

**Технический регламент
конкурса отдельных заданий
«ПРО//ЧТЕНИЕ. Номинация: Грамматика_Eng»**

Содержание:

Общая информация.....	3
Термины и определения.....	3
Участники конкурса и условия участия.....	6
Циклы и этапы конкурса.....	6
Конкурсная задача и оценивание решений Участников.....	7
Отборочный этап.....	7
Этап финальных испытаний.....	8
Судейская коллегия и Техническая комиссия.....	9
Права Оргкомитета.....	10
Протесты и апелляции.....	10
Участникам запрещается.....	11
Обеспечение безопасности и дополнительные условия.....	11
Приложение 1. Метрики точности разметки.....	12
Приложение 2. Программные модули.....	14
1. Программные модули для парсинга, сравнения и оценивания разметок.....	14
2. Программный модуль обмена данными между ПКУ и СПР.....	14
3. Алгоритм вычисления критериев и итоговой оценки за эссе.....	15
Приложение 3 Классификатор ошибок и язык разметки данных.....	18
1. Язык разметки данных для конкурса.....	18
2. Описание разметки данных в формате JSON.....	24
3. Классификатор ошибок.....	26

1. Общая информация

- 1.1. Настоящий Технический регламент определяет требования к процедуре проведения конкурса отдельных заданий «ПРО//ЧТЕНИЕ. Номинация: Грамматика_Eng» в целях реализации Национальной технологической инициативы (далее Конкурс), обеспечению и другим техническим параметрам Конкурса.
- 1.2. Технический регламент публикуется на официальном Сайте Конкурса (далее - Сайте) <http://ai.upgreat.one/>
- 1.3. Настоящий Технический регламент является документом, детализирующим Конкурсное задание Конкурса, опубликованное на Сайте Конкурса, по регламентам проведения испытаний и процедуре определения результата и победителя Конкурса. Настоящий документ не противоречит Конкурсному заданию, но решает задачу детализации Конкурсного задания до технического уровня; технические детали, указанные в настоящем документе являются необходимым и достаточным описанием задачи Конкурса. Технический регламент определяет требования к процедуре проведения Испытаний Конкурса, Программным комплексам Участников, процедуре контроля и алгоритму определения уровня разработок команд.
- 1.4. Основная цель настоящего документа: обеспечение единого набора документации и регламента участия в Конкурсе, обеспечивающих прозрачность и объективность как для Участников, так и для сторонних наблюдателей.

2. Термины и определения

- 2.1. В настоящем Техническом регламенте помимо терминов, перечисленных ниже, используются в том числе, термины в значении, определенном в п. 1.10. Конкурсного задания Конкурса
- 2.2. **Алгоритм разметки** – алгоритм на основе искусственного интеллекта, используемый в Программном комплексе Участника для преобразования файла с неразмеченным текстом эссе в файл с размеченным текстом того же эссе.
- 2.3. **Алгоритмическая разметка** – разметка текста эссе, сгенерированная алгоритмом разметки в Программном комплексе Участника.
- 2.4. **БД, База данных эссе** – совокупность всех текстов эссе на английском языке, как размеченных, так и неразмеченных, представленных в виде файлов, организованная по определённым правилам.
- 2.5. **Выборка данных, дата-сет (ДС)** – набор текстовых файлов эссе, размеченных или неразмеченных.
- 2.6. **Выборка для обучения алгоритмов разметки, Обучающая выборка, Размеченный эталонный открытый ДС** – набор размеченных текстовых файлов эссе, предоставляемый Участникам Конкурса на равных условиях, используемый Участниками для обучения своих алгоритмов разметки.
- 2.7. **Выборка для тестирования алгоритма разметки, Тестовая выборка, Размеченный эталонный закрытый ДС** – набор неразмеченных текстовых

файлов эссе, предоставляемый Участникам Конкурса на равных условиях для независимого оценивания точности алгоритмов разметки в рамках Отборочного этапа. В Системе проверки решений используется выборка размеченных файлов тех же эссе, доступная только Оператору Конкурса и недоступная Участникам.

- 2.8. **Выборка для финального испытания, Финальная выборка, Финальный ДС** – набор неразмеченных текстовых файлов эссе, предоставляемый Участникам Конкурса на равных условиях в рамках этапа Финальных испытаний. Разметка финальной выборки экспертами осуществляется после того, как от всех Участников будут получены алгоритмические разметки всех файлов финальной выборки.
- 2.9. **Испытания, Финальные испытания** – ограниченный по времени период, начинающийся в назначенный Оргкомитетом день, когда Участники проводят с помощью разработанных ими Программных комплексов потоковую разметку текстов, содержащихся в файлах Выборки финального испытания.
- 2.10. **Конкурс** – открытый конкурс отдельного задания конкурс «ПРО//ЧТЕНИЕ. Номинация: Грамматика_Eng», направленный на решение научно-технологической проблемы в области искусственного интеллекта. Входит в Систему конкурсов поддержки команд.
- 2.11. **КЗ, Конкурсное задание** – основной документ, определяющий цели, задачи и порядок проведения Конкурса. Конкурсное задание утверждается конкурсной комиссией технологических конкурсов в целях реализации Национальной технологической инициативы.
- 2.12. **КО, Классификатор ошибок** – перечень типов ошибок и смысловых блоков, которые могут использоваться в экспертных и алгоритмических разметках текстов эссе. Классификатор ошибок и язык разметки описаны в Приложении 3 к данному Техническому регламенту.
- 2.13. **Лидерборд** — список Участников, ранжированный по убыванию критерия средней точности алгоритмической разметки (СТАР).
- 2.14. **ЛК, Личный кабинет** – раздел пользовательского веб-интерфейса Платформы, в котором пользователь может получить доступ к данным, аналитической информации и сервисам Платформы в соответствии с его ролью в Конкурсе и правами доступа.
- 2.15. **Оценка за эссе** – оценка в баллах, вычисляемая на основании разметки текста эссе согласно правилам проверки эссе данного типа (например, оценка за мини-сочинение по обществознанию в рамках ЕГЭ). Оценка может быть вычислена как по экспертной разметке, так и по алгоритмической.
- 2.16. **Ошибка** – фрагмент текста эссе, обладающий обязательными параметрами начала, конца и типа ошибки согласно КО, а также необязательными параметрами подтипа ошибки согласно КО, комментария (для детализации подтипа ошибки без привязки к тексту эссе), пояснения (для объяснения ошибки с привязкой к тексту эссе), исправления и тега.

- 2.17. **Параметры Конкурса** — числовые параметры, влияющие на значения автоматически вычисляемых критериев. Параметры Конкурса объявляются перед началом цикла и перед Финальными испытаниями, а также могут быть изменены решением Оргкомитета при согласовании с Судейской коллегией и Технической комиссией Конкурса.
- 2.18. **Платформа** – сайт в компьютерной сети Интернет, обеспечивающий доступ к данным, аналитической информации и сервисам для всех категорий пользователей, вовлечённых в участие в Конкурсе и проведение Конкурса. В частности, Участникам предоставляется возможность тестирования Алгоритмов разметки, Экспертам предоставляется возможность размечать тексты эссе в Программе Разметчике, и т.д. Платформа имеет интерфейс на русском и английском языках.
- 2.19. **Программа Разметчик** – программное обеспечение с вэб-интерфейсом на базе Платформы, которое позволяет проводить разметку текстовых файлов, выделяя и описывая ошибки и смысловые блоки согласно КО. Доступ к полной функциональности Разметчика имеют определённые Оргкомитетом категории пользователей. Участники имеют права на просмотр разметки, но не могут самостоятельно делать разметку в интерфейсе Разметчика.
- 2.20. **ПКУ, Программный комплекс Участника** – разработанный Командой Участника Конкурса комплекс программного и аппаратного обеспечения, который в автоматическом режиме принимает на входе неразмеченный файл эссе и выдаёт на выходе размеченный файл эссе, сгенерированный Алгоритмом разметки.
- 2.21. **ПСР, Программа сравнения решений** – программное обеспечение с вэб-интерфейсом на базе Платформы, позволяющее пользователю визуально сравнить два размеченных файла одного и того же эссе (например, алгоритмическую разметку с экспертной или две экспертные разметки от разных Экспертов).
- 2.22. **Размеченный файл, разметка файла** – текстовый файл эссе, размеченный в соответствии с правилами, описанными в Приложении 3 к данному Техническому регламенту, имеющий расширение *.txt и кодировку UTF-8 без спецификаций. Разметка может быть экспертной или алгоритмической.
- 2.23. **СПР, Система проверки решений** – автоматическая система проверки, которая принимает выборку файлов с алгоритмической разметкой, сравнивает их с соответствующей экспертной разметкой и выдаёт результат сравнения в форме аналитического отчёта. Результат проверки может быть визуально проконтролирован для любого файла из выборки в Программе сравнения решений.
- 2.24. **Смысловой блок** – фрагмент текста эссе, обладающий обязательными параметрами начала, конца и типа смыслового блока согласно КО. Смысловые блоки не являются ошибками, но их наличие может влиять на оценку за эссе.
- 2.25. **Участник, Участник Конкурса** — российское или иностранное юридическое или физическое лицо, или объединение таких лиц, чья заявка на участие в Конкурсе прошла одобрение Оргкомитетом. Все требования положений Конкурсного задания распространяются на Участника также, как и на Команду.

