**Уровень 1**

(Уровень 1,2)

1. Найдите время

Для поиска времени в формате "часы:минуты" в строке можно использовать следующее регулярное выражение:

\b(?:[01]\d|2[0-3]):[0-5]\d\b

Это регулярное выражение будет искать время в формате "часы:минуты" в тексте.

2. Java[^script]

Да, данное регулярное выражение Java[^script] найдет совпадение в строке "Java", так как оно ищет слово "Java", за которым следует любой символ, кроме "s", "c", "r", "i" или "p". Однако, в строке "JavaScript" данное регулярное выражение не найдет совпадений, так как после "Java" идет "Script", что соответствует условию исключения символов после "Java".

3. Цвет

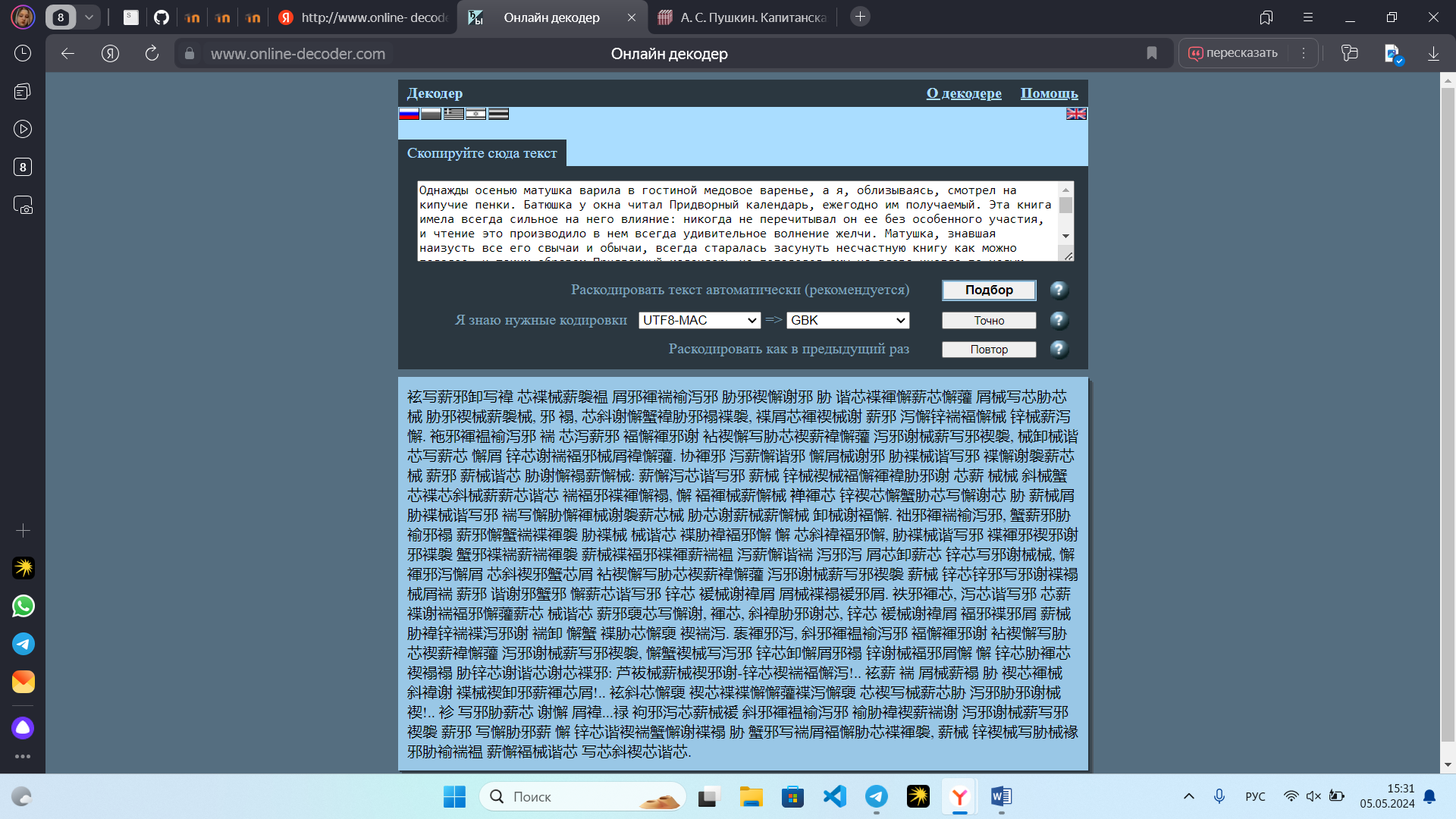
Для поиска HTML-цвета в формате [#ABCDEF](tg://search_hashtag?hashtag=ABCDEF) можно использовать следующее регулярное выражение: #([A-Fa-f0-9]{6})  
Это регулярное выражение означает: - # - символ решетки, который начинает HTML-цвет; - ([A-Fa-f0-9]{6}) - группа из 6 символов, где каждый символ может быть либо буквой от A до F (или a до f), либо цифрой от 0 до 9. Таким образом, это регулярное выражение будет соответствовать строкам вида "[#ABCDEF](tg://search_hashtag?hashtag=ABCDEF)", где A, B, C, D, E и F могут быть любыми шестнадцатеричными символами.

