

## 4Spectra - Initial Planning Document

Last Updated: 19 June 2002

Homepage: <http://spectra.sourceforge.net>

Author and Copyright:

Copyright © 2002 - Saurabh Gupta

This program is free software; you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation; either version 2 of the License, or (at your option) any later version.

This program is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details.

You should have received a copy of the GNU General Public License along with this program; if not, write to the Free Software Foundation, Inc., 59 Temple Place, Suite 330, Boston, MA 02111-1307 USA

## Table of Contents

1. About this document
  - 1.1. Intended Audience
2. Content Structuring
3. Tools

## 1. About this document

### 1.1. Intended Audience

This document is intended for Content Managers, Content Architect, Web Managers, Webmasters or anyone who deals with content. This document contains the details of how the content will be structured in Spectra.

## 2. Content Structuring

- Content Provider (CP)
  - ✓ A content provider is an entity that generates content. This entity provides content to us under a contract.
- Contract
  - ✓ All content is acquired under a contract. A contract usually defines terms of use, duration of use etc.
  - ✓ There may be one or more contracts with a CP.
- Contract Unit
  - ✓ Sometimes there is a need to separate content, even when it is from one CP and is acquired under one contract.
- Content Unit (CU)
  - ✓ A logical grouping of related Elemental Units.
  - ✓ This is the smallest saleable unit of content.
  - ✓ In tangible terms a Content Unit can be thought of as an XML file.
- Elemental Unit (EU)
  - ✓ Each CU is made up of many elemental units.
  - ✓ An EU is the smallest
    - For example the "Headline" and "Date" in a news story are different elemental units in a CU where CU is the news story.
- Information Unit
  - ✓ Meta-tagged CU
    - For example, defining that a news article belongs to the "Current Affairs" section.
- Search Unit
  - A Search Unit is a category/description pair, which is used to qualify a CU.
  - For example, Section=Current Affairs, where Section is the *category* and Current Affairs is the *description*
- Section Unit
  - A Grouping of Search Units and Elemental Units.
  - Search Units have a hierarchical structure i.e. they can have a parent-child relationship with other Section Units.
- Product Unit
  - Groups of Section Units together make a Product Unit.
- Product
  - Product is what is finally sold to a customer.
  - A customer may be sold one or more Products

### 3. Tools/Components

- 3.1. Schema Editor
- 3.2. Schema Instance Validator
- 3.3. XML Parser
- 3.4. XSLT Engine
- 3.5. Xquery Parser
- 3.6. Indexer/ Search Engine
- 3.7. Compression Library
- 3.8. Content Server/Daemon – SOAP Based
- 3.9. Replication Manager