- 2.26. **Файл, файл с текстом эссе** – размеченный или не размеченный текстовый файл, имеющий расширение *.txt и кодировку UTF-8 без спецификаций.
- 2.27. **Цикл** – единый логический блок Конкурса, который состоит из этапов квалификации, финального испытания и подведения итогов.
- 2.28. **Эксперт, Специалист** – привлечённый профильный специалист общеобразовательного учреждения, который проверяет текст эссе и формирует в результате проверки размеченный файл.
- 2.29. **Экспертная разметка** – разметка текста эссе, сформированная экспертом в результате проверки вручную не размеченного файла данного эссе. В Платформе реализован специальный сервис для автоматизации процесса экспертной разметки.
- 2.30. **Эссе, текст, текст эссе** – оцениваемая письменная работа учащегося в виде развернутого и аргументированного текста, написанная по заданной проблеме или теме (например, сочинение, написанное учеником 11 класса общеобразовательной школы в рамках единого государственного экзамена по английскому языку).

3. Участники конкурса и условия участия

- 3.1. Для участия в конкурсе Участник подаёт заявку путем заполнения электронной формы на Сайте конкурса, согласно п.3 и Приложению 1 КЗ. Порядок формирования Команды Участника описан в п.4 КЗ.
- 3.2. Участие Команды в Конкурсе возможно только при условии участия этой Команды в технологическом конкурсе в целях реализации Национальной технологической инициативы «Взаимное обучение на английском языке»

4. Циклы и этапы конкурса

- 4.1. Конкурс проводится в один цикл и включает в себя Отборочный этап (п. 6 КЗ), Финальный этап (п. 7 КЗ) и Подведение итогов (п. 8 КЗ).
- 4.2. Использование общих программных модулей (доступ к которым предоставляется Команде в рамках технологического конкурса в целях реализации Национальной технологической инициативы «Взаимное обучение на английском языке») является обязательным, поскольку обеспечивает равные условия для всех Участников Конкурса. Участники Конкурса не имеют право самостоятельно вносить изменения в общие программные модули. В случаях обнаружения ошибок в них Участник может подать в Оргкомитет по электронной почте в свободной форме заявку «О внесении изменений в общие программные модули», описав в ней обнаруженные ошибки и/или свои предложения по улучшению работы модулей.
- 4.3. Отборочный этап Конкурса заключается в определении команд, соответствующих двум критериям:
- корректно заполненная заявка;
 - успешное прохождение Командой квалификационного этапа технологического конкурса в целях реализации Национальной технологической инициативы «Взаимное обучение на английском языке»

- 4.4. На этапе Финальных испытаний Участник может открыть и провести до конца только одну сессию обмена данными с Платформой.
- 4.5. Функции для участия в Финальных испытаниях должны быть реализованы в ПКУ с использованием общего программного модуля обмена данными с Платформой. Для облегчения данной технологической задачи Участникам предоставляется пилотный программный пакет с минимальным набором функций ПКУ.
- 4.6. Оргкомитет оставляет за собой право обновлять Обучающую и Тестовую выборки. Увеличение Обучающей и Тестовой выборок возможно в течение всего Конкурса. Корректировка Обучающей и Тестовой выборок возможна не позднее, чем за 10 дней до Финальных испытаний.
- 4.7. Оргкомитет оставляет за собой право обновлять исходные коды общих программных модулей, но не позднее, чем за 10 дней до Финальных испытаний. В случае обновления Оргкомитет оповещает всех Участников путём публикации на сайте Платформы, сообщая цель и суть произведённых модификаций.

5. Конкурсная задача и оценивание решений Участников

- 5.1. Предметом Конкурса является выбор наилучшего решения для автоматического выявления фрагментов академических эссе и сочинений, содержащих в себе грамматические и лексические ошибки в текстовых документах с помощью программных комплексов, разработанных участниками Конкурса. Испытания проводятся путём выявления ошибок в текстах в формате академического эссе.
- 5.2. Оценивание решений Команд производится по показателю средней точности алгоритмической разметки (СТАР) в соответствии с методикой, описанной в Приложении 1, Приложении 2, Приложении 3 настоящего регламента.
- 5.3. В рейтинговой таблице ранжирование команд происходит по показателю СТАР. В случае совпадения целочисленной части значения СТАР у нескольких команд, положение команды в рейтинге определяется следующими десятичными разрядами дробной части

6. Отборочный этап

- 6.1. Данный раздел дополняет п.6 КЗ.
- 6.2. В Отборочном этапе могут участвовать все Команды зарегистрированные в конкурсе в целях реализации Национальной технологической инициативы «Взаимное обучение на английском языке»
- 6.3. Положительные заключения о степени готовности ПКУ к проведению Финального испытания выдаются в случае, если:
 - командой предоставлена корректно заполненная заявка;
 - пройдена Квалификация конкурса в целях реализации Национальной технологической инициативы «Взаимное обучение на английском языке»

- 6.4. Результаты Отборочного этапа публикуются на сайте Конкурса совместно с результатами Квалификационного этапа конкурса в целях реализации Национальной технологической инициативы «Взаимное обучение на английском языке»

7. Этап финальных испытаний

- 7.1. Данный раздел дополняет п.7 КЗ.
- 7.2. В этапе Финальных испытаний могут участвовать все Участники, успешно прошедшие Отборочный этап.
- 7.3. Финальные испытания проводятся на Финальной выборке неразмеченных эссе (финальный ДС). Экспертная разметка этих эссе производится в течение 15 рабочих дней после завершения алгоритмической разметки Участниками Конкурса в целях реализации Национальной технологической инициативы «Взаимное обучение на английском языке». Таким образом, во время проведения Финальных испытаний экспертная разметка Финальной выборки не существует.
- 7.4. Для участия в Финальных испытаниях Участник должен запустить сессию обмена данными средствами Программного комплекса Участника до момента старта Финальных испытаний.
- 7.5. Во время Финальных испытаний неразмеченные файлы из Финальной выборки становятся доступны по расписанию, начиная с момента старта, через равные промежутки времени в 1 минуту для обеспечения равных условий всем Участникам Конкурса и исключения любых возможностей для организации ручной разметки данных. Расписание Финальных испытаний (дата-время старта и периодичность отдачи файлов) объявляется не менее, чем за 10 дней до их начала.
- 7.6. В ходе сессии обмена данными Программном комплексе Участника (ПКУ) должен оперативно загружать и обрабатывать неразмеченный файлы с сервера Платформы, затем отдавать файлы с алгоритмической разметкой в Систему проверки решений (СПР). При превышении порогового времени задержки в 1 минуту (после открытия общего доступа к данному файлу) файл аннулируется для данного Участника и не учитывается при подсчёте критерия СТАР.
- 7.7. Если доля аннулированных файлов Участника во время Финальных испытаний превышает 5% от объёма Финальной выборки, то решение (вся совокупность отданных файлов) данного Участника полностью аннулируется и не участвует в подведении итогов Конкурса.
- 7.8. Техническая информация о порядке организации сессии обмена данными с Платформой во время Финальных испытаний приведена в Приложении 2 к настоящему Техническому регламенту.
- 7.9. По окончании обработки файлов всеми Участниками Финального испытания все сессии обмена данными закрываются, все файлы с алгоритмической разметкой, полученные от Участников, сохраняются в Базу данных эссе (формируя закрытый ДС с финальными решениями Участников). С этого момента стартует процесс экспертной разметки Финальной выборки. Каждый файл эссе проверяется двумя

независимыми экспертами. В случае существенного расхождения их оценок назначается третий эксперт, который формирует третью разметку. Правила назначения третьей проверки зависят от типа эссе и приводятся в Приложении 3. Все эксперты осуществляют проверку в условиях ограниченного времени (время проверки может ограничиваться сверху в зависимости от типа эссе). Третьему эксперту предоставляются данные двух предыдущих проверок и возможность сравнить их с помощью ПСР. Экспертные разметки двух или трёх экспертов сохраняются в Базу Данных эссе (формируя Закрытый финальный ДС) и используются для вычисления критерия СТАР по Финальной выборке.