4. Разобрать арифметическое выражение

Для разбора арифметического выражения с использованием регулярного выражения, учитывающего операнды и оператор, можно воспользоваться следующим шаблоном:

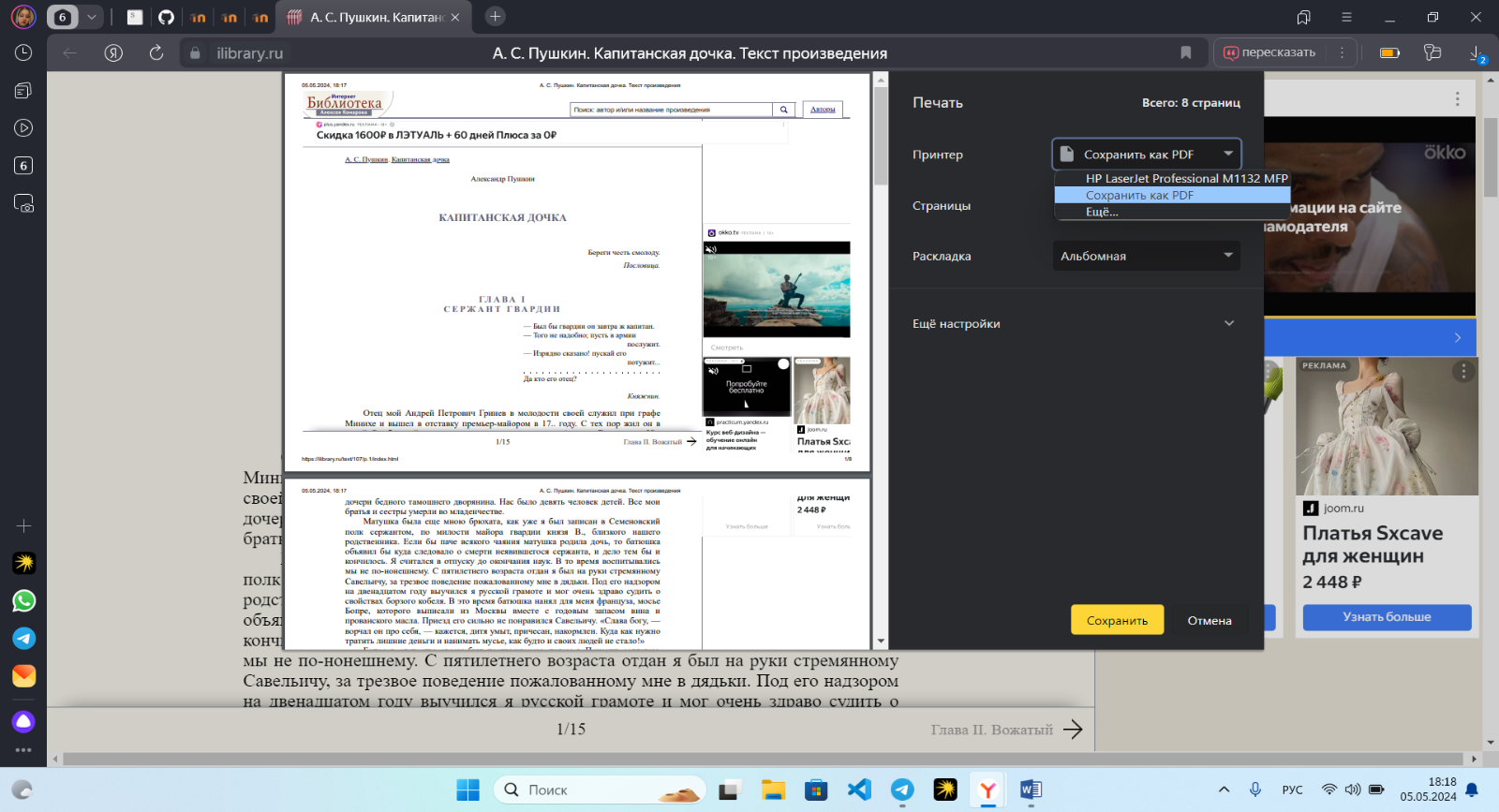
([-+]?\d\*\.?\d+)\s\*([-+\*/])\s\*([-+]?\d\*\.?\d+)  
Это регулярное выражение состоит из трех групп: 1. ([-+]?\d\*\.?\d+) - первый операнд. Здесь: - [-+]? - опциональный знак "+" или "-"; - \d\* - любое количество цифр; - \. - точка для десятичной части (если есть); - \d+ - одна или более цифр. 2. \s\*([-+\*/])\s\* - оператор. Здесь: - \s\* - ноль или более пробельных символов вокруг оператора; - [-+\*/] - сам оператор (один из +, -, \* или /); - \s\* - ноль или более пробельных символов после оператора. 3. ([-+]?\d\*\.?\d+) - второй операнд. Аналогично первому операнду. Это выражение найдет арифметические выражения типа "число оператор число", учитывая возможные пробелы вокруг операторов и чисел, а также учитывая знаки перед числами.

(Задание 1, 2,3)



