пишет издание US Weekly со ссылкой на свои источники .
5	А ведь прошло чуть меньше трех недель с момента расставания с Памелой Андерсон .
7	Избранницей продюсера стала некая Джулия Бернхейм .
8	О помолвке Питерс объявил во время заключительной церемонии бизнес-мероприятия NASDAQ в Нью-Йорке .
10	Напомним , Памела и Джон познакомились 30 лет назад в особняке основателя Playboy Хью Хефнера .

Рис. 25: Пример текста в интерфейсе **brat**

³Токены – отдельные слова и знаки пунктуации.

4.1 Уточняющие прилагательные

Прилагательные, уточняющие/конкретизирующие сущность (чаще всего они относятся к **Person Property**) входят в эту сущность, например:

российский премьер-министр, чеченский блогер, голливудский актер,
первый/бывший/будущий президент, питерский “Зенит”, младшая дочь

и т. п.

Обычные качественные прилагательные, типа *красивый, хороший, большой, дорогой* и т. п., которые дополнительно не конкретизируют объект, не входят в сущность.

4.2 Уточняющие именные группы

Уточняющие именные группы (существительное с зависимыми словами, если они есть) при именованных сущностях входят в эту сущность, например:

город Москва, радиостанция Эхо Москвы, издательство ЭКСМО,
американская авиакомпания Delta Air Lines, футбольный клуб “Зенит”,
река Волга и т. п.

4.3 Продукт, серия, уникальное название – как быть?

В тексте могут встретиться разные комбинации из упоминаний типа продукта, серии/производителя, уникального названия. В Таблице 3 приведены примеры, как следует выделять разные сочетания этих упоминаний.

Таблица 3: Разметка разных сочетаний Product и Person:Unconscious

Сочетание	Как выделять	Пример
тип продукта & серия/производитель	Product (производитель или торговая марка - Product:Trademark)	[самолеты Су-57] [космические корабли серии “Союз”] [автомобиль [Форд]]
тип продукта & уникальное название	Person:Unconscious	[броненосец “Потёмкин”] [американский космический корабль Орион]
тип продукта & серия/производитель & уникальное название	[[тип продукта & серия] название] → [[Product] Person:Unconscious]	[[Ледокол типа ЛК60Я] “Арктика”]
тип продукта & уникальное название & серия/производитель	[[тип продукта & название] серия] → [[Person:Unconscious] Product]	[[Ледокол “Арктика”] типа ЛК60Я]
уникальное название	целиком Person:Unconscious	[Потемкин], [Орион]
серия	Product	[Катюша], [Су-57], [Т34]

Примечание. **Person:Unconscious** – уникальное название только сущностей **Product**. Уникальные названия сущностей **Event** не выделяются как

Person:Unconscious. Следующие примеры выделяются целиком как **Event** без вложенных сущностей: “Ураган Катрина”, “Циклон Сара”, “Музыкальный фестиваль Юрмала-93” и т. п.

4.4 Соцсети: **Service** vs. **Service:Media** vs. **Org:Media**

При разметке упоминаний соцсетей, страниц/каналов/стен/аккаунтов в соцсетях следует опираться на следующие примеры:

- Акции Facebook_{Org} подскочили
- Об этом писали в социальной сети Facebook_{Service}
- Об этом писали в Facebook_{Service}
- Они написали об этом в своем Facebook_{Service:Media}
- Они написали обо этом на своей странице_{Service:Media} в Facebook_{Service}
- На канале Киркорова_{Service:Media} в YouTube_{Service} вышел новый клип

Примечание. Слова типа “страница”, “канал”, “аккаунт” размечаются как **Service:Media** только если в текущем абзаце есть уточнение, на какой платформе/сайте они находятся.

4.5 Предлоги

Начальные предлоги могут входить в ряд типов сущностей:

- Все типы сущностей групп **Date** и **Time**
- Все типы сущностей группы **Numeric**, кроме **Ordinal**

Примеры предлогов в составе сущностей изображены на Рисунке 26.