- 7.10. По окончании экспертной разметки Финальной выборки в течение не более чем 5 рабочих дней подводятся итоги Финального этапа. Всем Участникам предоставляется доступ к экспертным разметкам Финальной выборки.
- 7.11. По итогам Финального этапа строится финальный лидерборд — список Участников, ранжированный по убыванию критерия СТАР, рассчитанного по Финальной выборке. Аннулированные решения не участвуют в финальном лидерборде.
- 7.12. Порядок подведения итогов описан в п.8 КЗ.
- 7.13. В соответствии с п. 12.1.2 КЗ эссе размеченные ПКУ Участников во время Испытаний (но не само программное обеспечение, являющееся интеллектуальной собственностью команд) передаются Организаторам на условиях открытой лицензии Creative Commons Attribution 4.0 International (CC BY 4.0) для последующей публикации на Сайте конкурса с целью обеспечения прозрачности в определении победителей и призёров.
- 7.14. Порядок определения премий победителю и призёрам описан в п.10 КЗ.
- 7.15. Порядок перезапуска Испытаний:
 - 7.15.1. Испытания могут быть перезапущены в случае массового технического сбоя по решению Оргкомитета, на основании соответствующих протоколов Судейской Коллегии и Технической Комиссии.
 - 7.15.2. Перезапуск может быть назначен на следующий день, о чём сообщается отдельно на официальном сайте.
 - 7.15.3. Файлы на перезапуск новые. Уже отданные не учитываются.
 - 7.15.4. Результаты по отданным файлам не учитываются.

8. Судейская коллегия и Техническая комиссия

- 8.1. На Отборочном этапе:
 - 8.1.1. Судейская коллегия проверяет автоматически формируемые заключения о степени готовности ПКУ к проведению Финального испытания
 - 8.1.2. Результаты испытаний верифицируются Технической комиссией и передаются на утверждение Оргкомитету.
 - 8.1.3. Оргкомитет выносит окончательные решения о допуске Участников к Финальному испытанию.

- 8.2. На этапе Финальных испытаний
- 8.2.1. Техническая комиссия проверяет корректность полученных от Участников файлов, выявляет условия и причины аннулирований, если таковые имеются.
- 8.2.2. Техническая комиссия контролирует процесс распределения заданий для экспертной разметки Финальной выборки.
- 8.2.3. По окончании экспертной разметки Финальной выборки Судейская коллегия проверяет автоматически формируемые отчеты о решениях Участников
- 8.2.4. Результаты испытаний верифицируются Технической комиссией и передаются на утверждение Оргкомитету.
- 8.2.5. Оргкомитет выносит решения по итогам Финального испытания.

9. Права Оргкомитета

- 9.1. Оргкомитет оставляет за собой право обновлять Обучающую и Тестовую выборки. Увеличение Обучающей и Тестовой выборок возможно в течение всего Конкурса. Корректировка Обучающей и Тестовой выборок возможна не позднее, чем за 10 дней до Финальных испытаний.
- 9.2. Оргкомитет оставляет за собой право обновлять исходные коды общих программных модулей, но не позднее, чем за 10 дней до Финальных испытаний. В случае обновления Оргкомитет оповещает всех Участников путём публикации на сайте Платформы, сообщая цель и суть произведённых модификаций.
- 9.3. Оргкомитет имеет право проверить решение команды очно.
- 9.4. Оргкомитет имеет право запросить у команды дополнительные данные, обеспечивающие прозрачность и объективность судейства.
- 9.5. Оргкомитет имеет право аннулировать решения команды в случае нарушения настоящего Технического регламента и/или Конкурсного задания.

10. Протесты и апелляции

- 10.1. Если в ходе Финальных испытаний аннулировано более 5% файлов Участника, то данное решение Участника не может участвовать в подведении итогов Финального испытания без возможности подачи протеста или апелляции.

11. Участникам запрещается

- 11.1. Участнику запрещается осуществлять экспертную разметку Тестовой и Финальной выборки вручную, с помощью краудсорсинга или любыми иными способами с привлечением труда людей.
- 11.2. Запрещается привлекать труд людей для извлечения каких-либо дополнительных данных из текстов эссе в неразмеченных выборках – Тестовой и Финальной. Анализ и разметка этих текстов должны производиться Участником исключительно алгоритмически.
- 11.3. Запрещается самостоятельно модифицировать общие программные модули.

- 11.4. Запрещается на этапе Финальных испытаний отправлять файл с алгоритмической разметкой одного и того же эссе более 10 раз. Все последующие отправки будут проигнорированы, в качестве окончательного варианта будет рассматриваться последняя отправка.
- 11.5. Запрещается на этапе Финальных испытаний отправлять вместо алгоритмической разметки файла эссе какую-либо иную информацию, в том числе алгоритмическую разметку другого эссе. Такие отправки будут проигнорированы.
- 11.6. В случае нарушения перечисленных запретов Оргкомитет по запросу Судейской коллегии или Технической комиссии имеет право отстранить Команду от участия в Конкурсе

12. Обеспечение безопасности и дополнительные условия

- 12.1. Деятельность Участников в рамках Конкурса должна соответствовать действующим на территории Российской Федерации природоохранным нормам, требованиям техники безопасности.
- 12.2. В случае проведения очных мероприятий Оргкомитет Конкурса предоставляет Участникам правила по технике безопасности и охране окружающей среды, с которыми все члены команды Участника должны ознакомиться под подпись и соблюдать.
- 12.3. Все технические инструкции по работе Платформы, а также параметры и способы работы Платформы являются предметом отдельных документов, которые публикуются на официальном сайте Конкурсов, или на Платформе или ссылки на них ведут с Сайта или Платформы. Указанные документы не влияют на сутевую составляющую Конкурсного задания и Технического регламента, но описывают технические параметры функциональности Платформы и одинаково доступны всем Командам, в том числе для соблюдения равных условий участия Команд.

Приложение 1. Метрики точности разметки

Парная точность $M(X, Y)$ разметки X относительно разметки Y оценивается по метрикам $M_1(X, Y) \dots M_7(X, Y)$, которые затем усредняются с весами $w_1 \dots w_7$.

Метрики $M_1 \dots M_7$ измеряются в процентах от 0% до 100%. Чем больше, тем точнее совпадение. Итоговая метрика *парной точности* разметки X относительно разметки Y вычисляется как взвешенное среднее семи метрик:

$$M(X, Y) = \frac{\sum_{i=1}^7 w_i M_i(X, Y)}{\sum_{i=1}^7 w_i}$$

Весовые коэффициенты w_i позволяют задавать степень значимости каждой метрики. Возможно усреднение только по части метрик, тогда $w_i = 0$ для остальной части метрик.

В рамках конкурса отдельных заданий «ПРО//ЧТЕНИЕ. Номинация: Грамматика» и «ПРО//ЧТЕНИЕ. Номинация: Грамматика_Eng» весовые коэффициенты w_1 и w_7 принимаются равными нулю.

Набор весов $w_1 \dots w_7$ является параметром методики измерения точности разметки.

Для вычисления метрик $M_2 \dots M_6$ производится сопоставление фрагментов в двух разметках. Алгоритм сопоставления фрагментов описан в разделе 3 данного приложения.