袨写薪邪卸写褘 芯褋械薪褜褞 屑邪褌褍褕泻邪 胁邪褉懈谢邪 胁 谐芯褋褌懈薪芯懈虇 屑械写芯胁芯械 胁邪褉械薪褜械, 邪 褟, 芯斜谢懈蟹褘胁邪褟褋褜, 褋屑芯褌褉械谢 薪邪 泻懈锌褍褔懈械 锌械薪泻懈. 袘邪褌褞褕泻邪 褍 芯泻薪邪 褔懈褌邪谢 袩褉懈写胁芯褉薪褘懈虇 泻邪谢械薪写邪褉褜, 械卸械谐芯写薪芯 懈屑 锌芯谢褍褔邪械屑褘懈虇. 协褌邪 泻薪懈谐邪 懈屑械谢邪 胁褋械谐写邪 褋懈谢褜薪芯械 薪邪 薪械谐芯 胁谢懈褟薪懈械: 薪懈泻芯谐写邪 薪械 锌械褉械褔懈褌褘胁邪谢 芯薪 械械 斜械蟹 芯褋芯斜械薪薪芯谐芯 褍褔邪褋褌懈褟, 懈 褔褌械薪懈械 褝褌芯 锌褉芯懈蟹胁芯写懈谢芯 胁 薪械屑 胁褋械谐写邪 褍写懈胁懈褌械谢褜薪芯械 胁芯谢薪械薪懈械 卸械谢褔懈. 袦邪褌褍褕泻邪, 蟹薪邪胁褕邪褟 薪邪懈蟹褍褋褌褜 胁褋械 械谐芯 褋胁褘褔邪懈 懈 芯斜褘褔邪懈, 胁褋械谐写邪 褋褌邪褉邪谢邪褋褜 蟹邪褋褍薪褍褌褜 薪械褋褔邪褋褌薪褍褞 泻薪懈谐褍 泻邪泻 屑芯卸薪芯 锌芯写邪谢械械, 懈 褌邪泻懈屑 芯斜褉邪蟹芯屑 袩褉懈写胁芯褉薪褘懈虇 泻邪谢械薪写邪褉褜 薪械 锌芯锌邪写邪谢褋褟 械屑褍 薪邪 谐谢邪蟹邪 懈薪芯谐写邪 锌芯 褑械谢褘屑 屑械褋褟褑邪屑. 袟邪褌芯, 泻芯谐写邪 芯薪 褋谢褍褔邪懈虇薪芯 械谐芯 薪邪褏芯写懈谢, 褌芯, 斜褘胁邪谢芯, 锌芯 褑械谢褘屑 褔邪褋邪屑 薪械 胁褘锌褍褋泻邪谢 褍卸 懈蟹 褋胁芯懈褏 褉褍泻. 袠褌邪泻, 斜邪褌褞褕泻邪 褔懈褌邪谢 袩褉懈写胁芯褉薪褘懈虇 泻邪谢械薪写邪褉褜, 懈蟹褉械写泻邪 锌芯卸懈屑邪褟 锌谢械褔邪屑懈 懈 锌芯胁褌芯褉褟褟 胁锌芯谢谐芯谢芯褋邪: 芦袚械薪械褉邪谢-锌芯褉褍褔懈泻!.. 袨薪 褍 屑械薪褟 胁 褉芯褌械 斜褘谢 褋械褉卸邪薪褌芯屑!.. 袨斜芯懈褏 褉芯褋褋懈懈虇褋泻懈褏 芯褉写械薪芯胁 泻邪胁邪谢械褉!.. 袗 写邪胁薪芯 谢懈 屑褘...禄 袧邪泻芯薪械褑 斜邪褌褞褕泻邪 褕胁褘褉薪褍谢 泻邪谢械薪写邪褉褜 薪邪 写懈胁邪薪 懈 锌芯谐褉褍蟹懈谢褋褟 胁 蟹邪写褍屑褔懈胁芯褋褌褜, 薪械 锌褉械写胁械褖邪胁褕褍褞 薪懈褔械谐芯 写芯斜褉芯谐芯.

(Задание 1,2,3)



(Уровень 1,2,3)

*Хранение информации на диске*

Информация на цифровых носителях хранится в закодированном виде и представляет собой последовательность 0 и 1. В основе запоминающего устройства может лежать любой физический эффект, который обеспечивает приведение системы к двум устойчивым состояниям: одно из них будет обозначать 0, другое - 1 Так, в современных устройствах используются:•наличие/отсутствие заряда в конденсаторе

•направление намагниченности

•отражение/рассеяние света от поверхности CD, DVD или Blu-ray диска

Минимальная единица хранения информации - 1 бит.

Файловая система - это раздел жесткого диска, выделенный для хранения файлов на компьютере и другом электронном оборудовании: мобильных телефонах, фотоаппаратах и т.д.

Название FAT расшифровывается как File Allocation Table, или "таблица размещения файлов". Это значит, что "адрес" каждого файла на диске прописан в специальной таблице. Дисковое пространство разбито на кластеры, и каждый файл занимает не менее одного из них. Файловая система FAT состоит из следующих частей:

1)Загрузочная запись

2) FAT

3) Собственно файлы

Обычно большие файлы записываются в кластеры, идущие подряд, но, когда свободного места на диске остается мало, нужного количества свободных ячеек подряд может не быть. Однако если суммарное количество свободных кластеров достаточно для записи файла, он записывается по кусочкам в разные места диска. Это называется фрагментацией. Правда, разбитые на кусочки файлы читаются гораздо медленнее, поэтому существует дефрагментация — перераспределение файлов на диске таким образом, чтобы каждый из них хранился в непрерывной последовательности кластеров.