Примечание. В конец сущности могут входить уточняющие предлоги, например: “[в трёх километрах от]_{Quantity} деревни”, “[за два часа до]_{Time:Relative} начала”, “[через два дня после]_{Date:Relative} начала”.

4.6 Знаки препинания

Знаки препинания входят в сущность, только если являются частью сущности (запятые в адресах, дефисы в номерах телефонов, двоеточия и т. п.).

Конечные знаки препинания могут входить в сущность только если они являются частью названия объекта. Часто это восклицательные знаки, как, например, в *Мата Миа!* и *Кто боится Вирджинии Вульф?* (названия фильмов).

Запятые, точки, восклицательные, вопросительные и другие знаки препинания, не входящие в названия, не являются частью сущности (Рисунок 27).

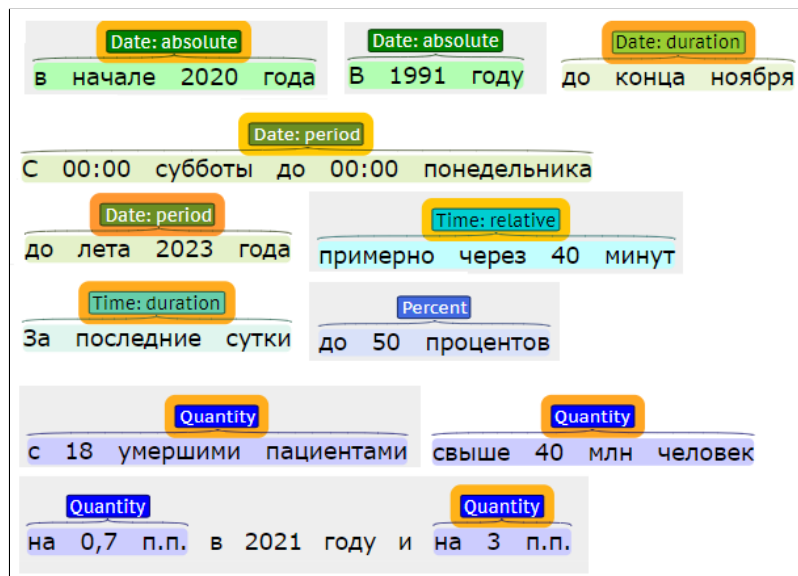


Рис. 26: Предлоги в составе сущностей

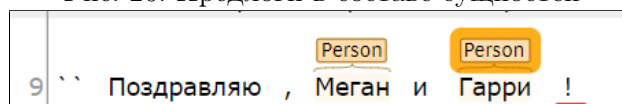


Рис. 27: Знаки препинания, как правило, не входят в сущность

4.7 Кавычки

В текстах начальные кавычки отображаются как ``, закрывающие кавычки - ".

Кавычки входят в сущность, если обособляют названия объектов (Рисунок 28). Кавычки, обозначающие начало/конец прямой речи или цитаты в сущности не входят (Рисунок 29).

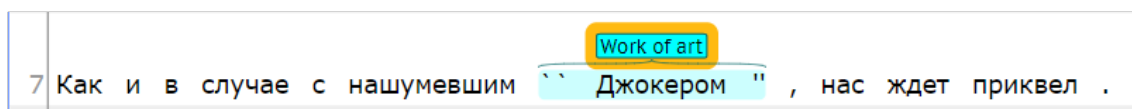


Рис. 28: Пример кавычек в названиях. Кавычки входят в сущность.

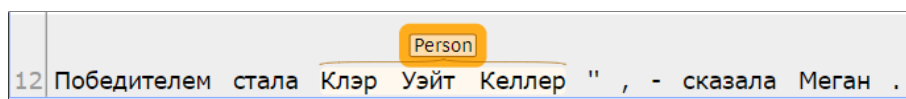


Рис. 29: Пример кавычек в прямой речи. Кавычки не входят в сущность.

4.8 Одиночные именованные сущности и анафора

Анафоричные упоминания именованных сущностей, не имеющие смысла без антецедента, не выделяются как именованная сущность.

