EXPWAY
MAKING XML EFFICIENT™

# Panorama des formats de description de documents audiovisuels

**Claude Seyrat**
CTO Expway

EXPWAY
MAKING XML EFFICIENT™

# Plan

- Introduction
- MPEG-7
- TV-Anytime
- ESG / ECG / EPG
- Conclusion

EXPWAY
MAKING XML EFFICIENT™

# Context

- Amount of audiovisual material is increasing
  - Example: Digital TV:
    - 300 h/h of TV today
    - 1000 h/h of TV programs in 2005 (24000 h/d !!)
- Finding, filtering and managing AV material is becoming a major issue

➔ MPEG-7 & TV-Anytime

# MPEG-7

EXPWAY
MAKING XML EFFICIENT™

# MPEG Standards

- MPEG-1 — 11/1992
  - Storage and coding of moving picture and audio
- MPEG-2 — 11/1994
  - Digital Television
- MPEG-4 (v2) — 11/1999
  - Coding of natural and synthetic media objects for multimedia applications
- MPEG-7 — 08/2001
  - Multimedia Content description for AV material
- MPEG-21 — coming soon
  - Multimedia Framework

# MPEG-7

- Standardize AV content metadata
  - To ease the management of AV documents
    - Fast and efficient search
    - User oriented filtering
    - Classification / organizations of AV DB
    - ...
  - By describing different type of features
    - Low level — colors, shapes, ...
    - Structural — scene, shot, ...
    - Semantic — relations, entities, ...
    - Organizational — collections, models, ...
    - ...

# Audio-Visual Metadata

Blue

Cut

T. L'hermitte

Cut

Eiffel Tower

Cut

Motion

« indien dans la ville »

AV Metadata

# MPEG-7 Technical overview

EXPWAY
MAKING XML EFFICIENT™

# MPEG-7 - Terminology

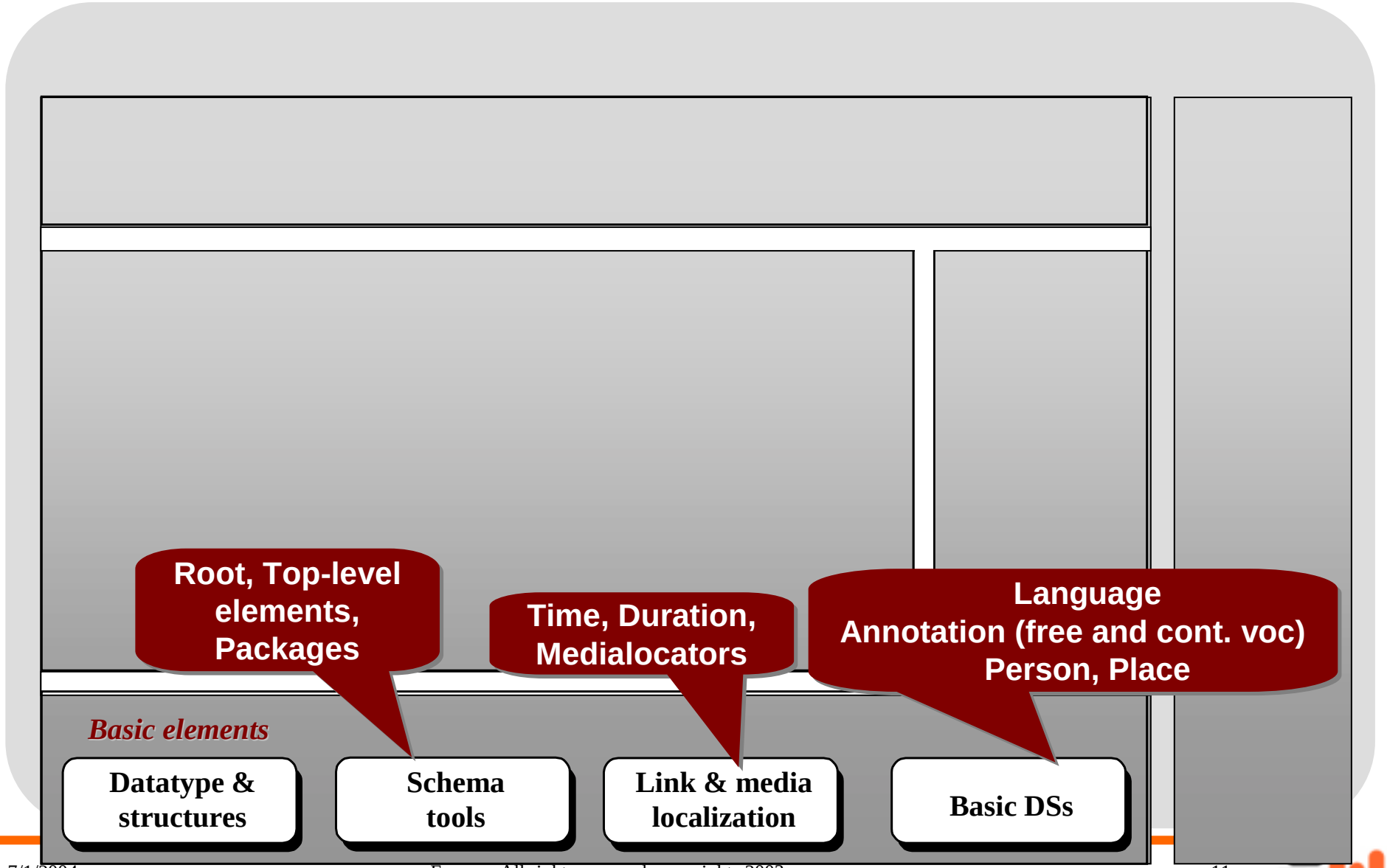
- Description
  - A metadata for AV material
- D : Descriptors
  - Syntax and semantic of representation AV features,
- DS : Description Schemes
  - Structure and semantics of relations between description components,
- DDL : Description Definition Language
  - Language to allow the creation and extension of DSs and Ds
- Systems tools
  - Encoding/decoding, compression and streaming of descriptions,