Оптимистичная относительная парная точность алгоритмической разметки одного эссе, при сравнении со всем множеством $\{\mathcal{E}\}$ экспертных разметок данного эссе:

$$M_{\text{опт}}(A, \{\mathcal{E}\}) = \frac{\max_{\mathcal{E}} M(A, \mathcal{E})}{\min_{\mathcal{E}, \mathcal{E}'} M(\mathcal{E}, \mathcal{E}')} 100\%,$$

где максимум в числителе берётся по всем экспертным разметкам, минимум в знаменателе — по всем парам экспертных разметок данного эссе.

Средняя относительная парная точность алгоритмической разметки одного эссе, при сравнении со всем множеством $\{\mathcal{E}\}$ экспертных разметок данного эссе:

$$M_{\text{ср}}(A, \{\mathcal{E}\}) = \frac{\text{avr}_{\mathcal{E}} M(A, \mathcal{E})}{\text{avr}_{\mathcal{E}, \mathcal{E}'} M(\mathcal{E}, \mathcal{E}')} 100\%,$$

где среднее в числителе берётся по всем экспертным разметкам, среднее в знаменателе — по всем парам экспертных разметок данного эссе.

Усредняя числители по выборке эссе, получим итоговый показатель *Среднюю Точность Алгоритмической Разметки*, с учётом параметра жёсткости H :

$$\text{СТАР} = \text{avr}_{\text{эссе}} \left(H \text{avr}_{\mathcal{E}} M(A, \mathcal{E}) + (1 - H) \max_{\mathcal{E}} M(A, \mathcal{E}) \right)$$

M1. Точность предсказания оценки за эссе.

Метрика не используется

M2. Точность и полнота поиска фрагментов.

Точность поиска определяется как доля фрагментов разметки X , имеющих сопоставленный фрагмент в разметке Y .

Полнота поиска определяется как доля фрагментов разметки Y , имеющих сопоставленный фрагмент в разметке X .

Агрегированная метрика *точности и полноты* определяется как их гармоническое среднее (F_1 -мера).

M3. Точность предсказания кодов.

Умноженная на $M2$ доля фрагментов разметки X , имеющих сопоставленный фрагмент в разметке Y с равным кодом фрагмента (типом ошибки или смыслового блока).

M4. Точность предсказания подтипов ошибок и комментариев.

Умноженная на $M2$ доля фрагментов разметки X , имеющих сопоставленный фрагмент в разметке Y с равным подтипом ошибки или с комментарием, который является парафразом.

M5. Точность локализации фрагментов.

Умноженная на $M2$ средняя точность совпадения фрагментов разметки X с сопоставленными им фрагментами разметки Y . Точность совпадения двух фрагментов вычисляется как мерой Жаккара – отношение числа слов в пересечении к числу слов в объединении двух фрагментов. Если фрагмент разметки X не имеет сопоставленного ему фрагмента разметки Y , то точность их совпадения принимается равной 0%.

M6. Точность исправлений ошибок.

Умноженная на $M2$ доля фрагментов разметки X , имеющих исправление и сопоставленный фрагмент в разметке Y с таким же исправлением.

M7. Точность пояснений.

Метрика не используется

Приложение 2. Программные модули

1. Программные модули для парсинга, сравнения и оценивания разметок

Следующие алгоритмы предоставляются Участникам Конкурса для обеспечения равных условий. Они должны быть разработаны и реализованы до начала Конкурса.

1. Алгоритм парсинга (синтаксического разбора) разметки:

Вход:

размеченный текст;

Выход:

список значений полей метаописания и список фрагментов;

для каждого фрагмента:

позиции начала и конца в исходном неразмеченном тексте,

текст фрагмента,

код(ы),

комментарий,

пояснение,

исправление,

тег.

2. Алгоритм поиска оптимального соответствия между фрагментами:

Вход:

две разметки одного и того же текста в виде двух списков фрагментов;

Выход:

список пар номеров соответствующих фрагментов из первого и второго списка.

3. Алгоритм вычисления критериев и итоговой оценки за эссе:

Вход:

список значений полей метаописания;

разметка в виде списка фрагментов;

Выход:

список значений критериев в баллах;

итоговая оценка согласно правилам оценивания данного типа эссе.

Формулы, по которым вычисляются критерии и итоговая оценка за эссе в баллах, зависят от типа эссе. В Приложении 3 к Техническому регламенту формулы оценки в баллах приведены для следующих типов эссе - сочинений ЕГЭ по английскому языку.

4. Алгоритм вычисления метрики парной точности разметки:

Вход:

две разметки одного и того же текста в виде двух списков фрагментов;

Выход:

значение парной точности разметки;

значения метрик M1–M7 (в рамках Конкурса учитываются только метрики M2-M6).

2. Программный модуль обмена данными между ПКУ и СПР

Программный модуль обмена данными между Программным комплексом Участника (ПКУ) и Системой проверки решений (СПР) предоставляется всем Участникам, подавшим заявку на Конкурс.

Основным назначением модуля является обеспечение синхронной раздачи неразмеченных файлов Участникам Конкурса и последующего приема размеченных файлов от Участников.

Модуль позволяет реализовать несколько сценариев коммуникации между ПКУ и СПР. В ходе сессии обмена данными ПКУ выступает в роли клиента, СПР – в роли сервера.

Сценарий сессии обмена данными на Отборочном этапе

- ПКУ: Запрос на открытие сессии.
- СПР: Проверка Участника и, в случае успеха, передача разрешения на открытие сессии.
- ПКУ: Запрос на получение следующего файла неразмеченного эссе.
- СПР: Передача файла неразмеченного эссе или сообщения о завершении выборки.
- ПКУ: Отправка размеченного файла того же эссе.
- СПР: Замер времени отклика, сохранение разметки, оценивание качества разметки.
- СПР: По завершении выборки генерация и сохранение аналитического отчёта.

Сценарий сессии обмена данными на этапе Финальных испытаний

- ПКУ: Запрос на открытие сессии (должен поступить не позднее 2 минут после старта Финальных испытаний).
- СПР: Проверка Участника и, в случае успеха, передача разрешения на открытие сессии.
- СПР: Передача файла неразмеченного эссе или сообщения о завершении выборки.
- ПКУ: Отправка размеченного файла того же эссе (должна произойти не позднее 60 секунд после отправки неразмеченного файла).
- СПР: Замер времени отклика, сохранение разметки, оценивание качества разметки. В случае, если разница между отправкой неразмеченного и получением размеченного файлов СПР превосходит 60 секунд, то данный файл не учитывается для определения результата команд в соответствии с: а) п.3.1.4. Конкурсного задания, ограничивающего время на обработку эссе в 30 секунд, б) определяемым Техническим регламентом временным интервалом на сетевые транзакции и задержки в передаче пакетов в размере 30 секунд.
- СПР: По завершении выборки генерация и сохранение аналитического отчёта.

Если СПР не получает от ПКУ запрос на получение файла неразмеченного эссе в течение отведённого времени или если СПР не получает от ПКУ размеченный файл в течение отведённого времени, то данный файл аннулируется для данного Участника и не будет учитываться при подсчёте усреднённых критериев СТАР и ОТАР.

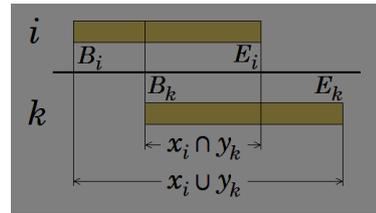
3. Алгоритм вычисления критериев и итоговой оценки за эссе

Разметкой будем называть последовательность n фрагментов $X = \{x_i = (B_i, E_i, C_i)\}_{i=1}^n$, где B_i – позиция начала фрагмента, E_i – позиция конца фрагмента, C_i – тип фрагмента.

Соответствием двух разметок $X =$ и $Y = \{y_k = (B_k, E_k, C_k)\}_{k=1}^m$ будем называть множество M пар фрагментов (i, k) , такое, что каждому x_i из X соответствует не более одного y_k и каждому y_k из Y соответствует не более одного x_i . Если для фрагмента x_i нет соответствия, будем записывать « $x_i \rightarrow \emptyset$ ».

