**XML**

XML – подмножество SGML для определения языков описания данных

SGML – сложная фигня

Область применения:

1. Создание DSL для описания данных
2. Альтернатива реляционной модели
3. Мультиплатформенный формат для передачи данных
4. Отделение семантики данных от их представления

HTML vs XML:

Xml чувствителен к регистру, нельзя опускать закрывающие теги, совмещение открытия и закрытия, значение атрибута в кавычках, атрибуты должны быть с значениями.

CDATA – данные, которые не будут парситься: <![CDATA[…]]!>

Декларация в XML. Standalone определяет местонахождение схемы.

Комментарии в XML: <!-- … -->

Символы: алфавит, арабские цифры, точки, дефисы, подчёркивание, двоеточие – используется только для именования пространств имён.

Вложенные теги и атрибуты.

**ВАЛИДАЦИЯ**

Корректность: well formed – все элементы имею закрытый тег, элементы являются строго вложенными, значение атрибутов в кавычках, корневой элемент всегда один, всё вышеперечисленное выполняется с учётом регистра.

Valid – распространение схемы: DTD, XSD (W3C XML Schema), RELAX NG и проверка этими схемами, иногда содержит схему в себе.

URI – Uniform Resource Indicator.

XSL – eXtensible Stylesheet Language – список рекомендаций консорциума W3C для обработки и визуализации XML. Не имеет отношение к CSS. Включает в себя: XSLT – eXtensible Stylesheet Language Transformation: шаблоны, правила –

1. Для обработки корневого узла: обработать каждого потомка коревого узла,
2. Для любых текстовых узлов или узловых атрибутов: сделать копию узла в результирующем дереве,
3. Для всех узлов комментария и узлов-инструкций обработки: никакой операции не производится;

XSL-FO, XPath.

**JAXP API**

Разбор xml документов и преобразование.

XML парсеры в JDK.

DOM – Document Object Model. Тоже можно валидировать.

SAX – Simple API for XML.

SAX: Features and Properties.

SAX vs DOM

StAX.