EXPWAY
MAKING XML EFFICIENT™

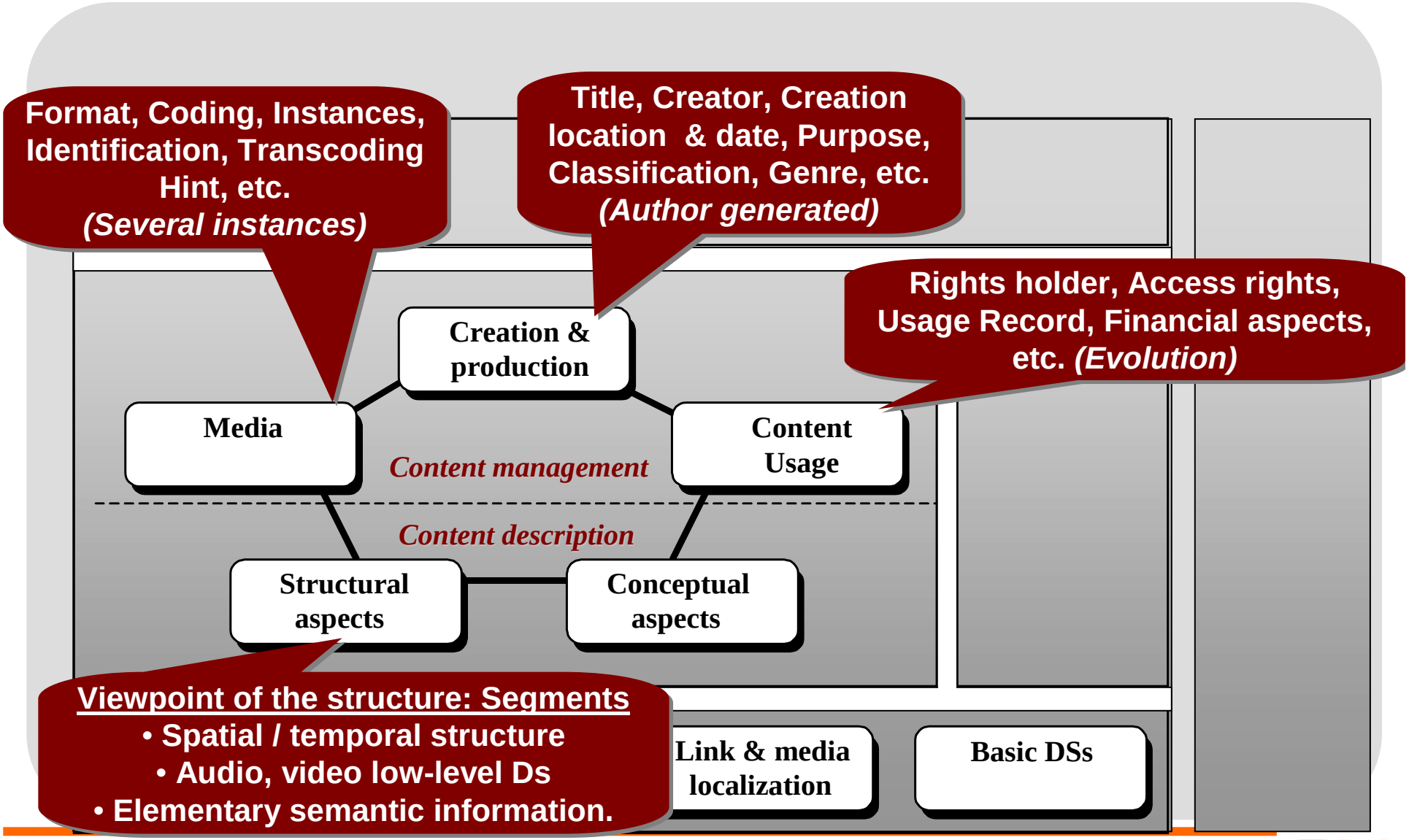
# MPEG-7 Standard

| | | |
|---|---|---|
| 1. | Systems | Transmission format |
| 2. | DDL | XML Schema |
| 3. | Audio | Audio MD |
| 4. | Visual | Visual MD |
| 5. | Multimedia DS | Structural MD |
| 6. | Reference Software | Open source soft |
| 7. | Conformance | Methods to test |
| 8. | Extraction and use of MPEG-7 descriptions | Algorithms and methods |
| 9. | MPEG-7 profiles and levels | Complexity reduction |

# Basic elements

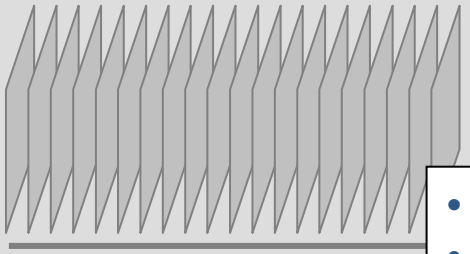


# Content Management & Description

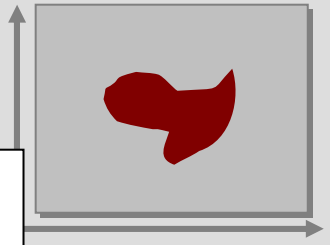


# Examples of Segments

## Video segments

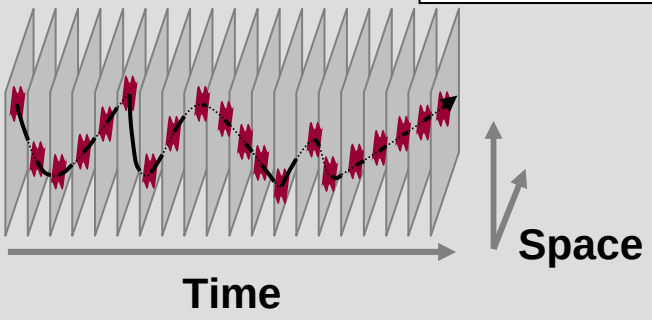



## Still regions

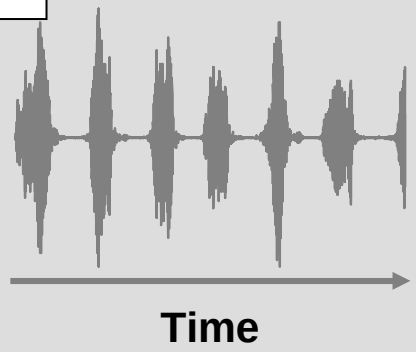



- Identification
- Creation Information
- Usage Information
- Media Information
- Annotation
- Segment Decomposition

## Moving regions

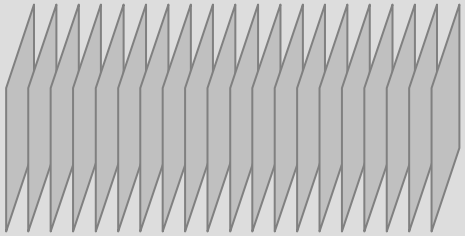



## udio segments



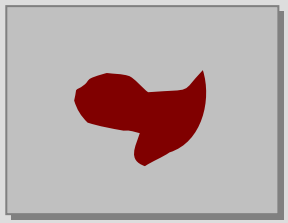
# Specific description of segments

## Video segments



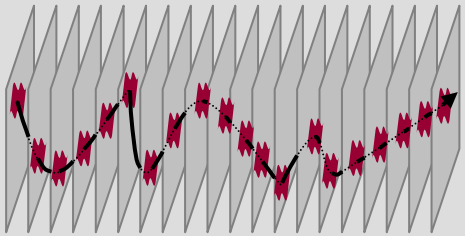
- Color
- Camera motion
- Motion activity
- Mosaic

## Still regions



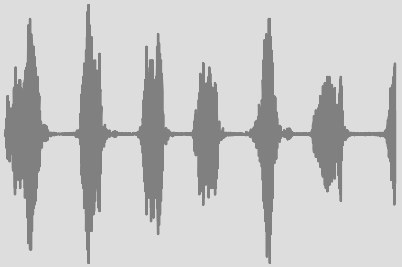
- Color
- Shape
- Position
- Texture

## Moving regions



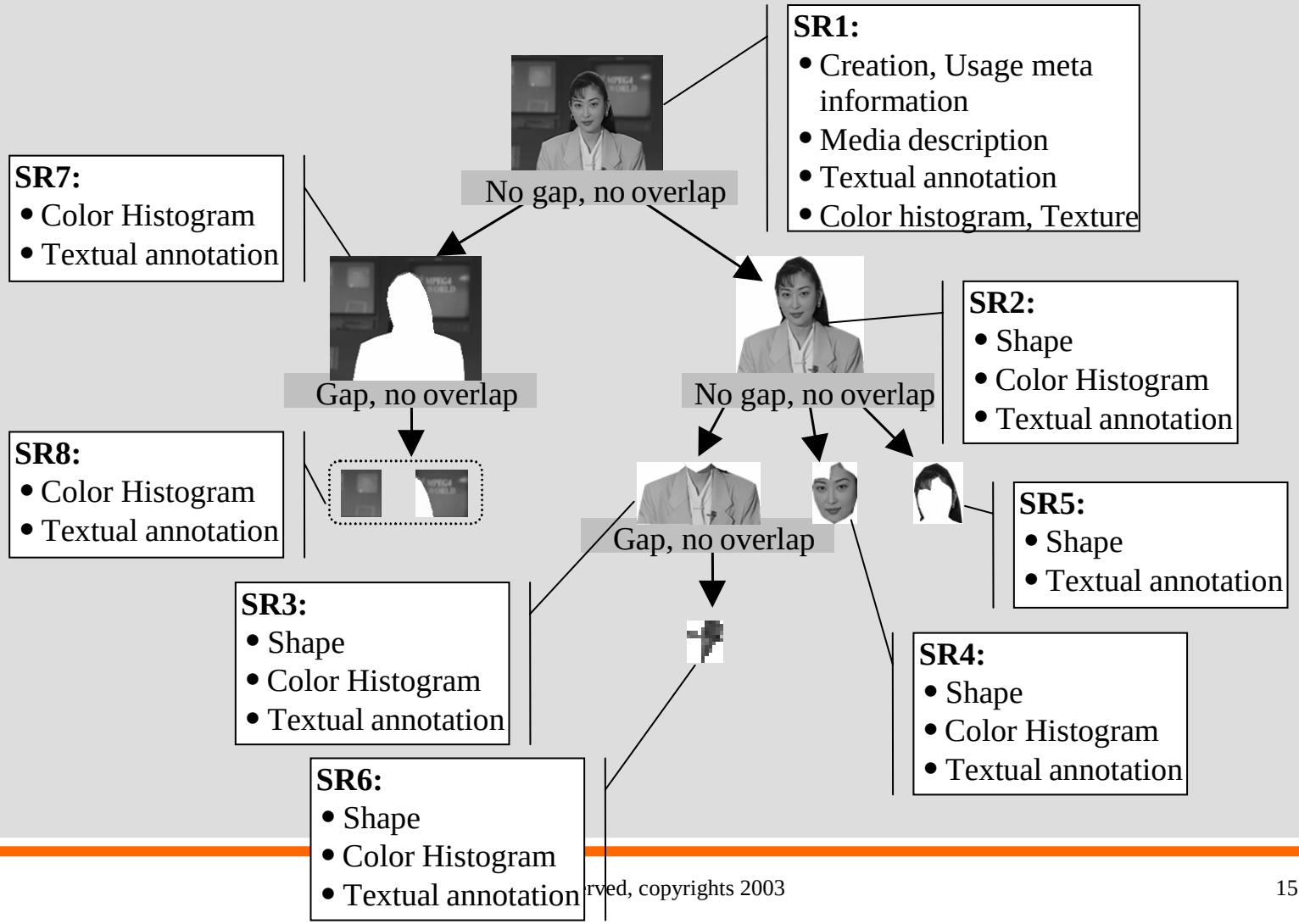
- Color
- Motion trajectory
- Parametric motion
- Spatio-temporal shape