Для произвольной пары фрагментов (x_i, y_k) определим *расстояние Жаккара*:

$$J_{ik} = 1 - \frac{|x_i \cap y_k|}{|x_i \cup y_k|}$$



Расстояние Жаккара J_{ik} принимает значения от 0 до 1. Если фрагменты x_i и y_k совпадают, то $J_{ik} = 0$. Если фрагменты не пересекаются, то $J_{ik} = 1$.

Определим *матрицу потерь* $L[i, k]$ размера $n \times m$:

$$L[i, k] = J_{ik} + [J_{ik} = 1] + [B_i \neq B_k] + [C_i \neq C_k].$$

$L[i, k]$

Примеры ситуаций соответствия пары фрагментов (x_i, y_k)

- | | | |
|------|---|--|
| 0 | фрагменты совпадают и имеют одинаковый тип | |
| 0..1 | фрагменты пересекаются, имеют общее начало и одинаковый тип | |
| 1 | фрагменты совпадают, но имеют разные типы | |
| 1..2 | фрагменты пересекаются и имеют либо разное начало, либо разные типы | |
| 2 | <i>фрагменты не образуют соответствия</i> | |
| 2..3 | фрагменты пересекаются, имеют разное начало и разные типы | |
| 3 | фрагменты не пересекаются и имеют одинаковый тип | |
| 4 | фрагменты не пересекаются и имеют разные типы | |

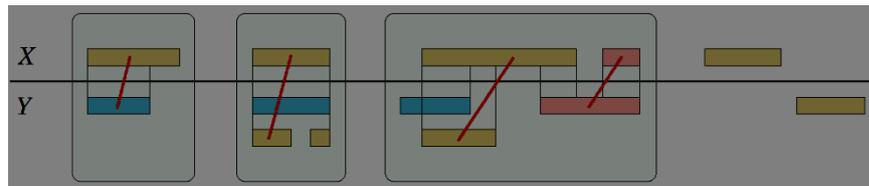
Требуется найти соответствие между разметками, минимизирующее сумму потерь:

$$Q(M) = \sum_{(i,k) \in M} L[i, k] + \sum_i [x_i \rightarrow \emptyset] + \sum_k [y_k \rightarrow \emptyset] \rightarrow \min_M$$

Задача поиска оптимального соответствия между разметками является обобщённой задачей о назначениях, для решения которой может быть применён один из вариантов венгерского алгоритма. Однако в нашем случае функция потерь устроена настолько специфически, что существует быстрый переборный алгоритм поиска оптимального соответствия.

Алгоритм проще описывать в терминах теории графов. Задан *двудольный граф* G , долями которого являются множества фрагментов X и Y , рёбрами соединяются пересекающиеся фрагменты (i, k) : $J_{ik} < 1$. Если пара фрагментов не пересекается, то она не может входить

в оптимальное решение, поскольку данную пару выгоднее вообще не связывать. Таким образом, задача минимизации суммарных потерь является задачей поиска оптимального паросочетания – подграфа M попарно несмежных рёбер в двудольном графе G .



На рисунке показан пример соответствия двух разметок: верхняя разметка X состоит из 5 фрагментов, нижняя Y состоит из 8 фрагментов. Оптимальное паросочетание состоит из 4 рёбер; несвязанными остаются 1 фрагмент из верхней разметки и 4 из нижней.

Задача поиска оптимального паросочетания решается по отдельности для каждой компоненты связности графа, что сильно сокращает перебор. Внутри каждой компоненты производится полный перебор паросочетаний. При переносе ребра из исходного графа в паросочетание компонента может распасться на ещё меньшие компоненты связности, которые в свою очередь можно обработать по отдельности, ещё сильнее сократив перебор. Эта идея реализуется рекурсивным алгоритмом полного перебора всех паросочетаний в связных компонентах заданного двудольного графа. Для ускорения перебора можно перебирать рёбра в порядке убывания количества смежных с ними рёбер.

Алгоритм:

1. задан двудольный граф G ; паросочетание M пусто; $Q(M) = n + m$;
2. перенести из G в M все рёбра (i, k) , для которых $L[i, k] = 0$;
3. $M_{min} := M$; $Q_{min} := Q(M)$;
4. перебрать_паросочетания (G, M);
5. вернуть наилучшее паросочетание M_{min} ;

Функция перебрать_паросочетания (G, M):

1. разбить граф G на связные компоненты;
2. для каждой связной компоненты G' из G :
 3. перебирать рёбра (i, k) , несмежные с M , по убыванию числа смежных ребер;
 4. перенести ребро (i, k) из G' в M ;
 5. вычислить $Q(M)$;
 6. если $Q(M) < Q_{min}$ то запомнить $M_{min} := M$; $Q_{min} := Q(M)$;
 7. перебрать_паросочетания (G', M);
 8. вернуть ребро (i, k) из G' в M ;

Приложение 3 Классификатор ошибок и язык разметки данных

В данном документе описана методика разметки текстов академических эссе для формирования обучающих и проверочных выборок данных для проведения испытаний по отдельному заданию ПРО//ЧТЕНИЕ. «Номинация: Грамматика_Eng»

Методика основана на формализации критериев проверки сочинений по английскому языку в рамках единого государственного экзамена.

Методика и классификатор ошибок основаны на материалах ЕГЭ 2020 года.

1. Язык разметки данных для конкурса

Основной задачей Конкурса является разработка моделей и алгоритмов понимания естественного языка (natural language understanding, NLU), способных к поиску и исправлению грамматических и речевых ошибок. Для создания таких алгоритмов необходимо иметь большую выборку текстов, в которых ошибки размечены экспертами.

Целью формирования размеченного корпуса академических эссе является обеспечение исходными данными как самого Конкурса, так и последующих исследований в области понимания естественного языка.

В данном документе описаны *правила разметки и типы ошибок* в сочинениях по английскому языку проверяемых в рамках единого государственного экзамена.

1.1 Общие принципы разметки

Следующие общие соображения были учтены при разработке правил разметки.

1. Основной целью разметки является указание *локализации и типа* каждой ошибки в тексте эссе, удобное как для человека, так и для машины.
2. Каждому типу эссе соответствует свой *классификатор типов ошибок*. Для грамматических и речевых ошибок используются отдельные классификаторы.
3. Фрагмент, локализирующий ошибку, должен быть непрерывным, достаточным для объяснения типа и сути ошибки, и безыбыточным, то есть он не должен захватывать избыточные участки текста, не имеющие отношения к данной ошибке.
4. При оценивании некоторых типов эссе учитываются не только ошибки, но также наличие определённых *смысловых блоков*. В частности, в сочинениях по русскому языку, обществознанию и истории должны приводиться примеры, причём количество корректно приведённых примеров влияет на оценку.
5. Разметка текста должна быть достаточной для автоматического выставления оценки согласно формальным критериям проверки эссе данного типа (за исключением орфографических и пунктуационных ошибок, если они не приводят к искажению смысла высказываний). Если эксперт снимает балл по некоторому критерию, то разметка должна содержать полную информацию о том, на каком основании был снят этот балл. Это либо ошибка, либо отсутствие смысловых блоков, обязательных для эссе данного типа.
6. Разметка может содержать исправления *непрофильных ошибок*, которые не учитываются при оценивании эссе данного типа, но могут затруднять автоматическую

обработку текста. Например, в сочинениях по истории и обществознанию опечатки, орфографические, пунктуационные, грамматические и речевые ошибки являются непрофильными.

7. Эксперты не должны вносить в тексты никаких собственных исправлений или комментариев, за исключением предусмотренных правилами разметки.
8. Синтаксис разметки не должен оставлять возможностей программе-парсеру спутать элементы разметки с исходным текстом эссе. Это обеспечивается выбором спецсимволов для выделения фрагментов, типов ошибок и комментариев эксперта.

1.2 Разметка фрагментов

Основной синтаксической конструкцией *языка разметки* является выделение фрагмента текста с помощью специальных скобок «(\ ... \)» или «(* ... *)», между которыми могут использоваться спецсимволы «\», «::», «>>», «#» для выделения типа ошибки и комментариев эксперта.

Пример ошибки «нарушение норм управления» с исправлением из сочинения по русскому языку:

до разметки:

Все удивлялись его силой.