Например: Белорусская оппозиция раскрыла план действий при отказе Лукашенко покинуть пост. Если требования [оппозиции] не будут выполнены, последует ряд массовых забастовок.

В примере выше только выражение “*Белорусская оппозиция*” выделяется как именованная сущность. Отдельно упомянутое слово “*оппозиция*” не является именованной сущностью, так как без антецедента нельзя установить, какая именно это оппозиция.

При разметке кореференции выражение “*оппозиция*” следует выделить как анафор.

Исключением из правила является выражения “*столица*”, “*граница*” и “*правительство*”. Если в тексте подразумевается территориальная граница и Правительство России, или в тексте есть уточнение, о столице какой страны идет речь (например, “американская/белорусская столица” или в абзаце указано название страны), “столица” размечается как GPE, “граница” – как Facility, а “правительство” – как Organization.

Кроме того, одиночные сущности Facility, Event и Service могут быть выделены как именованная сущность, если в рамках одного абзаца с ними есть конкретизирующее уточнение (Подробнее в Разделе 2.5).

Примечание. Одиночные слова, не указывающие на конкретный объект, не выделяются как сущность. К таким, например, относятся *полиция, армия, власти, маркетплейсы, соцсети, ТВ, радио*.

4.9 Numeric: Quantity vs. Cardinal в случае эллипсиса

Единичные числа могут отмечаться как Quantity, если в тексте объект счета пропущен, но может быть восстановлен из контекста. На Рисунке 30 выделенное выражение можно восстановить до “*умер 7071 человек*”.

зарегистрировано	450	436	новых заражений	, выздоровел	329	201	человек	умер	7071

Рис. 30: Пример единичного числа в роли Quantity

4.10 Выделение сущности, разбитой на две строки

Если в результате сегментации текста одна сущность оказалась на нескольких строках, необходимо сначала разметить часть сущности на первой строке, затем снова зайти в выбор типа сущности, нажать на Add Frag. и выделить оставшуюся часть сущности на второй строке (Рисунок 31).

После выделения фрагмента на второй строке, появится связь между двумя фрагментами (Рисунок 32).

Примечание. Если выделить разбитую на две строки сущность целиком, в интерфейсе появится ошибка (см. Раздел 4.11).