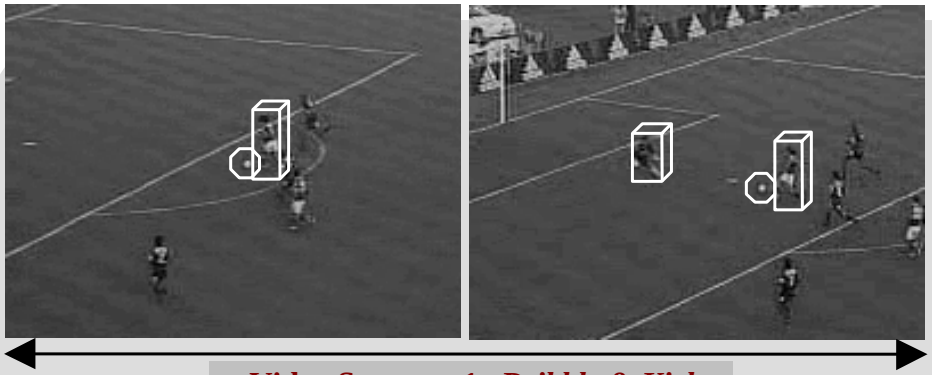
## Audio segments



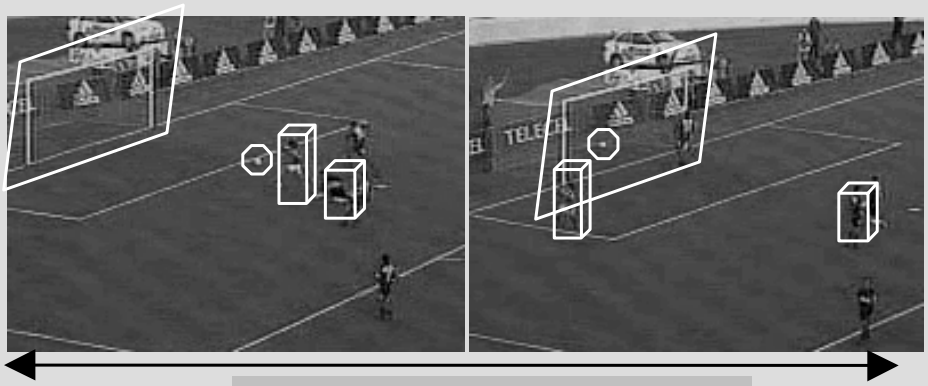
- Spoken content
- Spectral characterization
- Music: timbre, melody