после разметки:

Все удивлялись его (* Г.упр \ силой >> силе *).

Замечания о синтаксисе языка разметки. Спецсимволы «\», «*» и «(» выбраны потому, что их ввод не требует переключения регистра или языка на клавиатуре. Спецсимвол «\» или «*» рядом со скобками ставится для того, чтобы разметку невозможно было перепутать с авторскими скобками в исходном тексте. Косая черта используется именно обратная «\», а не прямая «/», поскольку прямая черта иногда встречается в текстах, например: «и/или», «д/р». Два типа скобок «(\ ... \)» и «(* ... *)» нужны для того, чтобы в случае вложенных фрагментов разметчику было удобнее отличать внешние фрагменты от внутренних. Открывающая и закрывающая скобки должны быть одного типа, то есть нельзя начать фрагмент скобкой «(\», а закрыть скобкой «*)».

После открывающей скобки «(\» или «(*)» записывается одно или два ключевых слов (кодов), обозначающих тип ошибки или смыслового блока и необязательный подтип ошибки.

Пример выделения смыслового блока «понятие» в мини-сочинении по обществознанию:

до разметки:

Познание - процесс постижения человеком окружающей действительности.

после разметки:

(*ПОНЯТИЕ\ Познание - процесс постижения человеком окружающей действительности.*)

Аналогичным образом выделяются ошибки. Пример ошибки типа «**О.теорсвязь**»:

(\ **О.теорсвязь**\ Деятельность – это процесс целенаправленной активности людей.
\)

Типы ошибок жёстко зафиксированы в классификаторе и связаны с формальными критериями оценивания. Ошибку можно детализировать с помощью кода подтипа ошибки, который записывается после кода ошибки через пробел, либо с помощью *комментария*, который записывается после выделенного фрагмента через косую черту «\»:

(\ **О.теорсвязь идея** \ Деятельность – это процесс целенаправленной активности людей. \)

(\ **О.теорсвязь** \ Деятельность – это процесс целенаправленной активности людей.

\Понятие не связано с основной идеей. \)

Подтип ошибки или комментарий кратко объясняет учащемуся суть ошибки. Для типа ошибки в классификаторе может быть предусмотрено несколько подтипов. Каждому подтипу соответствует свой стандартный комментарий (то есть *подтип ошибки* – это, по сути, аббревиатура для стандартного комментария). Если эксперт считает, что ни один из стандартных комментариев не подходит для данного случая, то он может записать свой комментарий. Текст комментария должен быть лаконичным и называть типовую ошибку, встречающуюся во многих работах. Комментарий не должен обращаться к тексту данной работы. Если в разметке указан и подтип ошибки, и комментарий, то комментарий имеет приоритет и подтип должен быть проигнорирован.

Если эксперт считает, что комментария недостаточно для объяснения сути ошибки автору эссе, то он может добавить развёрнутое пояснение, отделив его двумя двоеточиями «::». Также он может добавить свой вариант исправления, отделив его стрелкой «>>», составленной из двух символов «больше». В отличие от комментария, пояснение должно обращаться к тексту данной работы и конкретно к выделенному фрагменту. Пример:

(\ **О.теорсвязь идея**\ Деятельность – это процесс целенаправленной активности людей.

:: Здесь следовало бы рассмотреть особенности процесса познания. Более общее понятие деятельности уводит в сторону от основной идеи высказывания. \)

Для разметки ошибки, не профильной для данного типа эссе, используется предопределённый код «ИСП». Исправление, отделяемое стрелкой «>>», указывается обязательно. Комментарий, пояснение и тег, как правило, опускаются. Пример:

Все удивлялись его (\ИСП\ силой >> силе\).

Выделяемые фрагменты могут быть вложенными. Наиболее типичный случай – когда выделяется длинный смысловой блок, внутри которого могут выделяться как ошибки, так и другие смысловые блоки. Пример:

(* **ПОНЯТИЕ** \ (\ **О.теорсвязь** \ Деятельность – это процесс целенаправленной активности людей. \ Понятие не связано с основной идеей. \) *)

В таких случаях надо внимательно следить за балансом скобок: сколько было открывающих скобок «(\», столько же должно быть и закрывающих «\)». Выделяемые скобками фрагменты могут вкладываться друг в друга, но не могут перекрываться. Для визуального выделения вложенных фрагментов можно использовать разные скобки: «(*... (\ ... \) ... *)», однако эта рекомендация не является обязательным требованием.

Некоторые типы ошибок относятся не к определённому фрагменту, а ко всему тексту. В таком случае ошибка может быть указана в конце всего текста без выделения фрагмента. Примеры:

(* С.тема *)

(* С.тема \:: Тема осталась нераскрытой, для раскрытия темы следовало ... *)

1.3 Разметка фрагментов: формальное определение

Фрагмент с ошибкой или смысловой блок выделяется скобками «(\ ... \)» или «(* ... *)». Внутри скобок перед фрагментом указывается код или несколько кодов, разделяемых пробелами. После фрагмента могут идти комментарий, пояснение, исправление и тег, каждый из которых начинается своим стартовым спецсимволом:

(* Коды \ ТекстФрагмента \ Комментарий :: Пояснение >> Исправление #
Тег *)

Условные обозначения:

- **Коды** – одно или несколько ключевых слов (аббревиатур), обозначающих *тип смыслового блока, тип ошибки* или *подтип ошибки* по классификатору; коды ошибок имеют вид «х.zzz», где «х» – буква, указывающая, к какому типу эссе относится данная ошибка, «zzz» – аббревиатура ошибки; при записи кода регистр символов не важен: можно написать **Ф.факт**, **ф.Факт**, **Ф.ФАКТ**, **ф.факт** – все эти варианты правильные;
- **ТекстФрагмента** – неизменённый фрагмент исходного текста эссе; границы фрагмента должны точно указывать локализацию ошибки или смыслового блока; если **ТекстФрагмента** отсутствует, то ошибка относится к тексту в целом (таких случаев желательно избегать);
- **Комментарий** – подтип ошибки, более детально объясняющий учащемуся суть ошибки; для типовых ошибок используются, по возможности, стандартные комментарии или их аббревиатуры (подтипы ошибок) по классификатору; **Комментарий** может отсутствовать;
- **Пояснение** – развёрнутый комментарий, обращающийся непосредственно к тексту данного эссе и конкретно к выделяемому фрагменту; **Пояснение** может отсутствовать;
- **Исправление** – предлагаемый экспертом правильный вариант написания, замещающий **ТекстФрагмента** целиком; **Исправление** может отсутствовать;
- **Тег** – строка из букв или цифр; служит для связывания нескольких фрагментов, относящихся к одной и той же ошибке или смысловому блоку; **Тег** может отсутствовать, если ошибка локализуется только в одном фрагменте.

Элементы разметки **\Комментарий**, **::Пояснение**, **>>Исправление** и **#Тег** могут отсутствовать. Если элемент разметки отсутствует, то его стартовый спецсимвол также опускается. Пробелы вокруг стартовых спецсимволов «\», «::», «>>», «#» игнорируются.

Конструкция «**#Тег**» используется только в тех (довольно редких) случаях, когда ошибку или смысловой блок невозможно локализовать в одном фрагменте, и требуется выделить несколько фрагментов, связав их друг с другом. Связывающий их тег должен быть уникальным, то есть отличаться от других тегов в том же тексте.

1.4 Формат представления данных

Каждая работа сохраняется в отдельном текстовом файле (с расширением .txt) в кодировке UTF8. Использование других форматов, в том числе MS Word (с расширением .doc или .docx) недопустимо.

Если проверка орфографических, пунктуационных, грамматических и речевых ошибок не предполагается для данного типа эссе, то при переносе текста ученической работы из рукописного варианта в txt-файл эти ошибки, включая явные опечатки, рекомендуется исправлять. В таких случаях можно использовать спеллер.

Имя файла должно состояться из двух частей – уникального идентификатора работы и уникального идентификатора проверяющего эксперта. Тем самым по имени файла можно выделить все проверки данной работы и все работы, проверенные данным экспертом.

