

Технология DITA: обзор возможностей и основные преимущества

Корьевкина М. М.

PhiloSoft Technical Communication

Август, 2008



Что такое DITA?

Darwin:

*в технологии реализованы принципы
специализации и наследования*

Information Typing:

*типизация информации и
структурированное содержание тем*

Architecture

История развития

DITA Evolutionary Timeline

IBM GML 1960s

ISO SGML 1986

W3C XML 1998

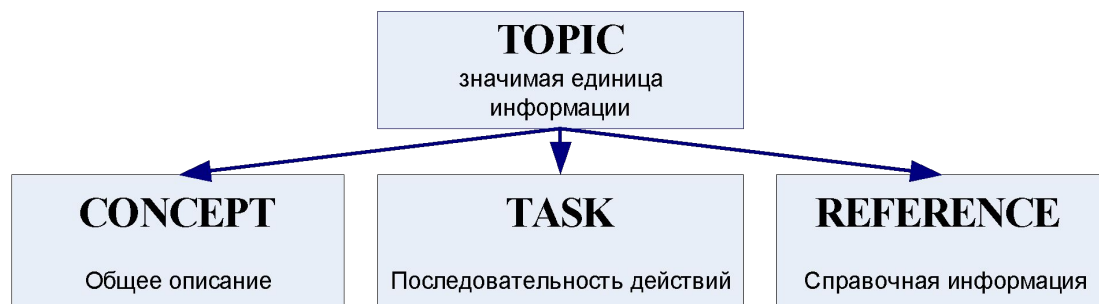
IBM DITA 2001

DITA v1.0 2005

- Технология разработана в IBM в 1999-2000 г.
- О ней впервые объявлено в 2001 г.
- 4 апреля 2005 г.: документация на DITA для Adobe Creative Suite 2
- 9 мая 2005: OASIS (Organization for the Advancement of Structured Information Standards) выпускает первый официальный релиз DITA
- В настоящее время развивается при поддержке OASIS DITA Technical Committee и IBM
- Используется в IBM, ADOBE, NOKIA, etc

Концепция представления информации

- Технология основана на XML, единый источник
- **Тема (topic)** – основная единица информации
- **Concept, Task, Reference** – специализированные типы тем



- **Domain** – набор элементов, ассоциированных с предметной областью
- **Map** – организует темы в последовательность. Представляет собой коллекцию ссылок на топики (**topicref**).
- **Conref** – возможность многократного использования содержимого
- Специализация

Инструментарий

■ DITA Open Toolkit

- DITA DTD
- XSL-стили
- Шаблоны ANT-файлов
- Документация и примеры
- ANT
- Java-классы, используемые при сборке

■ Java Developer Kit 1.4.2_08

■ Дополнительный инструментарий

- XEP+Idiom FO Plugin+библиотека ICU4J
- HTMLHELP
- XML-редакторы (oXygen, Syntext Serna, XMLMind, etc.)

Специализация

- Позволяет создавать новые типы информации на основе существующих.
- **Два вида специализации:**
 - специализация структурных типов (map, topics);
 - специализация словарей (domain, attribute).
- **Принцип наследования:**
 - специализация задачи (task) по-прежнему является задачей, специализация словаря интерфейса пользователя (ui-domain) по-прежнему является словарем интерфейса пользователя

Специализация структурных типов

- **Основные правила специализации структурных типов:**
 - Каждый новый тип информации должен отображаться в один из существующих.
 - Каждый новый тип информации должен быть более ограниченным, чем его родитель.
 - Каждый новый тип информации должен быть определен в отдельном DTD модуле.
 - В модуле должны быть определены только новые элементы типа информации.
- **Процедура специализации:**
 1. Определить структуру типа информации.
 2. Определить соответствие типа информации более общему типу (mapping).
 3. Проверить, что содержимое (content model) специализированных элементов более ограничено, чем их более общий эквивалент
 4. Создать модуль с описанием (type module), который будет содержать декларацию структуры специализированного элемента и его атрибутов (.mod).
 5. Создать DTD модуль (dtd).
- **Для специализированных типов информации может (желательно) использоваться свой процесс обработки (XSLT)**

Пример

Документ

Документ	
Base type	Specialized Type
topic	tdocument
(topic.mod)	(tdocument.mod)
topic	tdocument
title	tdoctype
body	tdocbody
section	tdocdesc
fig	tdocform
ul	tdocfields
li	tdocfield
ul	tdocbuttons
li	tdocbutton
relatedlinks	

