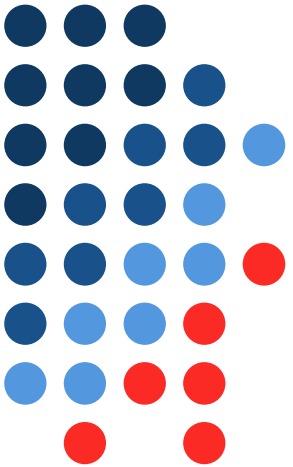
From video information retrieval to hypervideo management

Olivier Aubert
Yannick Prié

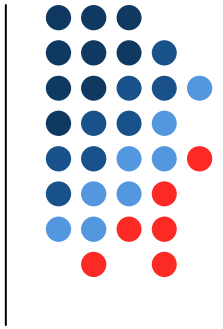
CORIMEDIA

International workshop on multidisciplinary image, video and audio retrieval and mining
25-26 oct 2004 – Sherbrooke



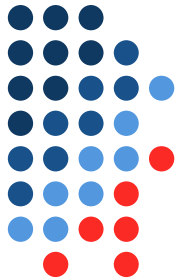
LIRIS
Lyon Research Center for Images and Intelligent Information Systems
FRE 2672 – CNRS/INSA-Lyon/Université Lyon 1/Université Lyon 2/Centrale Lyon

Overview



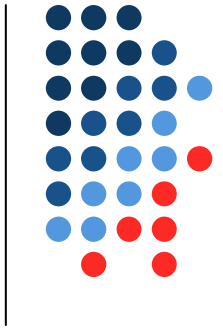
- Audiovisual information systems and information retrieval
- Hypervideos
- The advene model and application
- Discussion

AVIS
audiovisual information systems



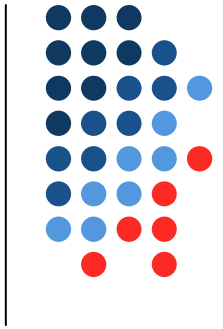
- Several activities...
  - Video indexing and retrieval
    - high / medium level features
  - Query results selection
    - video skim, surrogates
  - Query results exploitation
    - watch, reuse video
- ... different tools and descriptors

AV information usage



- Video usage
  - from simple visualisation to video reuse in other documents
  - retrieval is determined by video usage
- We should
  - seek for interoperability of systems
  - integrate video usage in AVIS
- cf. Web innovation
  - easiness of use
  - interoperability of tools
    - document visualisation, search, refering, manipulation, etc.

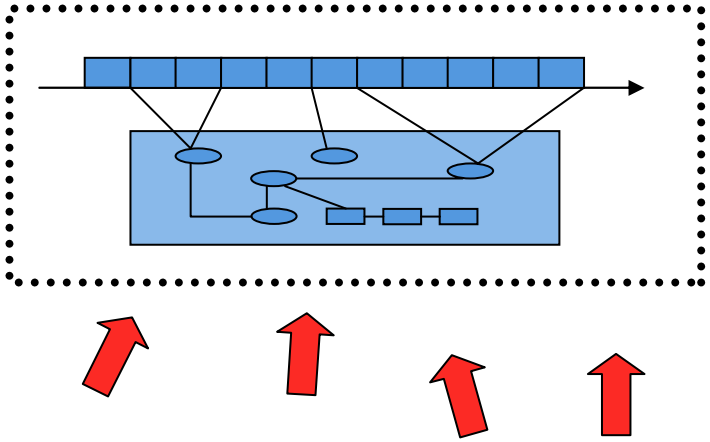
AV descriptors

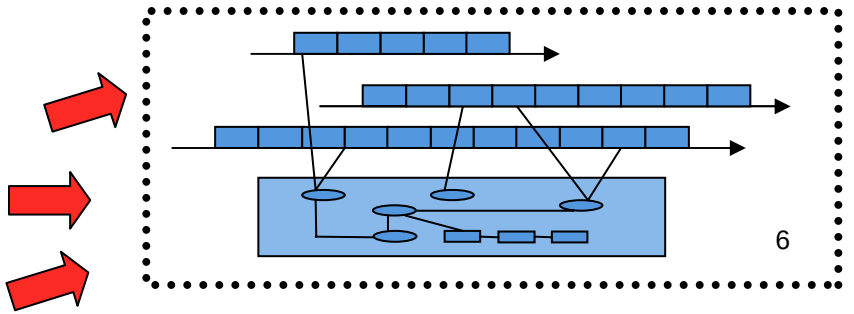


- Video documents do not provide minimal AV units such as characters and words
- AV descriptors are the key to
  - retrieval
  - visualization
- We need to find useful descriptors
- A way of doing it : build systems
  - that are fully descriptor-based
  - that integrate AV information usage as a whole
  - that facilitate the emergence of new ideas