1.5 Метаописание файла

В начале текстового файла должны быть записаны строки с полями метаописания. Метаописание отделяется от основного текста пустой строкой. Каждая строка метаописания имеет вид

Поле : Значение

где **Значение** – текст до конца строки, **Поле** – одно из predetermined ключевых слов:

Поле	Значение поля
Тема	тема сочинения или текст, по которому написано эссе
Исходный текст	полный текст задания, по которому было написано данное сочинение
Предмет	тип эссе или название предмета, определяющий возможные типы ошибок: русский, литература, обществознание, история, английский, английский-свободное, русский-свободное.
Линия	область, к которой относится тема мини-сочинения; например, для обществознания это одна из шести линий: философия, экономика, социология, социальная психология, политология, правоведение
Класс	строковое значение, например «11», «5», «2 курс»
Год	четырёхзначное число (например, 2018, 2019, 2020); год важен, поскольку правила проверки данного типа эссе могут быть разными для разных лет
Тест	вид экзамена, определяющий критерии оценивания работы: огэ, егэ, тренировка и т.д.
Эксперт	уникальный идентификатор эксперта, проверившего данную работу
KN	оценка работы по критерию с номером <i>N</i>

Если **Значение** содержит несколько строк (например, текст, по которому написано сочинение по русскому языку), то оно заключается в скобки «(\ ... \)» или «(* ... *)»:

Поле: (* Значение *)

Значение любого поля может быть пустым.

Пример метаописания:

Тема: (* В чём заключаются сила и слабость базаровского нигилизма? *)

Класс: 11

Год: 2020

Предмет: литература

Тест: егэ тренировка

Эксперт: СеверусСнейп2020

1.6 Синтаксические ошибки

Синтаксический парсер – это программа, принимающая на входе размеченный txt-файл эссе, и выдающая на выходе список полей метаописания и список фрагментов. Синтаксический парсер может обнаружить следующие ошибки в разметке (в скобках указано действие парсера):

1. Неизвестное поле метаописания (данное поле игнорируется).
2. Неизвестный код, либо забыт разделитель «\» (данный код принимается за начало фрагмента).
3. Не указан код фрагмента (создаётся фрагмент с пустым кодом).
4. Открывающая скобка без закрывающей (считается, что закрывающая стоит в конце текста).
5. Закрывающая скобка без открывающей (закрывающая скобка игнорируется).
6. Закрывающая скобка не соответствует открывающей (считается, что соответствует)
7. Фрагмент «ИСП» не содержит конструкцию «>>Исправление» (исправление игнорируется).
8. Исходный текст эссе искажен (предупреждение, игнорируется).

1.7 Алгоритм работы эксперта-разметчика

Действия эксперта с txt-файлом эссе, полученным на разметку.

1. Занести в текстовый файл все поля метаописания, по которым имеются достоверные данные.
2. Если в тексте присутствует смысловой блок, то его обязательно надо выделить независимо от того, насколько правильно его содержимое.
3. Если в тексте имеется ошибка, за которую эксперт собирается снизить балл по некоторому критерию, то он должен:
 - точно локализовать начало и конец ошибочного фрагмента в тексте;
 - выбрать из классификатора код подходящего типа ошибки;
 - выбрать из классификатора код подходящего подтипа ошибки; если ни один из них не подходит, то написать комментарий;
 - если суть ошибки может быть неправильно понята автором эссе, то добавить пояснение и/или исправление;
 - если фрагмент требует связывания с другими фрагментами, то придумать

- уникальный тег и проставить его во всех связанных фрагментах.
4. Проставить оценки по всем формальным критериям (K1, K2, и т.д.), проверив, что оценки в баллах однозначно вычисляются по разметке согласно правилам проверки данного типа эссе.

1.8 Визуальный интерфейс разметки

При разметке файла эссе с помощью визуального веб-интерфейса действуют следующие правила.

1. Смысловой блок отмечается фоновым светло-зеленым цветом.
2. Ошибка отмечается фоновым светло-красным цветом.
3. Ошибка, совмещенная со смысловым блоком, отмечается фоновым светло-жёлтым цветом.
4. Вложенные блоки одного цвета отмечаются увеличенной интенсивностью фонового цвета.
5. Исправления «ИСП» отмечаются фоновым светло-голубым цветом.
6. Связанные фрагменты помечаются в конце фрагмента тегом, похожим на сноску.
7. Окно ввода атрибутов фрагмента (коды типа и подтипа, комментарий, пояснение, исправление, тег) выводятся на полях по клику на данном фрагменте.
8. Если фрагмент совмещает несколько смысловых блоков и ошибок, то при клике на данный фрагмент выводятся окна атрибутов для всех типов ошибок.

2. Описание разметки данных в формате JSON

Формат JSON используется для представления размеченных текстов эссе в виде, более удобном для компьютерной обработки.

Описание полей формата JSON

Наименование поля	Тип данных	Пустое значение	Соответствие синтаксическим конструкциям языка разметки
meta	объект	не допускается	Метаописание
meta.id	строка	не допускается	Уникальный ID текста, который используется для обращения в поддержку. Может быть пустым для непубличных текстов.
meta.uuid	строка	не допускается	Уникальный технический ID текста
meta.theme	строка	не допускается	Тема
meta.class	строка	пустая строка	Класс
meta.year	число	не допускается	Год
meta.subject	строка	не допускается	Предмет допустимые значения: rus, eng, lit, social, hist, rus-free, eng-free
meta.taskText	строка	пустая строка	Исходный текст — текст задания, по которому написано сочинение
meta.category	строка	пустая строка	Линия
meta.expert	строка	может	Эксперт — поле обязательно для

		опускаться или пустая строка	файлов с экспертной разметкой
meta.test	строка	не допускается	Тест
criteria	массив	допускается	КМ
selections	объект	отсутствует в неразмеченных файлах эссе	последовательность размеченных фрагментов
selections[N].id	число	не допускается	номер фрагмента
selections[N].startSelection	число	не допускается	позиция начала фрагмента посимвольно, начиная с нуля относительно text. Переносы строк считаются за символ.
selections[N].endSelection	число	не допускается	позиция конца фрагмента, посимвольно
selections[N].comment	строка	пустая строка	Комментарий
selections[N].explanation	строка	пустая строка	Пояснение
selections[N].correction	строка	пустая строка	Исправление
selections[N].tag	строка	пустая строка	Тег
selections[N].group	строка	не допускается	тип фрагмента: error — ошибка meaning — смысловой блок
selections[N].type	строка	не допускается	код типа ошибки или смыслового блока по классификатору
selections[N].subtype	строка	пустая строка	код подтипа ошибки по классификатору
Text	строка	не допускается	Текст эссе, написанный учащимся. Не имеет пробельных символов в начале и конце. Параграфы размечены символом переноса строки в unix формате (\n)

Пример

```
{
  "meta": {
    "subject": "литература",
    "test": "егэ тренировка",
    "category": null,
    "year": 2017,
    "class": "11",
    "theme": "Проблема преодоления страха в обычной, мирной жизни",
    "taskText": "Выдержка из текста Ф.А.Вигдоровой"
  },
  "selections": [
    {
      "id": 424,
      "tag": "",

```

```

    "type": "Р.прост",
    "group": "error",
    "comment": "",
    "subtype": "",
    "correction": "",
    "explanation": "",
    "endSelection": 211,
    "startSelection": 53
  },
  {
    "id": 529,
    "tag": "",
    "type": "ИСП",
    "group": "error",
    "comment": "",
    "subtype": "",
    "correction": "Текст correction",
    "explanation": "",
    "endSelection": 447,
    "startSelection": 330
  },
  {
    "id": 530,
    "tag": "",
    "type": "АРГУМЕНТ",
    "group": "meaning",
    "comment": "",
    "subtype": "",
    "correction": "",
    "explanation": "Текст explanation",
    "endSelection": 627,
    "startSelection": 523
  }
],
"text": "Действительно ли, одним из самых тяжёлых испытаний в жизни человека является испытание страхом? В предложенном для анализа тексте автор поднимает проблему... "
}

```

3. Классификатор ошибок

Классификатор ошибок, исправляемых и учитываемых при оценивании работы учащегося, основан на методических рекомендациях ФИПИ [1] и опубликованных демонстрационных вариантах контрольных измерительных материалов ФИПИ [2, 3, 4, 5].