Tdocument.mod

```
...
<!ENTITY % tdocument " tdocument" >
...
<!ELEMENT tdocument ((%tdoctype;), (%tdocbody;),
(%related-links;)?) >
<!ATTLIST tdocument
    %univ-atts;
    outputclass
        CDATA                #IMPLIED
    domains CDATA            "&included-domains;" >
...
<!ATTLIST tdocument %global-atts; class CDATA "- topic/topic
tdocument/tdocument" >
...
```

Tdocument.dtd

```
...
<!ENTITY % hi-d-dec PUBLIC
"-//OASIS//ENTITIES DITA Highlight Domain//EN"
"highlightDomain.ent" >
%hi-d-dec
...
<!ENTITY % tdocument-info-types "no-topic-nesting" >
...
<!ENTITY % tdocument-typemod
PUBLIC "-//PH//ELEMENTS DITA Tdocument//EN"
"tdocument.mod" >
%tdocument-typemod;
```


Специализация словарей (доменов)

- Специализация доменов позволяет определить новые типы тех элементов, которые не **должны зависеть от структуры информации**.
- Каждый домен реализован с помощью двух файлов:
 - Файл, в котором декларируются сущности домена (**.ent**).
 - Файл, в котором декларируются элементы домена (**.mod**).
- Правила специализации доменов:
 - **Иерархия специализации топиков и доменов должна быть отдельной.** Специализированный тип информации не должен содержать элементы доменов, а домены могут быть специализированы только от элементов типа информации или других доменов.
 - Для элементов с внутренней структурой (списки, таблицы) следует специализировать весь элемент, а не его составляющие.
- Нельзя специализировать элементы highlight domain.



Специализация атрибутов

- Используется для создания собственных атрибутов путем специализации атрибутов **base** и **props**.
- Используется для **conditional processing**.
- Процедура специализации заключается в создании файла с объявлениями сущностей (**.ent**).

Пример

Workflow domain elements

document
process
status
action
role

Userrole attribute

userrole

workflowDomain.mod

```
...
<!ENTITY % document    "document" >
...
<!ELEMENT document     (#PCDATA)* >
<!ATTLIST document
      keyref    CDATA    #IMPLIED
      %univ-atts;
      outputclass
              CDATA      #IMPLIED >
...
<!ATTLIST document %global-atts; class CDATA "+ topic/keyword
wf-d/document " >
```

workflowDomain.ent

```
...
<!ENTITY % wf-d-keyword  "document | process | docstatus | action | role" >
...
```

userrolePropsDomain.ent.mod

```
...
<!ENTITY % userrole-props-d-attribute "userrole CDATA #IMPLIED">
...
<!ENTITY userrole-props-d-att "a(props userrole)" >
```