*Форматы данных и кодировки*

Практически любая файловая система иерархична: файлы объединяются в каталоги (папки), которые образуют одно или несколько деревьев. Расположение файла в такой структуре описывается с помощью пути. Кроме того, у файла есть расширение - несколько символов после точки, которые обозначают его тип (формат): txt, doc, jpg, mp3, html... Проще говоря, расширение указывает на то, какая информация (текстовая, графическая и т.д.) находится в файле и как правильно ее раскодировать.

*Кодировка текста*

Для перевода последовательности нулей и единиц в текстовые символы используется специальная таблица, которая называется кодировкой; если текст был зашифрован с помощью одной кодировки, а расшифрован с помощью другой, то мы увидим кракозябры. португальском языке в UTF-8, декодированный.

*Краткая история кодировок*

1 ASCII (1963 год)

2 ISO/IEC 646 (1972 год)

3 KOI8 (1974 год, СССР)

4 ISO/IEC 8859 (середина 80-х).

5 Windows-125x (начало 90-х)

6 UTF-8 (Unicode Transformation Format, 8-bit — «формат преобразования Юникода, 8-битный», 1993 год)

Unicode – это стандарт кодирования символов, включающий в себя знаки почти всех письменных языков мира.

Формат PDF Portable Document Format (PDF) —межплатформенный формат электронных документов, разработанный фирмой Adobe Systems. "Межплатформенный" в данном случае означает то, что документ будет одинаково отображаться на любом устройстве с любой операционной системой вне зависимости от того, где он был создан.

**Уровень 2**

Язык разметки — набор символов или последовательностей, вставляемых в текст для передачи информации о его выводе или строении. Текстовый документ, написанный с использованием языка разметки, содержит не только сам текст (как последовательность слов и знаков препинания), но и дополнительную информацию о различных его участках.

TEI (Text Encoding Initiative) — это

1 Cтандартизированный язык разметки

2 Cпособ электронной публикации, при которой основные элементы

документа/текста выделяются с помощью разметки и могут быть

легко идентифицированы (поисковой системой, программой

обработки и т.д.)

3 Cпособ представления документов и коллекций, обеспечивающий

их взаимную совместимость (правильно структурированный

документ xml)

4 Cтруктура (framework) для стандартизованной разметки

документа

5 Библиотека готовых компонентов

Из чего состоит стандартный документ TEI?

1 TEI Header - паспорт документа

2 FACSIMILE (копия источника) - если есть

3 Текст

Структура заголовка (header) в TEI, которая должна быть в каждом из файлов перед текстом грамоты, представляет собой дерево, где теги вложены друг в друга.

Теги закрываются в обратном порядке, т.е. тег, который был открыт первым, закрывается последним.

<fileDesc> (описание файла) содержит исчерпывающее библиографическое описание источника файла.

<titleStmt> группирует информацию о названии произведения и о тех, кто несет ответственность за его содержание.

<title> содержит название произведения, будь то статья, книга, журнал или серия, включая все альтернативные названия или подзаголовки.

<respStmt> содержит информацию о работе над текстом.

<persName> содержит собственное имя или название.

<resp> содержит информацию, описывающую тип персональной интеллектуальной ответственности за приводимые в тексте сведения.

<sourceDesc> содержит библиографическое описание экземпляра текста (или текстов), который служил оригиналом для создания или генерирования электронного варианта.

Одна из причин, по которой мы пользуемся текстовыми редакторами - Sublime, Notepad++ и т.п. - в том, что они позволяют подсвечивать синтаксис языков программирования и языков разметки.