В классификаторе описаны типы ошибок, учитываемых при проведении конкурса отдельных заданий ПРО//ЧТЕНИЕ. «Номинация: Грамматика_Eng»

Ошибки в сочинениях по английскому языку

В сочинении по английскому языку (задание №40 в ЕГЭ 2020) от учащегося требуется

высказать своё мнение по следующему плану:

- (1) перефразированная формулировка проблемы,
- (2) личное мнение с приведением 2-3 аргументов;
- (3) противоположное мнение с приведением 1-2 аргументов;
- (4) обоснование несогласия с противоположным мнением;
- (5) заключение с подтверждением позиции автора.

Эти пять пунктов называются *аспектами*, шестым аспектом считается стилевое оформление с соблюдением нейтрального стиля. Оценивается решение коммуникативной задачи, организация текста, лексика, грамматика, орфография и пунктуация [6,7].

Ошибки, описанные в данном разделе, относятся только к типу эссе **английский**.

Код	Стандартные комментарии	Примеры Тема: (*Sport unites people*)
Лексические ошибки (критерий К3, от 0 до 3)		
А.лекс	Лексическая ошибка. конт Неправильное употребление слова в контексте сочет Нарушение лексической сочетаемости проп Пропуск слова, когда это не влияет на грамматическую структуру предложения слобр Ошибка в словообразовании, не меняющая часть речи послелог Послелог во фразовом глаголе слово Ошибка в написании слов, которая меняет его значение	(\А.лекс слобр\ unregular >> irregular\ (\А.лекс слово\ think >> thing\)
А.запас	Используемый словарный запас не соответствует высокому уровню сложности задания. повтор Повтор одних и тех же слов син Отсутствие синонимов ант Отсутствие антонимов некорр Некорректное употребление слова сочет Нарушение сочетаемости.	(\А.запас повтор\ Scientists think think think that \) (\А.запас син\ Many people think that sport unites people, but some people do not agree. :: 3 раза people \)
Грамматические ошибки (критерий К4, от 0 до 3)		
А.грамм	Грамматическая ошибка. видовр Видовременная форма глагола нел Неличная форма глагола мод Модальный глагол множ Форма множественного числа прит Форма притяжательного падежа существительного	...you will be able to make a good choice of a career only when you (\А.грамм видовр\ will be older >> are older\ ...it will be helpful for you to find (\А.грамм ант\ right job >> the right job\ (\А.грамм слобр\ actively >> activity\)

	<p>сравн Форма степени сравнения прилагательного или наречия</p> <p>арт Артикль</p> <p>пред Предлог</p> <p>мест Местоимение</p> <p>поряд Порядок слов в предложении</p> <p>проп Пропуск слова (подлежащего или сказуемого), влияющий на грамматическую структуру предложения</p> <p>слобр Ошибка в словообразовании, если меняется часть речи</p>	
А.уров	<p>Используемые грамматические средства не соответствует высокому уровню сложности задания.</p> <p>повтор Повтор одной и той же конструкции</p> <p>прост Использование только простых коротких предложений</p> <p>распр Отсутствие распространенных предложений</p> <p>некорр Некорректное использование грамматических средств</p>	

Замечание 1. В экспертной разметке ошибка **А.объем** может отсутствовать, поскольку объём вычисляется автоматически.

Замечание 2. *Повторяющиеся ошибки*, т.е. ошибки на одно и то же правило грамматики, лексики, орфографии, пунктуации, считаются как одна ошибка. Для выделения повторяющихся ошибок используются теги. Все фрагменты одной повторяющейся ошибки должны иметь один и тот же тег. Разные повторяющиеся ошибки должны иметь разные теги. В качестве тега рекомендуется взять аббревиатуру типа ошибки с порядковым номером ошибки (например, **#множ1**).

Формула оценки в баллах

Правила начисления баллов из [6,7] не допускают однозначного вычисления критериев на основе подсчёта числа смысловых блоков и числа ошибок, поскольку содержат интервальные пороговые значения и двусмысленные условия «И/ИЛИ». В предлагаемых формулах неоднозначности устраняются исходя из *принципа неапеллируемости* оценки (что приводит в среднем к смещению оценки в пользу учащегося). Отдельно производится подсчёт числа нераскрытых аспектов (переменная *ошАспекты*) и числа отклонений от предложенного плана (переменная *ошПлан*). Запись [x] означает единицу, если условие x истинно и ноль, если условие x ложно.

если (**А.лекс**≤1) и (**А.запас**=0) то К3 = 3;
 иначе если (**А.лекс**+3***А.запас**≤3) то К3 = 2;
 иначе если (**А.лекс**≤4) и (**А.запас**≤1) то К3 = 1;
 иначе К3 = 0;
 если (**А.грамм**≤2) и (**А.уров**=0) то К4 = 3;

иначе если (А.грамм<=4) и (А.уров=0) то К4 = 2;
иначе если (А.грамм<=7) и (А.уров<=1) то К4 = 1;
иначе К4 = 0;
итоговая оценка К = К3+К4

Условие назначения третьей проверки

Третья проверка осуществляется при расхождении итоговых оценок К двух экспертов по другим типам ошибок, входящим в состав Технического регламента Технологического конкурса «Взаимное обучение на английском языке» на 4 или более баллов.

При этом третьему эксперту предоставляются данные двух предыдущих проверок [2]:

Пример разметки сочинения по английскому языку

Тема: (*Sport unites people*)

Класс: 11

Год: 2017

Предмет: английский

Тест: егэ тренировка

Эксперт: СеверусСнейп

(*ПРОБЛЕМА\ Some people (\А.пункт\ think, >>think \) that sport unites people, while the others (\А.пункт\ think, >>think \) that it is definitely not right. So, it is a great topic to (\А.орф\ disquss >>discuss \). *)

(*ЛМНЕНИЕ\ As for me, I believe that sport (\А.орф\ can not >>cannot \) work in the bad side, only in the good side of our life. (*АРГУМЕНТ\ Because it is a great way to spent time together, side by side. *) Also, I think, that (*АРГУМЕНТ\ with the help of sport you can meet new friends. *) Besides, I am convinced that if you and your friend have the same sport activity, you become (\А.грамм\ more closer \). *)

(*ПРМНЕНИЕ\ Nevertheless, some people believe that (*АРГУМЕНТ\ sport (\А.грамм множ\ make >>makes \) people very nervous and (\А.орф\ exousted >>exhausted \) *). More than that they think that (*АРГУМЕНТ\ sport is just a waste of time*). *)

(*ОБОСНОВАНИЕ\ I do not fully agree with that, because sport (\А.грамм\ help >>helps \) people to relax physically and mentally. Also, sport is the best thing for time spending for (\А.грамм\ your >>you \) and your family, because there are a lot of different kinds of sport which (\А.грамм\ specially were made >> were specially made \) for family *)

(*ВЫВОД\ In conclusion, I would like to say that sport is great thing, that allows us to spend our time not only just for fun and joy, but also for our good health and our happy future. *)

Литература

1. Шерстобитова И.А., Белокурова С.П., Гвоздинская Л.Г. Классификация ошибок, исправляемых и учитываемых при оценивании работы учащегося. ФИПИ, 2012.
2. Демонстрационный вариант контрольных измерительных материалов единого

- государственного экзамена 2020 года по русскому языку. ФИПИ, 2019.
3. Демонстрационный вариант контрольных измерительных материалов единого государственного экзамена 2020 года по литературе. ФИПИ, 2019.
 4. Демонстрационный вариант контрольных измерительных материалов единого государственного экзамена 2020 года по обществознанию. ФИПИ, 2019.
 5. Демонстрационный вариант контрольных измерительных материалов единого государственного экзамена 2020 года по истории. ФИПИ, 2019.
 6. Демонстрационный вариант контрольных измерительных материалов единого государственного экзамена 2020 года по английскому языку. ФИПИ, 2019.
 7. *Вербицкая М.В., Махмурян К.С., Курасовская Ю.Б.* Методические материалы для председателей и членов предметных комиссий субъектов Российской Федерации по проверке выполнения заданий с развернутым ответом экзаменационных работ ЕГЭ 2020 года. Английский язык. Раздел «Письмо». ФИПИ, 2020.