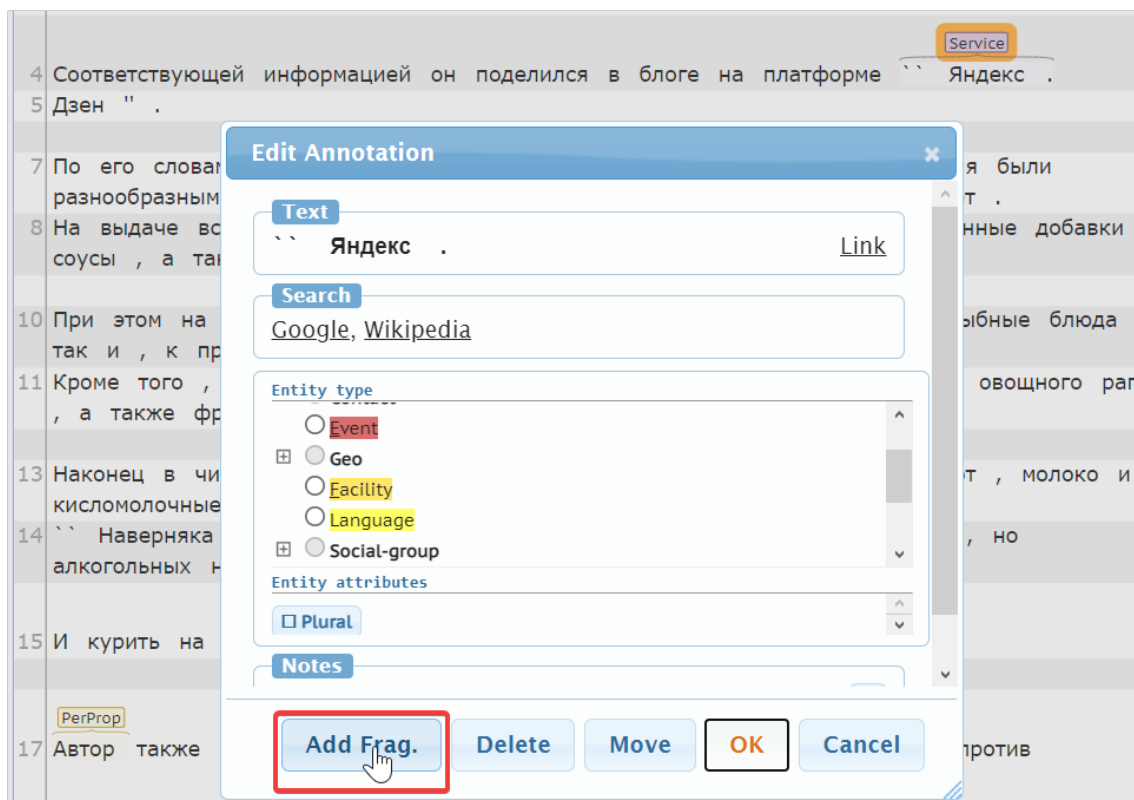


Рис. 31: Добавление фрагмента к сущности

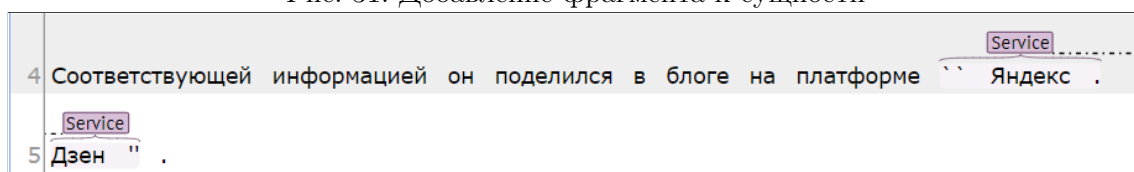


Рис. 32: Связь между фрагментами

4.11 Возможные ошибки и предупреждения

Примечание. Некоторые критичные ошибки возникают из-за разметки в браузерах типа Chrome, Opera и др. Убедитесь, что вы установили и используете **Mozilla Firefox** для разметки в **brat**.

В некоторых случаях могут появляться ошибки или сообщения с предупреждениями. Возможные случаи:

- **Пересекающиеся сущности.** Пересекаться с другими сущностями могут только сущности **Quantity**. В остальных случаях при пересечении метки подсвечиваются красным (Рисунок 33). Необходимо переразметить сущности таким образом, чтобы они не пересекались.

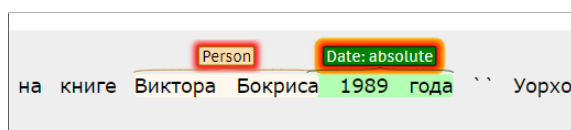


Рис. 33: Пример ошибки при пересечении сущностей.

Ошибки также возникают, если на пересечении вложенных сущностей оказываются лишние пробелы. Такие случаи тоже необходимо переразмечить.

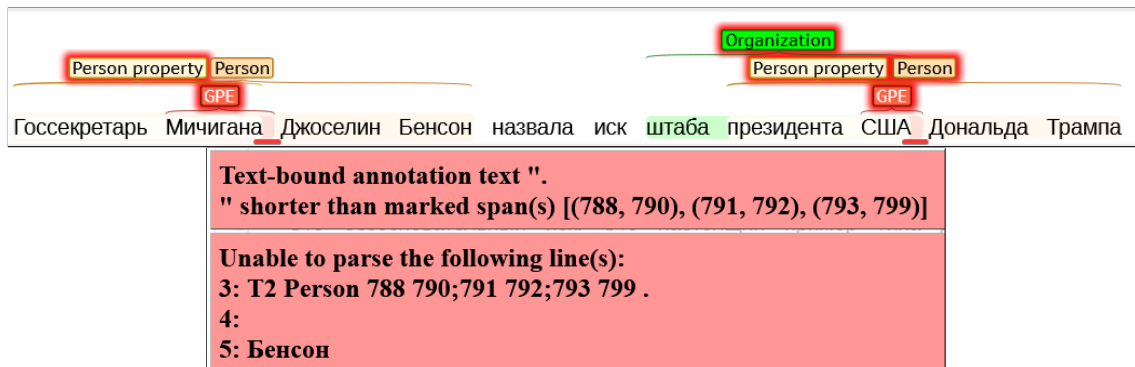


Рис. 34: Лишние пробелы на пересечении сущностей вызывают ошибки.