Разметка структуры документа

Создаем TEI документ и выделяем в нем две части: TEI-Header (там мы сложим наши метаданные) и собственно TEXT. В xml документе все наши размышления оформляются в виде комментариев с помощью такой разметки <!-->.

Header

Расставляем теги в TEI HEADER. Там будет и информация из описания документа в издании, и служебная информация о том, кто работал с файлом, необходимая для нашего курса и вашей оценки по нему (кто делал оцифровку и откуда брал документ)

Именованные сущности

Во всех документах необходимо отмечать упоминания лиц (<pers>), географических объектов (<place>) и институций (<inst>, в случае АСЭИ это будут в основном монастыри и церкви). Если при названии объекта или институции присутствуют родовые определения (город, деревня, пустошь, озеро, река, болото, мох, камень, монастырь и т.п.), они включаются в тэги <place> и <inst>. Аналогичным образом, в тэг <pers> включаются титулы (князь). Метафорические именования институций (дом Пресвятой Троицы) рассматриваются наравне с обычными. Если метафорическое и обычное именования помещены рядом, так, что между ними нет других слов (в дом Пречистой в Кириллов монастырь) тэг <inst> используется один раз. Для старорусского языка было обычно повторение предлогов (в монастырь в Кириллов, у Ивана у Петрова сына у Долгополова). Для единообразия открывающие тэги надо ставить перед первым предлогом (т.е. в <inst>монастырь в Кириллов<inst> — неправильно, <inst>в монастырь в Кириллов<inst> — правильно). Но если предлог один, то его надо оставить вне тэга (<place>по реку Шексну</place> — неправильно, по <place>реку Шексну</place> — правильно).

Структурные элементы акта

Не должно быть элементов вне тегов!

Теги, обозначающие части акта, не могут быть вложены друг в друга!

Знаки препинания не отделяются тегом от слова!

Invocatio

\_<inv> \_

В классических западных документах это фраза вроде «Во имя Отца, и

Сына, и Святого Духа».

Intitulatio

<intit>

В этом элементе указывается лицо (или лица), от которых исходит

документ.

Inscriptio

<inscr>

В этом элементе указывается лицо (или) лица, которым документ адресован: От игумена Никифора старцу Исакию.

Salutatio

<sal>

Актовые источники выросли из «обычных» писем, где рядом с обращением естественным образом шло приветствие.

Arenga  
<ar>  
Документ может содержать преамбулу (в дипломатике — аренгу). Она бывает двух видов. Первый — это указание на высшие силы, ради милости которых совершается описываемое далее действие (Пречистыя ради Богородцы, Всемилостивейшего ради Спаса и т.п.). Второй — описание факторов, которые повлияли на появление документа, в т.ч. распоряжений (по слову [имярек]) и согласований с властью (доложа князя/боярина [имярек]).  
Promulgatio  
<prom>  
Promulgatio — это объявление о факте совершения того или иного юридически значимого действия, без конкретного описания его причин, характера и условий, которые излагаются в следующих трех элементах: пожаловал есми Нифонта игумена Кирилова монастыря с братьею.  
Narratio  
<narr>  
Narratio — детальное описание событий, которые вызвали появление документа.

Dispositio

<disp>

Этот элемент включает в себя распоряжения по сути излагаемого в акте решения, в т.ч. объектов, над которыми совершается сделка, их характеристики (если речь идет о деревнях и

землях — детальное описание границ) и условия использования.

Sanctio

\_<sanc> \_

Санкции — это последствия, которые, по мысли составителя документа, должны постигнуть человека, нарушившего условия сделки.

Corroboratio

<corr>

Документ должен был нести на себе признаки, позволяющие удостовериться в его подлинности.

Datum

\_<dat> \_

В этот элемент входит указание на дату и место составления документа, помещаемое до или после сведений об удостоверяющих признаках документа.

Apprecatio

<appr>

Поскольку актовые источники выросли из писем, в них мог сохраняться и такой элемент письма, как заключительное благопожелание (далекий предок современных Искренне Ваш, Best и т.п.).