# Example of Segment trees





**Video Segment 1: *Dribble & Kick***



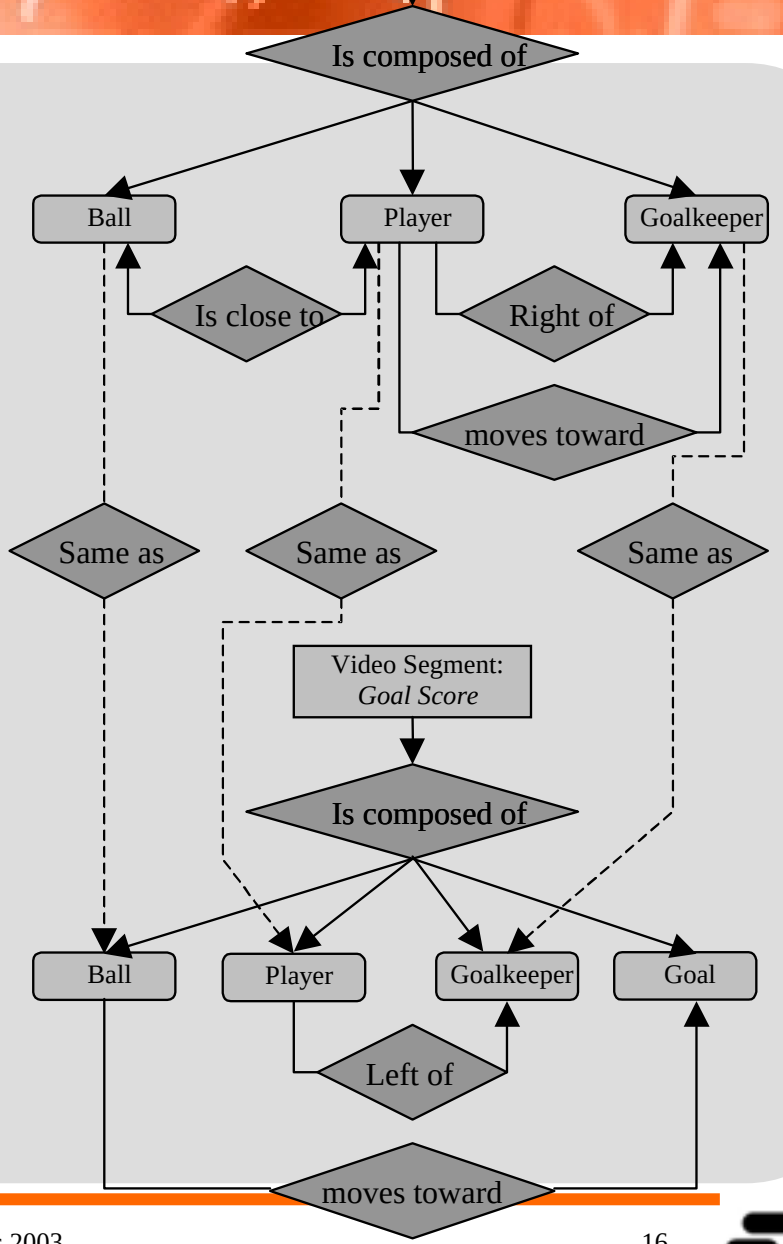
**Video Segment 2: *Goal Score***

Moving Region: Player

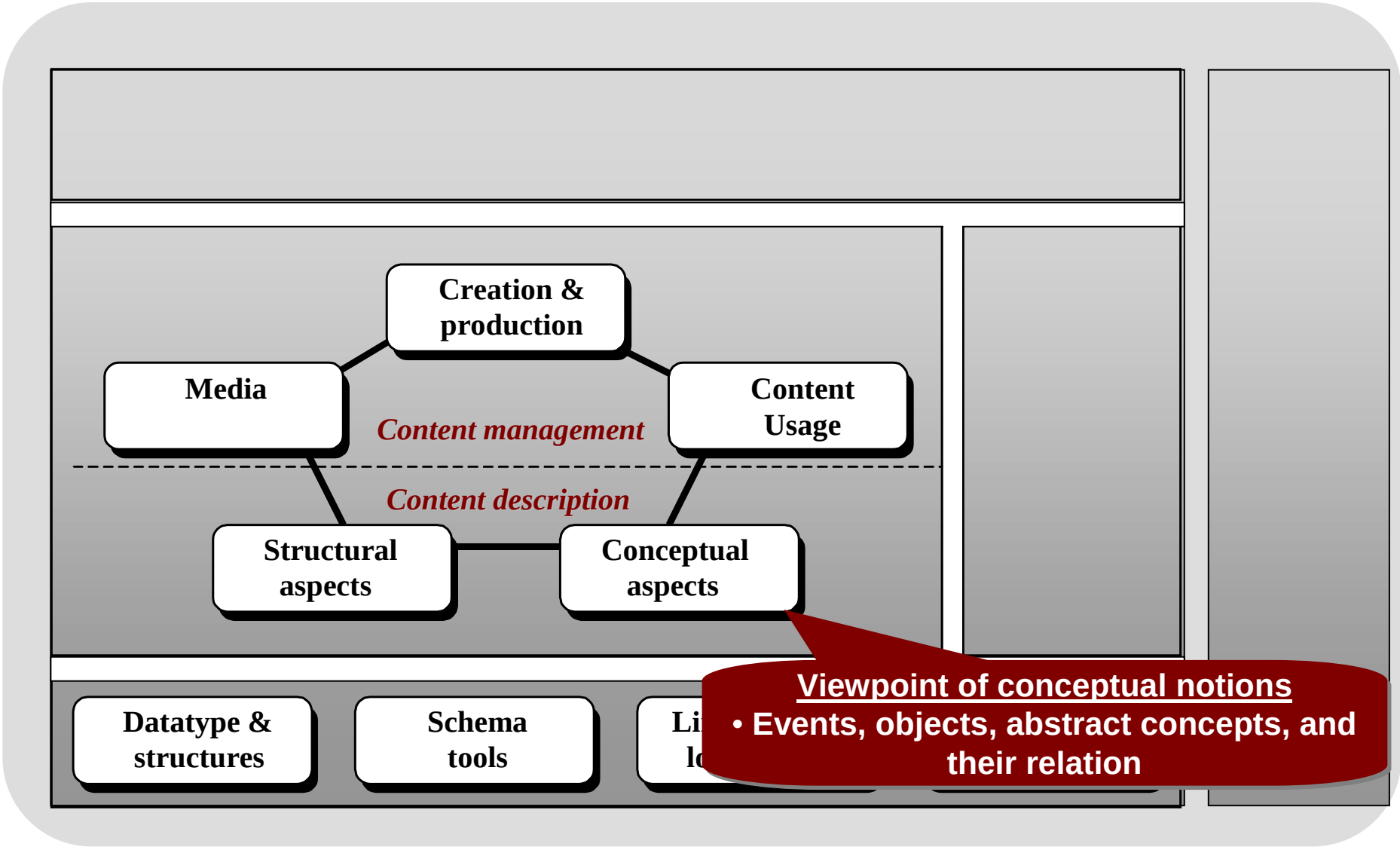
Moving Region: Ball

Moving Region: Goal Keeper

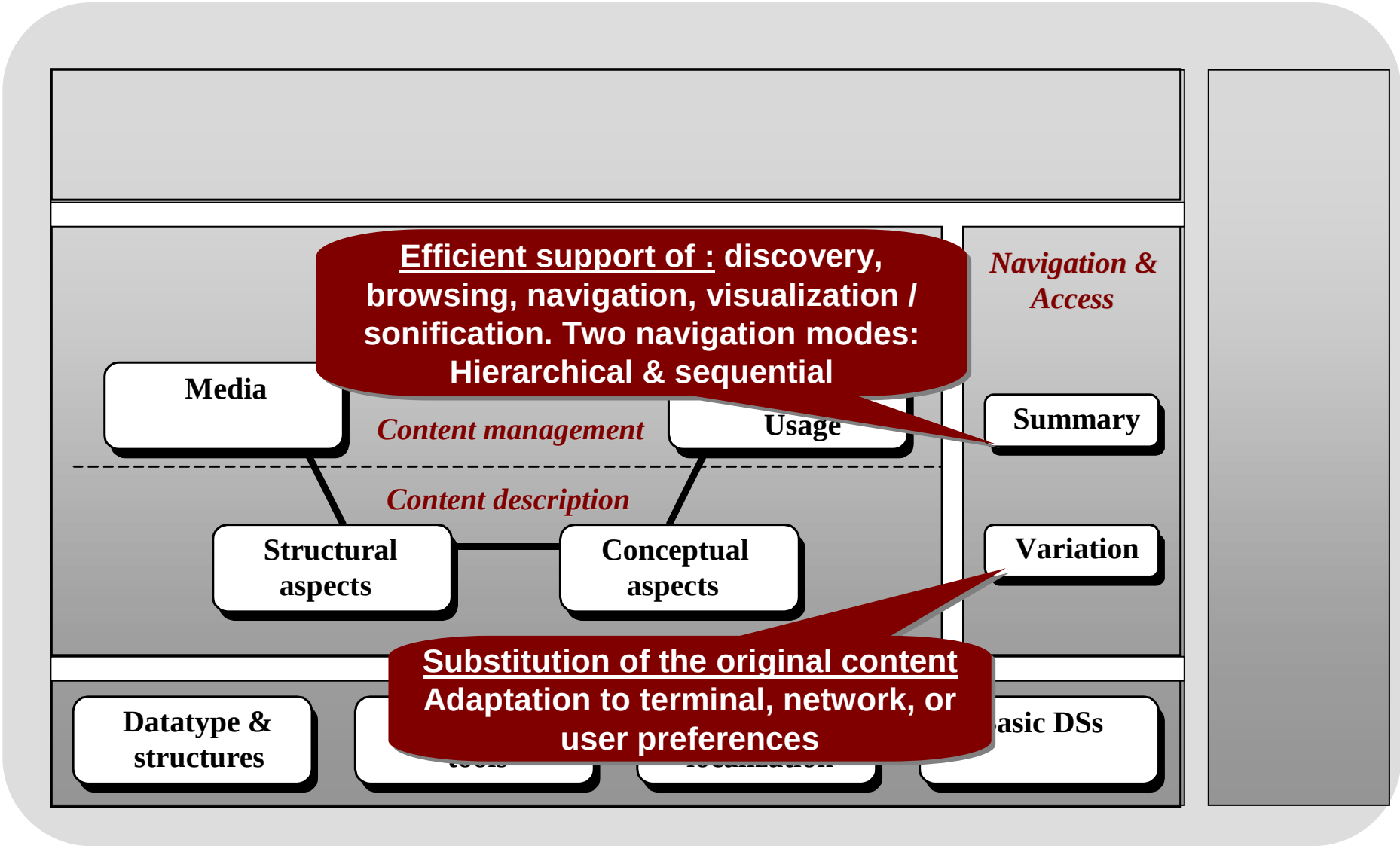
Still Region: Goal



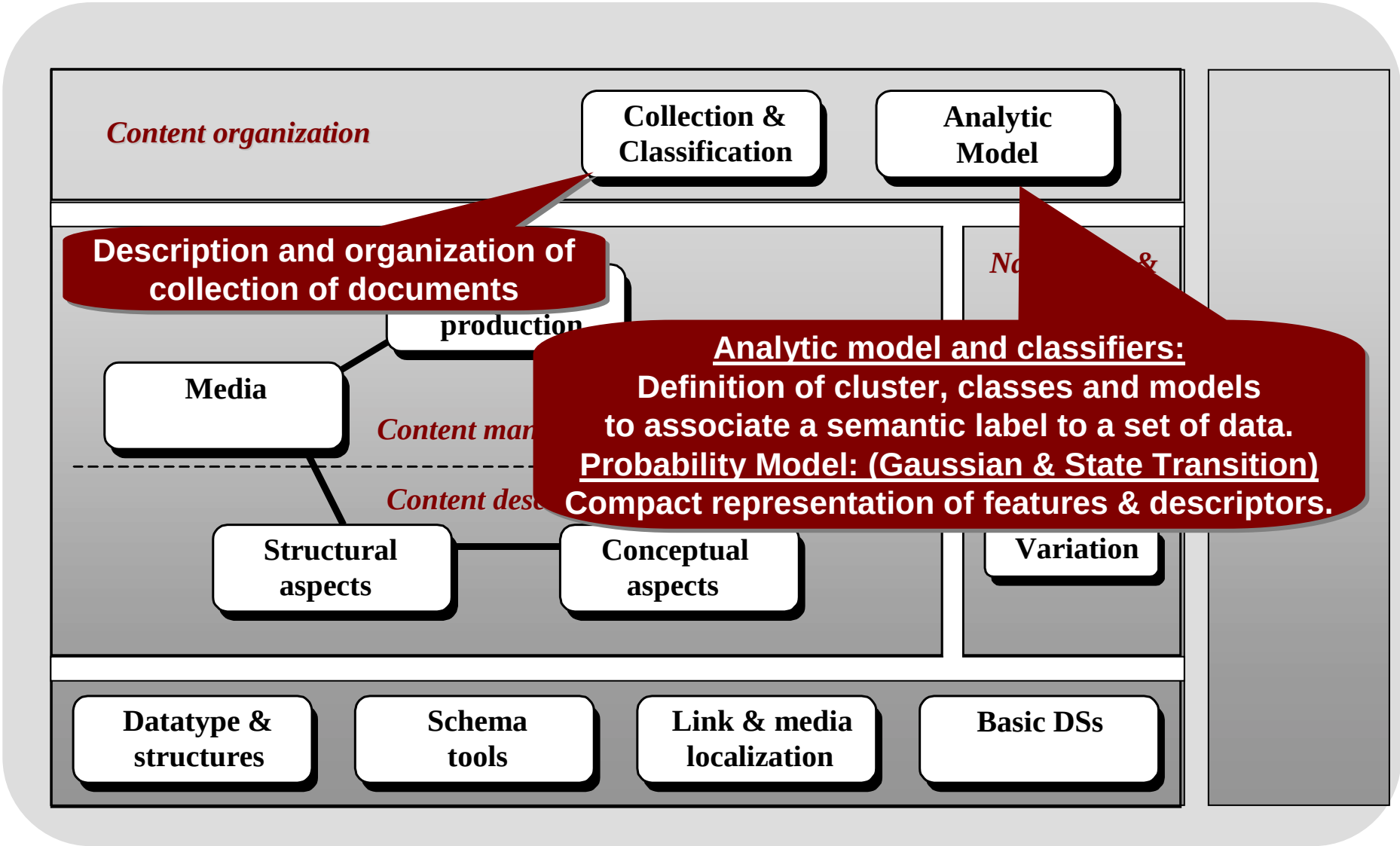
# Content Management & Description



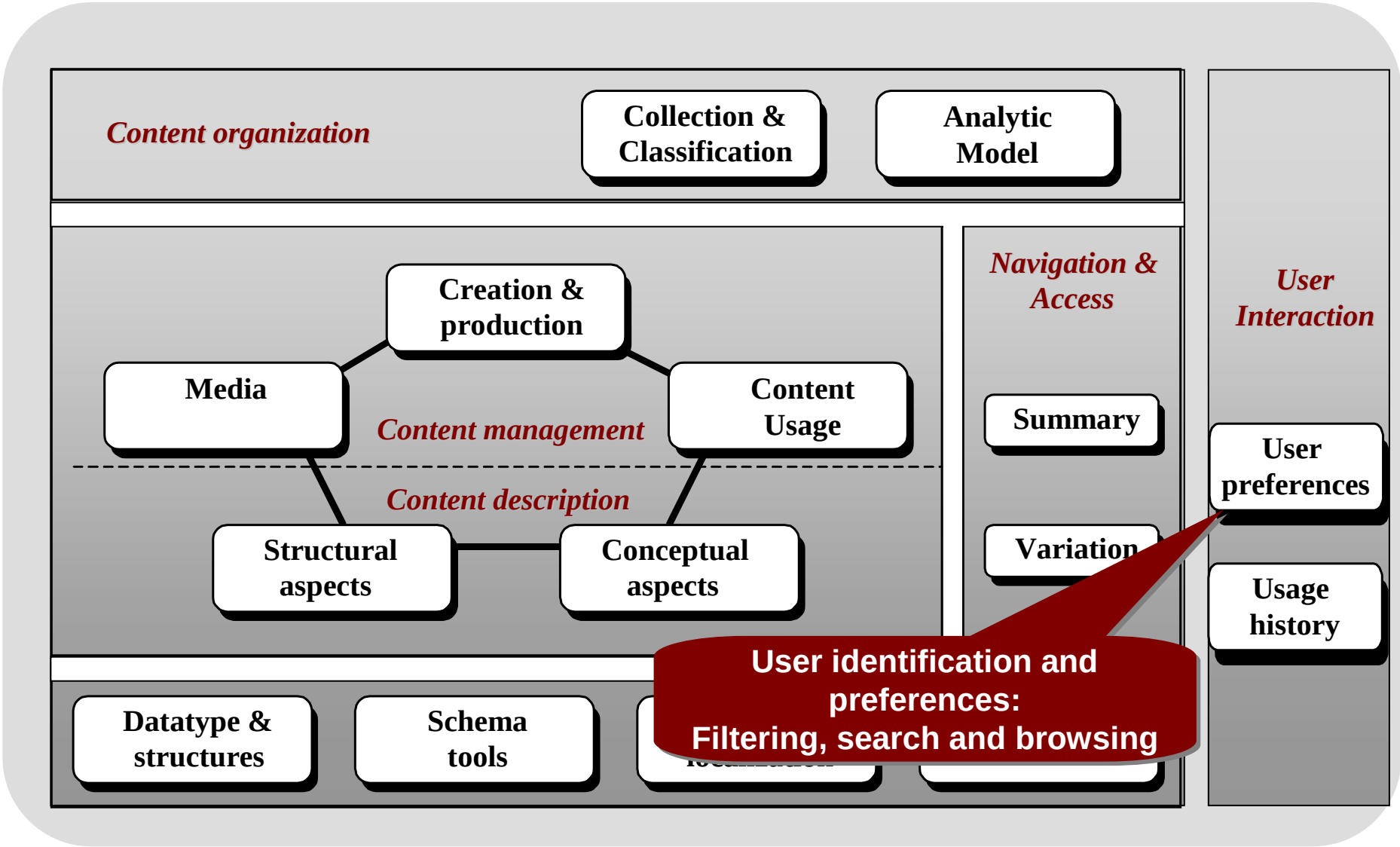
# Navigation and Access



# Content Organization



# User Interaction

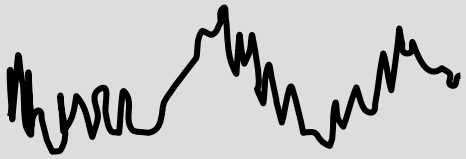


# MPEG-7 Standard

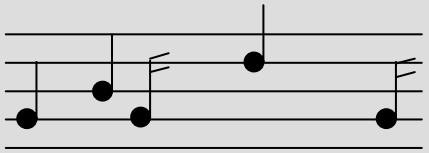
| | | |
|---|---|---|
| 1. | Systems | Transmission format |
| 2. | DDL | XML Schema |
| 3. | Audio | Audio MD |
| 4. | Visual | Visual MD |
| 5. | Multimedia DS | Structural MD |
| 6. | Reference Software | Open source soft |
| 7. | Conformance | Methods to test |
| 8. | Extraction and use of MPEG-7 descriptions | Algorithms and methods |
| 9. | MPEG-7 profiles and levels | Complexity reduction |

# Part3 – Audio
# Some description tools

- Sound effects
  - Indexing and categorization of general sound effects
- Musical Instrument Timbre
  - Perceptual features of instrument sounds



- Spoken Content
  - Word and phone lattices for each speaker
- Melody Contour
  - Compact representation of melody

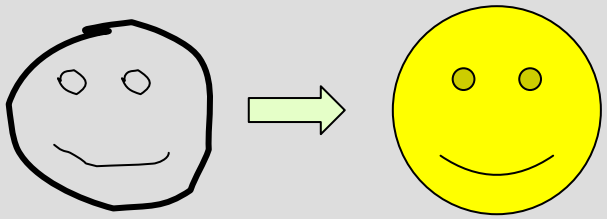


- Silence
  - Attach silence semantics to an audio segment

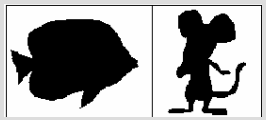
# Part 4 – Visual
# Some description tools

- Colors
  - Filter images by colors, ambiance, ..