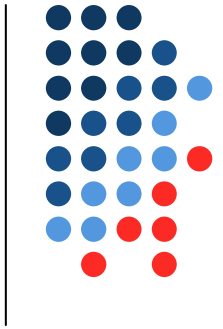
Hypervideos

- Annotated AV documents
  - AV document
  - *plus* annotation structure
- View ➡
  - « way of visualizing » an annotated AV document
- Hypervideo
  - A view that uses information from both the document and the annotation structure, giving access to the document as a stream
  - Hyper → random access
  - Video → temporal stream
- Generalization for several AV documents



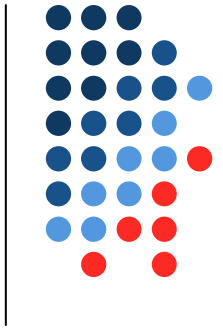


Views and hypervideos: some examples



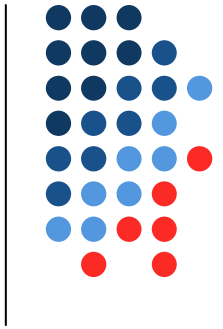
- Annotated AV document
  - Movie + shot/sequence decomposition
- Non-hypervideo views
  - The movie itself
  - Table of contents (with or without images from the movie)
- Hypervideo views
  - Movie with
    - shot number inserted on the stream
    - link to « next shot » / « next sequence »
  - Shot/sequence tree-view with links to the film at each shot transition
- Hypervideos in the wild

Advene



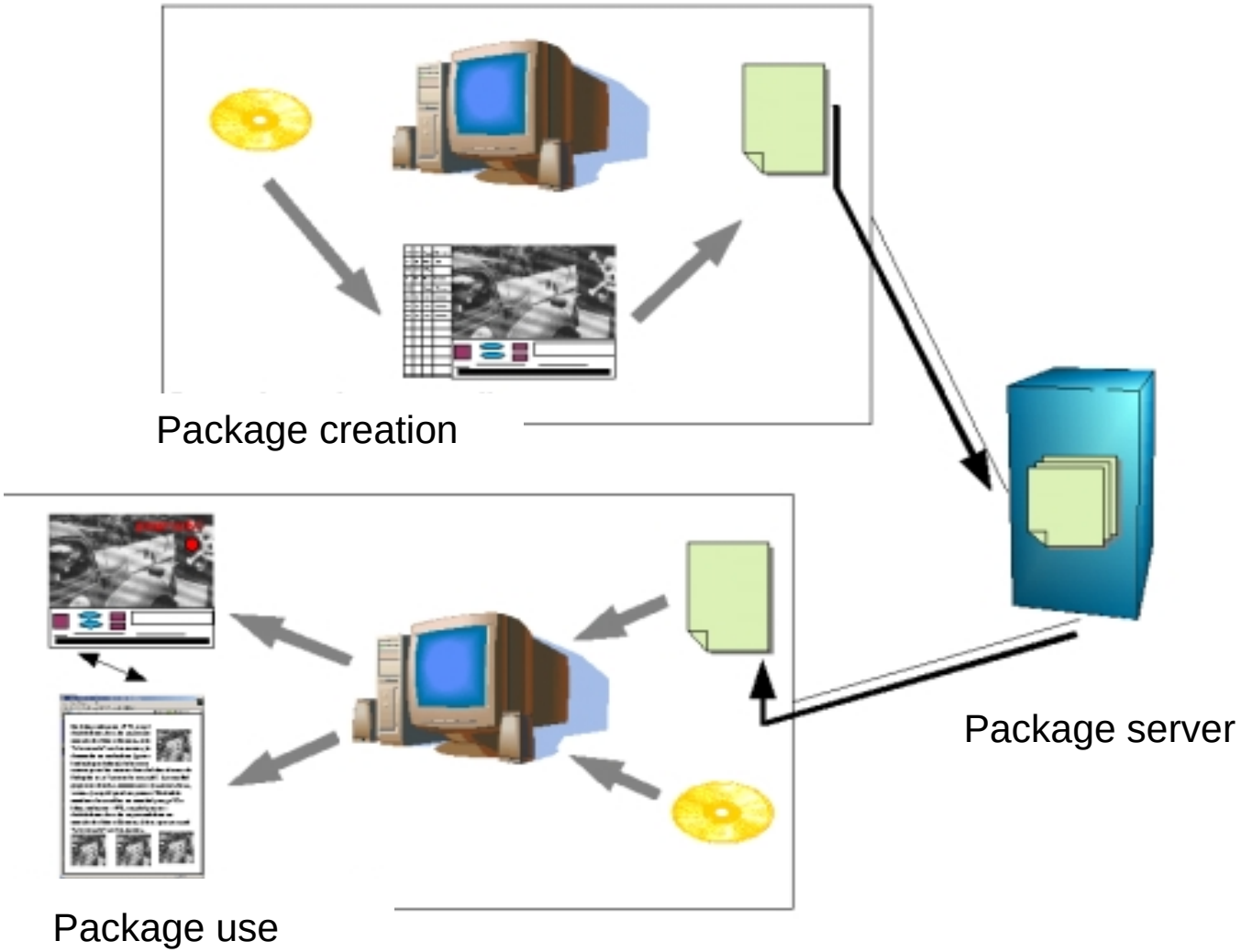
- « advene: to be added to something or become a part of it, though not essential » (Webster 1913)
- Advene: Annotate DVds, Exchange on the Net
- Objectives
  - Be fully descriptor-based
  - Allow new usages for audiovisual documents amongst different communities
    - Cinema / language teachers
    - Humanities scientists
    - Movies goers (sharing comments)
  - Better define and study hypervideos
    - Graphical interfaces
    - Documents and sharing