Пример использования специализированного типа

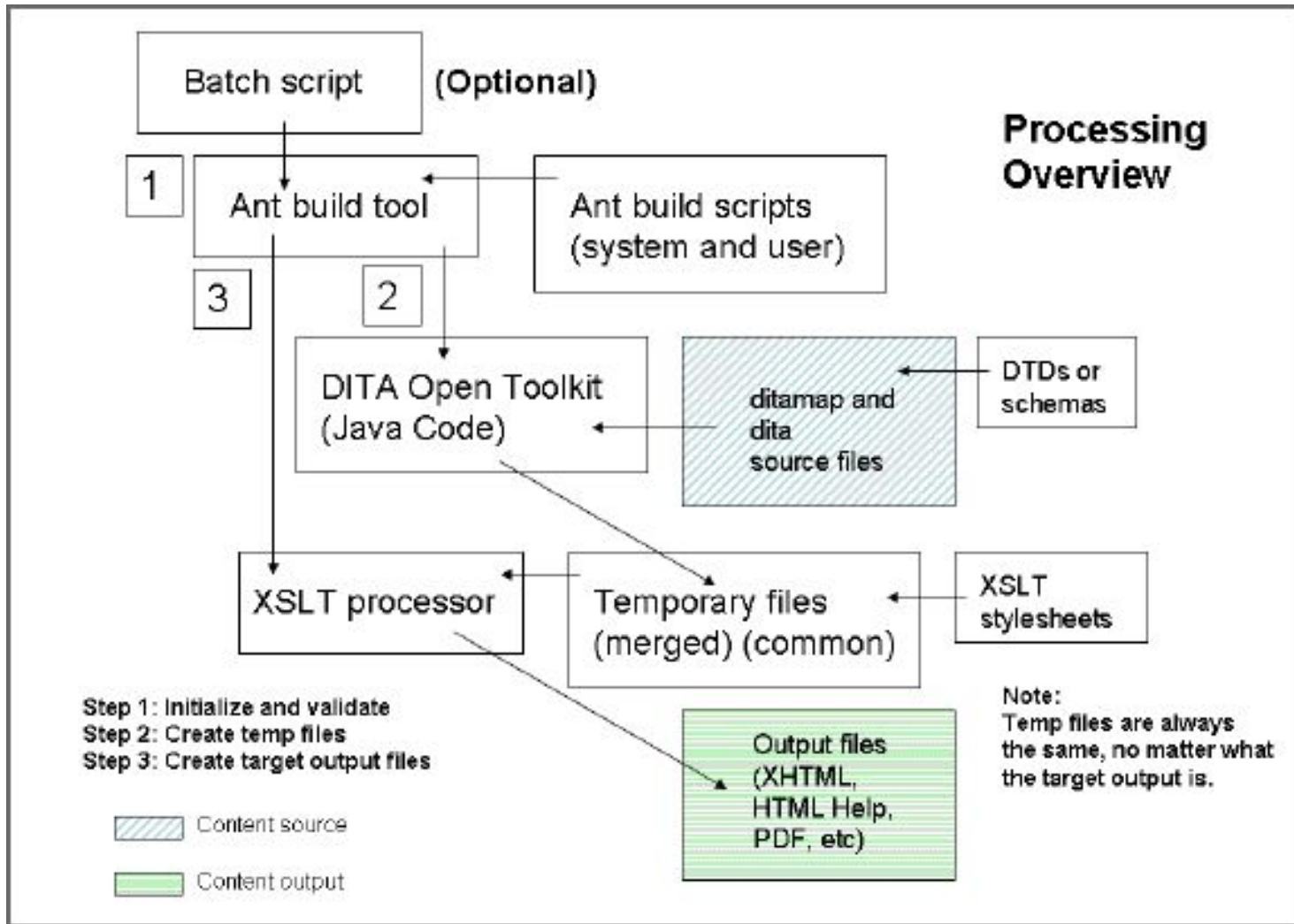
```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE tdocument PUBLIC "-//PH//DTD DITA Tdocument//EN"
"../dtd/tdocument.dtd">
<tdocument id="clientcard" xml:lang="ru-ru">
  <doctitle>Карточка клиента</doctitle>
  <tdocbody>
    <tdocdesc>
      <p>В системе регистрационной форме заказчика соответствует документ
<document>карточка
      клиента</document>. Документ создается на основании информации,
полученной от
      заказчика.</p>
    </tdocdesc>
    <p><document>Карточка клиента</document> представляет собой форму,
показанную на рис. <xref
      href="#clientcard/pic_ClientCard" />.</p>
    <tdocform id="pic_ClientCard">
      <title><document>Карточка клиента</document></title>
      <image href="images/ClientCard.png" />
    </tdocform>
    ...
  </tdocbody>
</tdocument>
```



Настройка оформления

- Редактирование CSS
- Редактирование XSL
- Настройка конфигурационного файла для сборки документов ANT

Общая схема сборки



Сборка документов

- Возможность собирать документы в форматах: **XHTML**, **HTML Help**, **PDF2 (RenderX XEP)**, Eclipse Help, DocBook, Eclipse content, JavaHelp, PDF (Apache FOP), troff, Word RTF.
- **Ant Build Script** – конфигурационный файл для сборки документов. Параметры сборки.
- **Запуск сборки:**
 - Запустить **startcmd.bat**
 - В командной строке набрать:

Ant -f <путь_к_конфигурационному_файлу_сборки>

Ссылки

- **Общая информация**

- <http://dita.xml.org/>
- <http://xml.coverpages.org/dita.html>
- <http://www.ibm.com/developerworks/xml/library/x-dita3/index.html>
- <http://docs.oasis-open.org/dita/>
- <http://www.lone-dita.com/>

- **Инструментарий**

- <http://sourceforge.net/projects/dita-ot/>

- **Tutorials**

- <http://www.ditainfocenter.com/eclipsehelp/index.jsp>
- <http://www.lone-dita.com/DITATutorial/tabid/53/Default.aspx>

- **Здесь отвечают на вопросы**

- <http://tech.groups.yahoo.com/group/dita-users/>
- <http://www.ditausers.org/>
- <http://dita.xml.org/forum>