- Texture
  - Distinguish clouds, walls, grass, ...
- Edges
  - Targets image-to-sketch matching

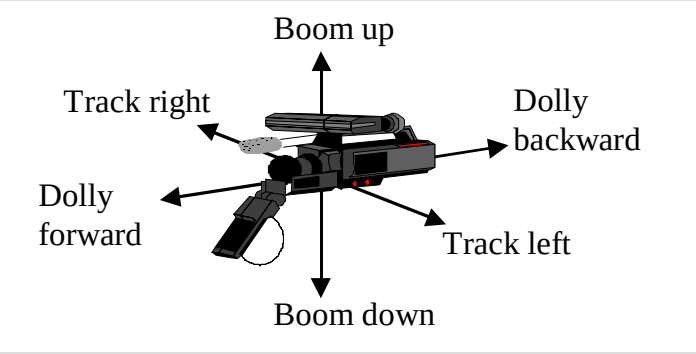


- Shapes
  - Describe visual object shapes



# Part 4 – Visual Other description tools

- Camera Motion
  - Pan, tilt, boom, track, dolly
- Motion Trajectory
  - Precise localization in time and space of an object
- Parametric Motion
  - Description of the motion of an object
- Motion Activity
  - Retrieve high speed car chase, interview, ...




# TV-Anytime

# The TV-Anytime Forum

- **Aim**: produce standards to **enable Personal Video Recorders (PVR)**

- **TVA recommendation on 4 main areas**
  - **Metadata**
    - For program descriptions, EPGs, and other information
    - **To help the viewer choose before acquiring content**
  - **Content Referencing**
    - Mechanisms for locating, tracking and record content
    - **To help the PVR to effectively acquire the content**
  - **System aspects**
    - Mechanisms for efficient transmission & processing of data
    - **To minimize network, CPU and memory ressources**
  - **Content Rights and Home Networking**
    - On going work known as TVA Phase II

# TVA Scenarios – base line

- A TV-Anytime scenario follow these steps:

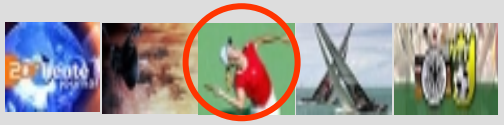
1. Based on the metadata, the viewer (or a user agent) selects a content to record. The result of this phase is a CRID that identifies the content

2. The CRID is resolved into a locator or a set of locators that identifiy where and when the content is available (DVB, IP, ...)

3. The device uses the locator(s) to record the content

# Ex. of TV-Anytime Services

- Broadcast EPG
  - Service Information
  - Programme Location
  - Programme Information
- Video-on-Demand
  - OnDemandProgram / service
- Advanced Applications
  - Group recording / trailer
  - Highlights viewing
  - Virtual Programmes
  - Indexes and chapters
    - e.g. for non linear navigation
  - Insertion of pre-recorded / cached related content
    - e.g. Target advertising / User Preferences / History

# TV-Anytime around the world

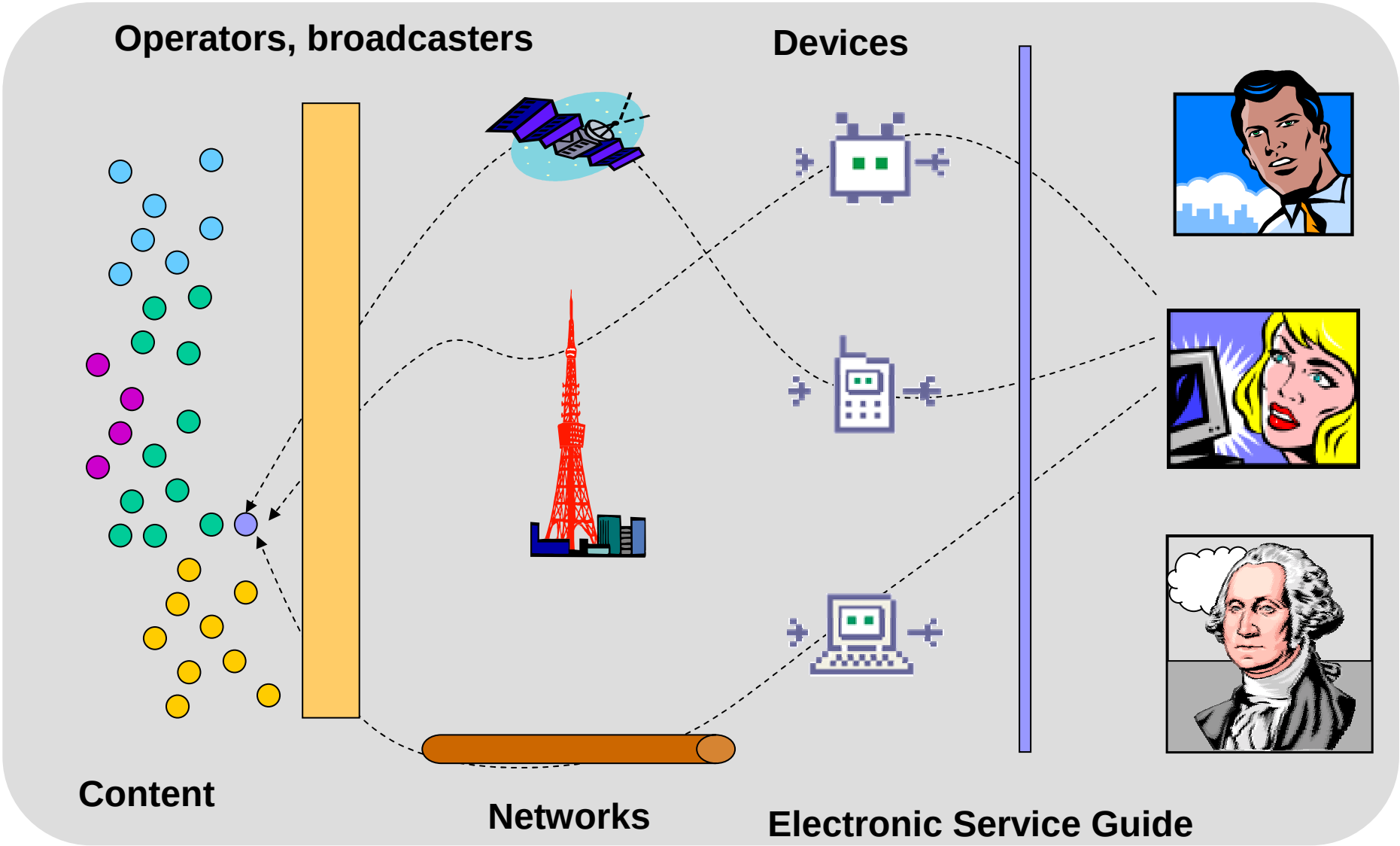
- DVB - Europe
  - DVB GBS: transport specification finalised
  - **UK DTG TVA Test Bed**: target for **live Service by Q1 2005**
    - EXPWAY currently working with 2 major STB manufacturers, 2 middleware editors and 2 head end tool providers
- ARIB - Japan
  - ARIB-STD B38: Coding, Transmission and Storage Specification for Broadcasting System Based on Home Servers
  - Several experiments:
    - **Tokyo Pilot** & **eConvergence** (NHK, NTT, Mitsubishi, Hitachi, etc.), Mobile DTT - **CRID Akasaka** (TBS, NTT, Hakuhodo, EXPWAY)
- Several experiments done by vertical operators

# Metadata for end users

About EPG, ESG and ECG

Stakes and challenges

# What, where, when, how, how much ??



# Electronic Service Guides

- With the convergence of networks and services
- Operators and Broadcasters face new challenges
  - Business models
    - Revisit their advertisements policy
  - Ease user choices in the service/content offer
    - Access to the most appropriate service
    - Provide personalized access
  - Manage heterogeneity of services, networks and terminals
  - Manage & market their service offers
- Electronic Service Guide (ESG) is the direct link between the operator, the broadcaster, the content owner and the user
  - Fast and smart metadata management is the enabling technology

# Electronic Service Guide

- Allows to
  - Describe the service or content
  - Attach / link media content to it
  - Transmit the service description
  - Help locate the content
  - Market and collect consumption usage
  - Infer user preferences to promote the proper content
  - Insert advertisements / targeted advertisements
- While, satisfying the end user
  - Privacy
  - Quality of service
  - Awareness of consumption