Advene
principles



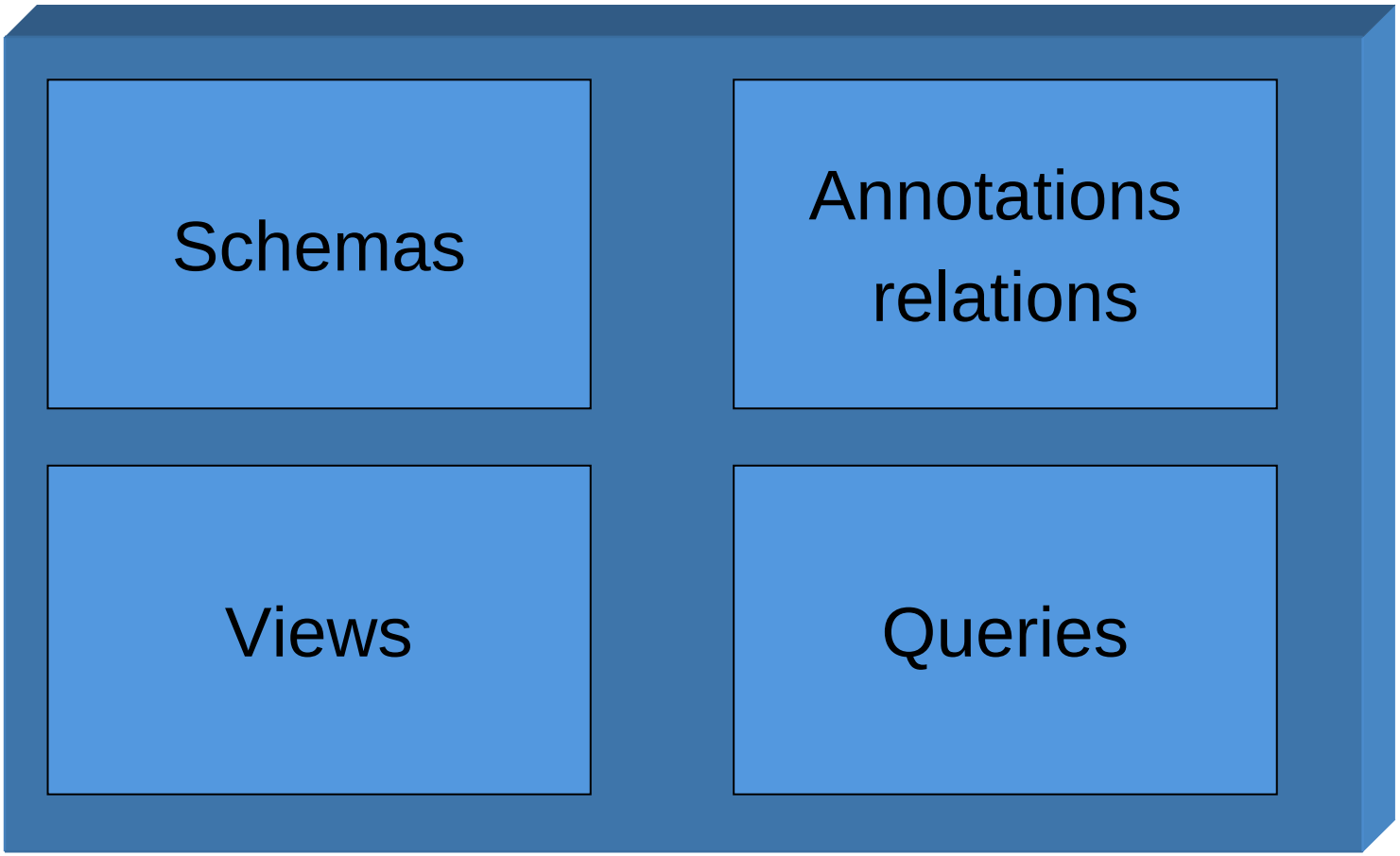
- Objective
  - facilitate the development of new ideas and usages
- Development principles
  - simplicity : wide audience
  - open source : movie-goers community developpers
- AV Documents (films) on DVD
  - alleviate rights problems
  - wide accessible corpus
- Package
  - Self-contained document with annotation structure, annotation definitions, views.
  - Hypervideo generation from package + DVD