- **Пробелы в начале сущности.** Если в начало выделенной сущности попали пробелы, высветится предупреждающее сообщение (Рисунок 35). Чтобы предупреждение пропало, необходимо переразмечить сущность (с помощью **Move** или **Delete**), при выделении не включая лишние пробелы в начало.

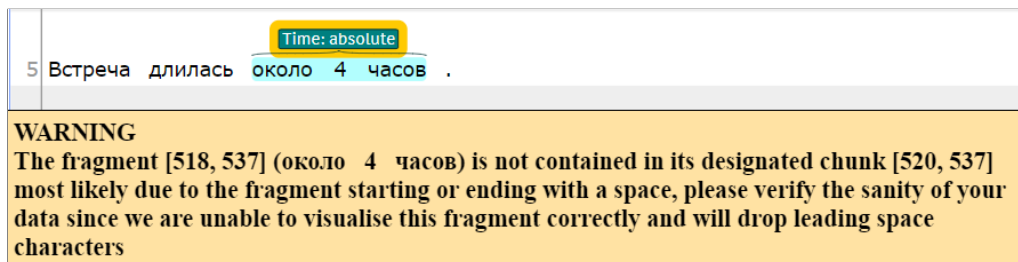


Рис. 35: Пример предупреждения при наличии пробелов в начале сущности.

Примечание. Лишние пробелы в конце сущностей будут удалены автоматически, они не вызывают сообщений с предупреждениями или ошибок.

- **Случайное выделение части второй строки при выделении сущности** вызовет серию ошибок, как изображено на Рисунке 36. Наиболее вероятная причина - разметка в нерекомендованном браузере. Удалить такую сущность самостоятельно и убрать ошибку не получится, обратитесь к администратору. В разделе 4.10 описано, как выделить сущность, разбитую на две строки, так, чтобы не возникало ошибок.

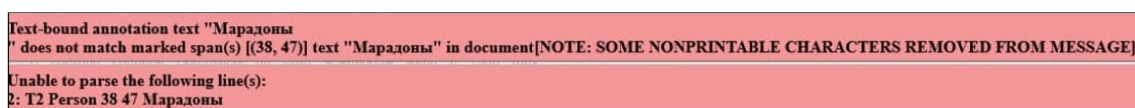


Рис. 36: Пример ошибки при выделении второй строки.

- **Ошибка сегментации.** Могут встречаться тексты, в которых цельные предложения разбиты на несколько строк (например, как на Рисунке

32). Если на границе строк одна сущность разбита на фрагменты, необходимо разметить ее, как описано в разделе 4.10. **Редко:** Если деление предложения не проходит по именованной сущности, в первой строке разбитого предложения отметьте любое слово, не являющееся сущностью, как `!!!INVALID DOCUMENT!!!` и напишите комментарий “Ошибка сегментации”. При этом остальной текст следует разметить как обычно.

A Changelog

3.02.2021

1. **Обновлена Таблица 1.** В примеры `Date:relative` добавлены “на 5-й день, сегодня/завтра/вчера”. В примеры `Time:relative` добавлено “на 56-й минуте”.
2. В разделе 26 добавлен пример “через два дня после”. Добавила подчеркивание конечных предлогов.
3. В разделе 2.4.2 в ряд примеров добавлены “завтра, вчера”.

4.02.2021

1. **Обновлена Таблица 1.** В примеры `Money` добавлены “от 20 до 30 тыс. руб., от \$200 до \$300”.