# Examples of ESG applications

- **Electronic TV-Program guide**
  - Receive every morning the TV-Program guide for the next days
  - Select, find your favorite program
  - Download trailers, related material
  - View advertisements
  - Record video clips/ Send email to your PVR
- **Video clips electronic guide**
  - Access to a large video clips database
    - sport events, music, news, ..
  - Find the clips that best suites to your taste / preview
  - Download it
  - Send an email to your friend to recommend it
- **Targeted and synchronized advertisements**
  - View your favorite show
  - Synchronized promotion and advertisements
  - Targeted to your taste

# Technical challenges

- Technical challenges
  - XML metadata are voluminous
  - XML processing is memory and CPU intensive
  - Various Transmission models
  - Manage server side/client side processing balance

- TV-Anytime, DVB, MPEG-7 uses
  - An efficient XML binary encoding format
  - Metadata fragmentation

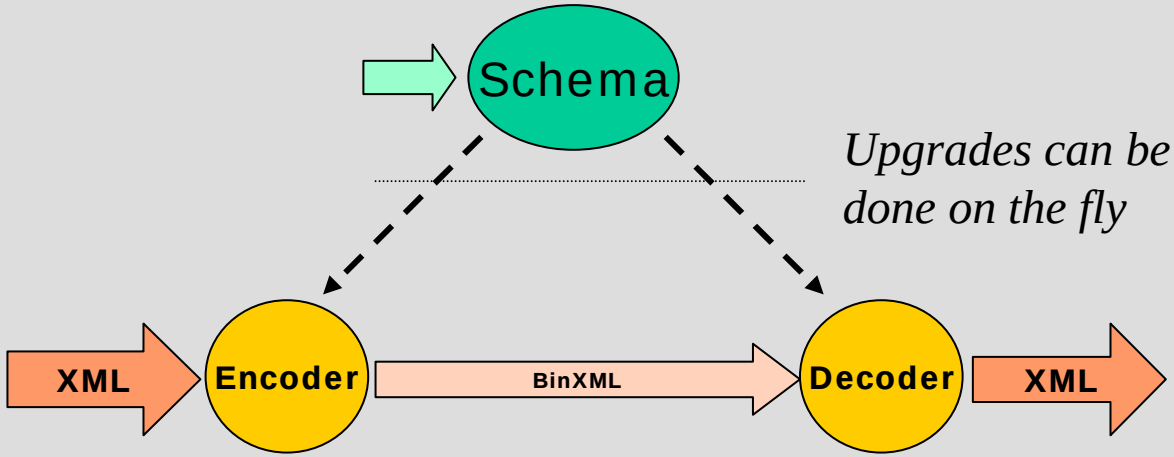
EXPWAY
MAKING XML EFFICIENT™

# MPEG-7 Standard

| | | |
|---|---|---|
| 1. | Systems | Transmission format |
| 2. | DDL | XML Schema |
| 3. | Audio | Audio MD |
| 4. | Visual | Visual MD |
| 5. | Multimedia DS | Structural MD |
| 6. | Reference Software | Open source soft |
| 7. | Conformance | Methods to test |
| 8. | Extraction and use of MPEG-7 descriptions | Algorithms and methods |
| 9. | MPEG-7 profiles and levels | Complexity reduction |

# BiM - Schema analysis

- XML Schema grammar analysis automatically generates encoding/decoding rules

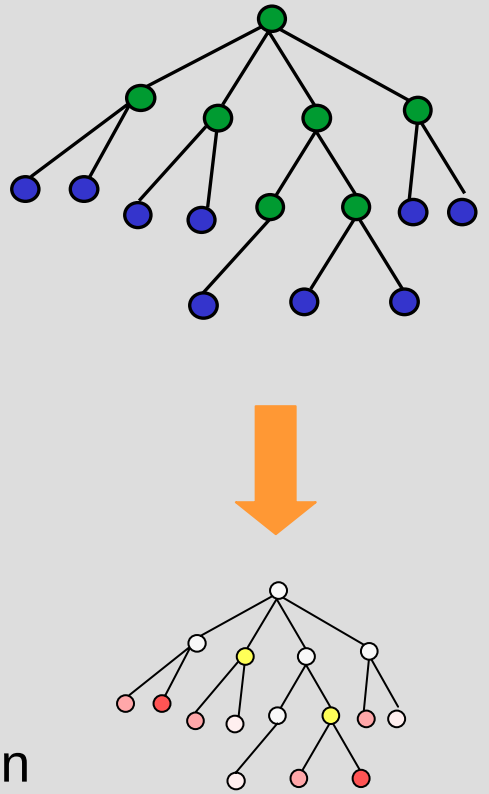



➔ Avoid to deal both with XML performance and flexibility

➔ Avoid to develop ad-hoc, stagnant binary formats

# BiM - Encoding methods

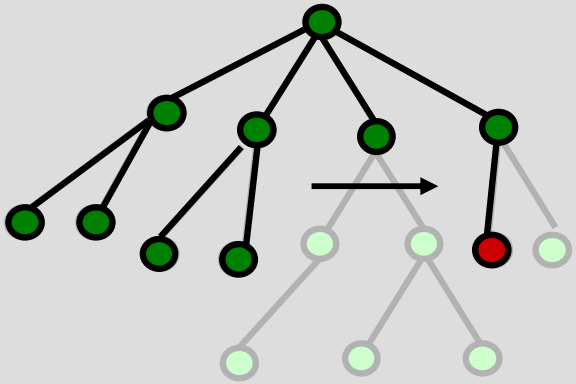
- Use of dedicated encoding methods
  - XML structure
    - Finite State Automaton based compilation
      - Simple, efficient, scalable
      - Enables validation & provides compression
  - XML content
    - Efficient standard compression methods
      - Statistical, Quantization, Dictionary, …
    - Standard International encoding formats
      - IEEE-754, UTF-8, UTF-16, …



➔ Intelligent adaptive and customizable compression

# BiM - Partial Decoding

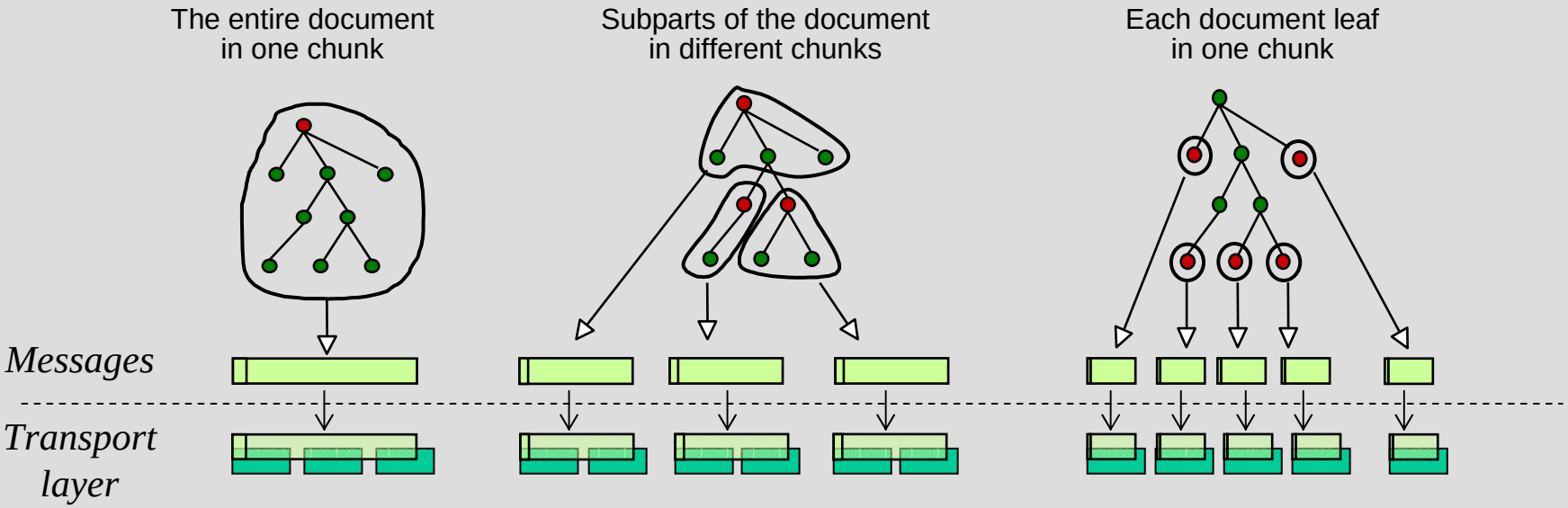
- An application can skip the decoding of some document subparts



➔ Parsing / filtering speed improvements

# BiM - Flexible Transmission

- Document can be cut into several pieces transmitted independently




➔ Optimize message mapping to the underlying transport
  ➔ Reduce processing latency
➔ Progressive & delta transmission
➔ Fragment based processing (indexing, encryption, ..)