Advene overview

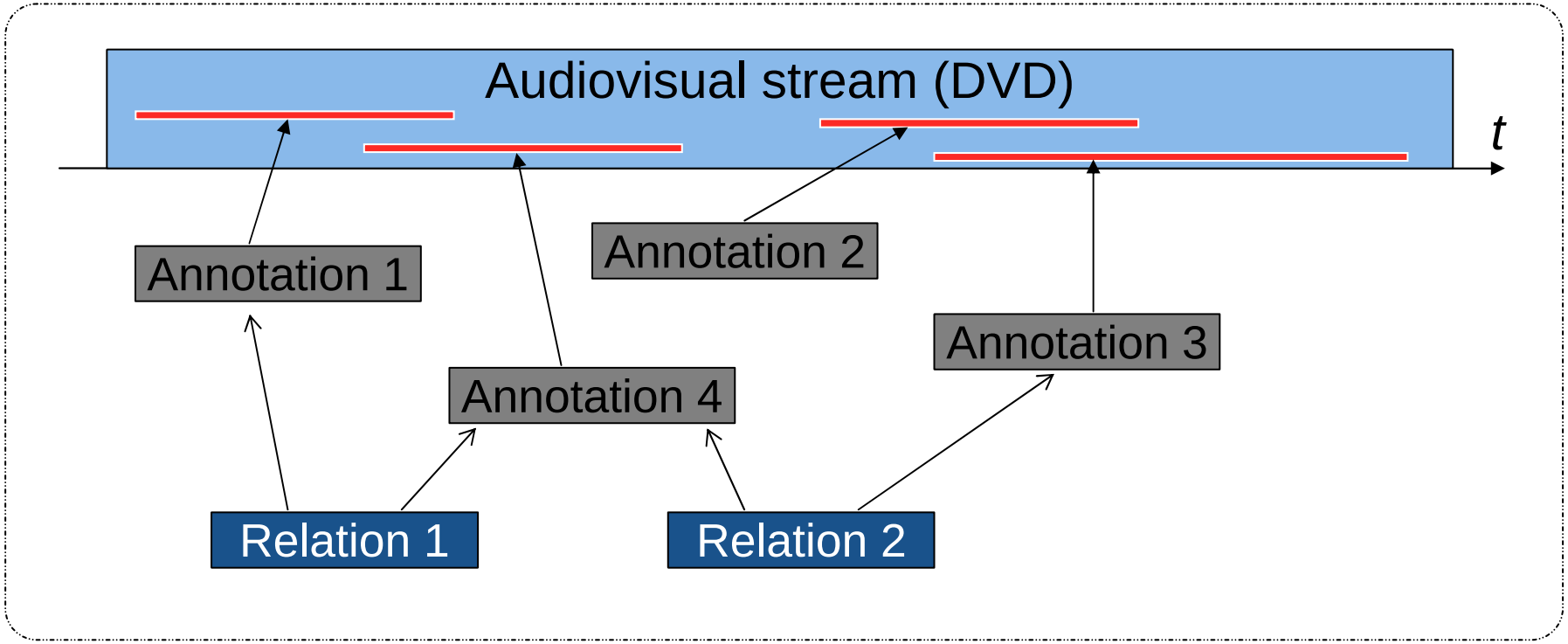


Advene

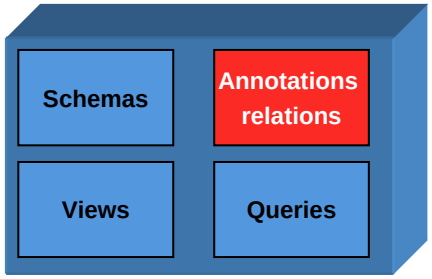
packages



Advene
annotated AV document

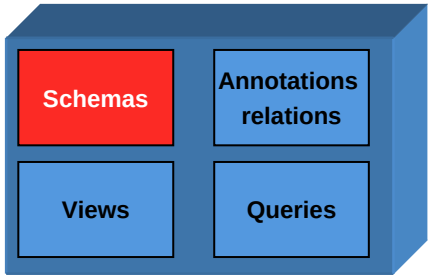


Advene
annotations and relations



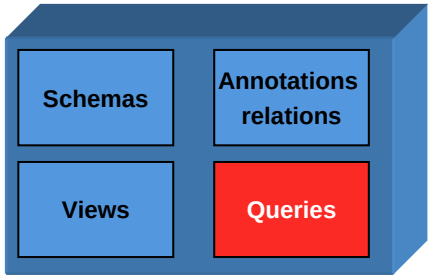
- Annotation
  - type (ex : shot, character, episode...)
  - content (text, image, sound...)
  - fragment (stream-id, temporal interval, other...)
  - context (subtitles, language, angle...)
- Relation
  - members (two or more annotations)
  - type
  - content (text, image, sound...)

Advene
schemas



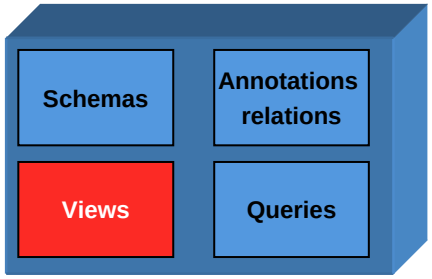
- Annotation type
  - name
  - constraints on the content
    - MIME types
    - Structured XML types (XML-Schema)
- Relation type
  - name
  - constraints on the types of the annotations belonging to the relation
  - constraints on the content (if needed)
- Schema
  - useful set of annotation and relation types
- Package imports
  - schema reuse
  - annotation / relation reuse

Advene queries



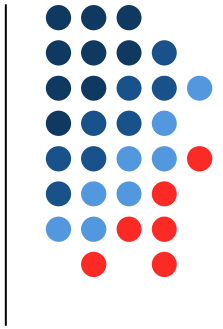
- Dynamic extraction of significant Advene elements (annotations, types, schemas,…)
  - Applies on all the elements of a package
  - The result is a set of items
- For the moment
  - Not a full-featured query language (limited set of conditions)
  - Rather an implementation of a filtering
    - Fragment duration, annotations content, etc.

Advene
views



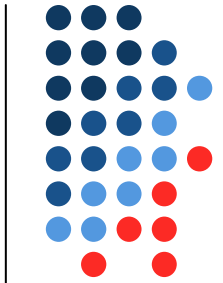
- Ways of visualizing and interacting with an element or set of elements from the annotated AV document