6.02.2021

1. В разделе 2.5 *Некоторые правила вложенности и разделения сущностей* пункты выделены в подразделы.
2. **Косметическая чистка текста**, в т.ч. теперь все названия сущностей в тексте, за некоторыми исключениями, выделены **таким шрифтом**.
3. **Добавлен подраздел 2.5.5** *Про сочетания Family с Person и Person Property*, в т.ч. Рисунок 19. + Добавила приложение и Рисунок 18.

7.02.2021

1. **Обновлена Таблица 1.** В сущности `Person property: Plural` отредактировано описание, примеры заменены на “главы России, Азейбарджана и Армении”. В `Cardinal` добавлен пример “5:0”.

9.02.2021

1. **Обновлена Таблица 1.** В примеры `Service` добавлена “скорая помощь”. В примеры `Person Property` добавлены “чиновник, специалист”.

13.02.2021

1. **Обновлена Таблица 1.** В примеры `Numeric` -> `Money` добавлены в примеры “более 2 млн рублей, почти \$100, около 50 евро”.
2. В разделе 4.3 добавлено **Примечание** о том, что `Person:Unconscious` относится только к `Product` и не относится к `Event`.

14.02.2021

1. **Обновлена Таблица 1.**
 - В примеры `Percent` добавлена “половина”.
 - В примеры `Ordinal` добавлены “повторно, первичный”.
 - Слово “сутки” перенесено из категории `Date` в `Time`. Обновлено описания и примеры.
2. **Добавлен новый подраздел 4.4, Coцcetu: Service vs. Service:Media vs. Org:Media**, с примерами и примечанием.
3. В раздел 2.4 добавлено примечание о разнице сущностей `Date` и `Time`.

15.02.2021

1. Обновлено Таблица 1.

- В примеры **Family** добавлены “сожителю, пара, молодожены”.
- В примеры **Quantity** добавлены “пара кексов, тройка лошадей”.
- В примеры **Cardinal** добавлены “пара (оценка), пятерка”. В описание добавлено “(нельзя восстановить единицу измерения)”.
- В примеры **Time:period** добавлены “темное время суток, днем, ночью, вечером, утром”.
- В примеры **Date:period** добавлено “холодное время года”.

16.02.2021

1. **Обновлена Таблица 1.** В примеры **Date:Period** добавлены “*лето, зима, весна, осень*”.
2. **Удален старый раздел 4.8** (перемещен в `deleted_sections.tex`), переписан раздел про время 2.4, добавлено деление на подразделы. В подразделе 2.5.3 добавлено примечание со ссылкой на новый подраздел 2.4.2, *Выражения с неявными временными рамками*.
3. **Добавлен раздел 2.5.4** *Про Ordinal в Time/Date:Relative*.
4. Косметическая чистка текста, разделение картинки с примерами про семью на две.

17.02.2021

1. **Обновлена Таблица 1.** В целом исправлены некоторые примеры и описания.
2. **Добавлен раздел 1.9** *Поиск по документу и коллекции*, в т.ч. Рисунок 12.
3. В Google-таблице удалены неактуальные листы, добавлен чеклист и паттерны регуляров для числительных.
4. Старые changelog убраны в изолированный файл. Уменьшен размер шрифта и межстрочный интервал для changelog.

18.02.2021

1. **Изменения в разделе 2.4.2** *Выражения с неявными временными рамками*:
 - явно прописано, что примеры типа “в этом/прошлом/следующем году, два года назад, через два года” и времена суток – **Date: relative**.
 - слова “сейчас, сегодня” и их синонимы, “в начале/конце ...”, а также “теперь, ранее, позднее, дальше, затем, потом” и др. примеры – **Date: relative**.
 - выражения, начинающиеся с “до, после” – **Date: period**.
2. Исправлены Таблицы в Google в соответствии с уточненной договоренностью по разметке времени (описана в разделе 2.4.2).

19.02.2021

1. В разделе 4.8 добавлено, что одиночное слово “столица” мб где, если рядом есть указание страны, в т. ч. “американская/германская и т.п. столица”

20.02.2021

1. **Обновлена Таблица 1.** Для **Date:relative** и **Date:period** добавлены Hotkey Z и X соответственно.