# Conclusion

# AV MD Standards evolution

- MPEG-7 is defining simple profiles
  - Simple Metadata Profile
    - simple metadata tagging for single instances of multimedia clip
    - Mapping from/to ID3 and EXIF
  - User Description Profile
    - to describe the personal preferences and usage patterns of users of multimedia content
  - Core Description Profile
    - to describe general multimedia content such as images, videos, audio, and collections.
- TV-Anytime is targetting new contents and new business models
  - content types: games, interactive app
  - devices: Mobile phones, PDA
  - storage: DVD-R
  - Content sharing

# Opportunities

- DVB & ARIB framework have adopted TV-Anytime
  - DVB-GBS / DVB-IPI
- Practical implementation of metadata standards has started
  - UK DTG Testbed using TV-Anytime/DVB (BBC)
  - eConvergence project in Japan (NTT / NHK)
- Technical solutions are mature enough
  - XML → flexibility
  - BiM → efficiency
- And also
  - Metadata providers already generate TV-Anytime
  - Storing content is increasing
    - PVRs / iPod / mobile phones
  - Datacast distribution is expected to gain over UMTS
    - DVB-H / SDMB

# Expway
## Mission & Business

- Mission
  - Faciliter **l'accès au contenu** quel que soit le type de contenu, l'environnement ou les terminaux en proposant des solutions innovantes, évolutives et standards.
- Produits
  - Expway développe et distribue des **solutions logicielles d'ESG** permettant la valorisation des bases de contenus multimédia sur tout type de réseaux et de terminaux.
- Marchés
  - Télévision Numérique
  - Téléphonie mobile
  - Editeurs de solutions logicielles

# Enfin la fin

- Merci de votre attention
- Plus d'information
  - MPEG : http://www.cselt.it/mpeg
  - TV-Anytime : http://www.tv-anytime.org
  - BiM et BinXML: http://www.expway.com
- Merci à Philippe Salembier pour son aide sur la présentation MPEG-7

Claude Seyrat
Expway, CTO
claude.seyrat@expway.fr

# Expway
## Mission & Business

- Mission
  - Faciliter **l'accès au contenu** quel que soit le type de contenu, l'environnement ou les terminaux en proposant des solutions innovantes, évolutives et standards.
- Produits
  - Expway développe et distribue des **solutions logicielles d'ESG** permettant la valorisation des bases de contenus multimédia sur tout type de réseaux et de terminaux.
- Marchés
  - Télévision Numérique
  - Téléphonie mobile
  - Editeurs de solutions logicielles

# ANNEXES

# La TV personnalisée - quelques chiffres

*Sources TiVo & replay TV*

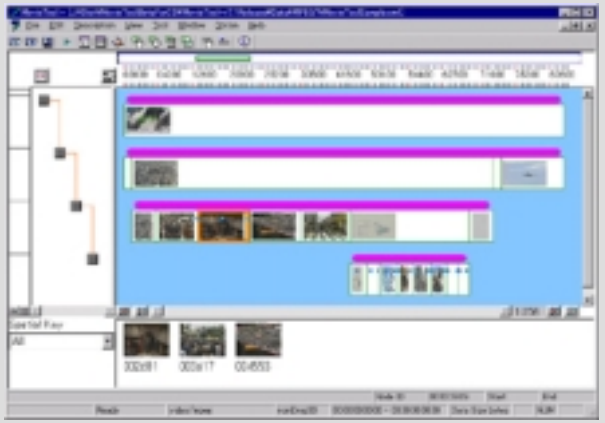
- Danger principal
  - Modèle de revenu actuel fondé sur la publicité
    - Avec les PDRs, 88% des publicités ne sont pas regardées
- Mais
  - 6 utilisateurs sur 10 de TiVo regardent plus la TV
  - En moyenne les utilisateurs regardent 3 heures de plus par semaine (ReplayTV)
  - 60% des utilisateurs regardent des séries qui n'étaient pas visibles à cause des horaires peu favorables
  - 1/3 des utilisateurs de TiVo considèrent la TV comme leur source essentielle de loisir (2* plus que ceux sans TiVo)
  - La location de films représente 2*plus de revenu que le cinéma
- De nouveaux modèles de revenus doivent être imaginés
  - Des publicités plus pertinentes, personnalisées
  - Des événements live payants
  - Des services de location / chargement à la demande

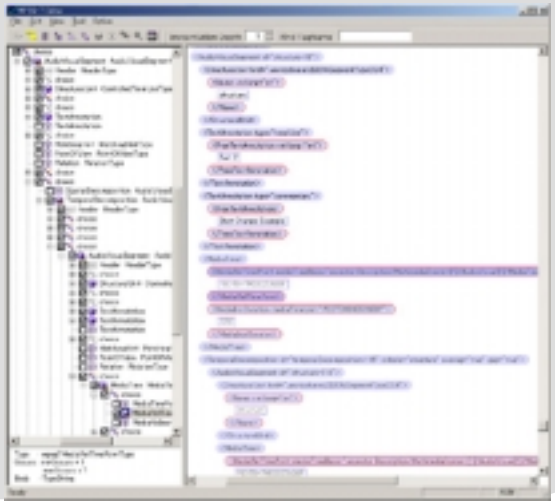
# Exemples d'application MPEG-7

# Outils d'annotation MPEG-7

## IBM - VideoAnnEx Annotation Tool


## Ricoh – Movie Tool

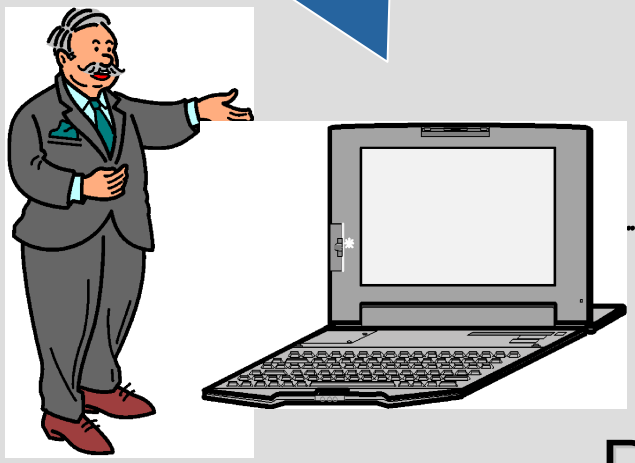




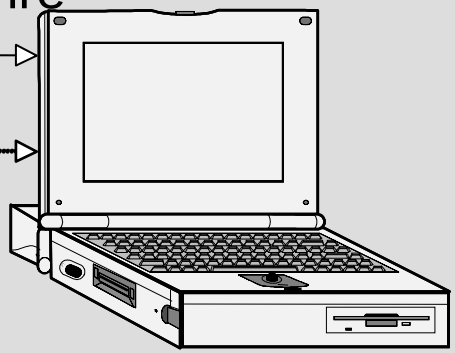
# Création automatique de présentations indexées

Presenter:
- Make the Presentation as usual
- Recorded the Presenter's Action
- One-person control is possible

Recording Operator:
- Direct HDD Recording
- Automatic content transfer and Web content generate within 1 min.
- Content Retrieval is possible using MPEG-7



FireWire



Wifi

PC Presentation

Ricoh - MPMeister

PC Capture

# Autres applications MPEG-7

- Fraunhofer
  - Technologie MPEG-7 d'identification de morceaux de musique
- EPFL
  - Camera MPEG-7
    - Détection de mouvements / formes naturelles
- NTT Docomo (projet de recherche)
  - Distribution de contenus personnalisés sur téléphone portable 3G
    - La description MPEG-7 des vidéos est transmise
    - Le téléphone retourne une scénario personnalisé
    - La vidéo est générée et streamée au terminal