- Three main types in Advene
  - standard web navigator view (static view)
  - enhanced video player view
  - ad-hoc interface view
- Possibility to switch from one view to another

Advene
navigator views



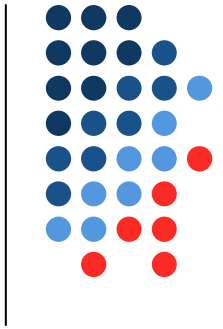
- Detemporalized views
  - UTBV (user-time based view)
  - Mainly images + text in XHTML
  - Images extracted from the stream
- Template attribute principles
  - X(HT)ML file
  - TAL Syntax
    - element replacement
    - iterations
  - TALES expressions
    - queries on the package, expressed as paths
    - Ex : /package/annotations/annotation121/type

Advene
augmented video player views



- Views in relation with a classical video player
  - Extended and automated actions
  - Control of the played stream
  - Augmentation of the stream
- First exploration in Advene
  - Rule-based model
    - ECA = event – condition – action
  - View = set of rules
- Example
  - If the player arrives to the beginning of a fragment associated with an annotation (event), if the annotation type is *character* (condition), then print on the stream ‘Character: character_name‘ (parameterized action)

Advene
ad-hoc views



- Complex views
  - Explicit programming
  - Apply to a set of package elements
  - Examples :
    - Timeline
    - Tree-like view
    - Text synchronised with video

File Edit View Web Go Bookmarks Tabs Help

Wallace and Gromit, in The Wrong Trousers - Exercises

Browser view

by Pierre-Antoine Champin

About the movie

Main characters

Wallace

Gromit

the penguin

the chicken

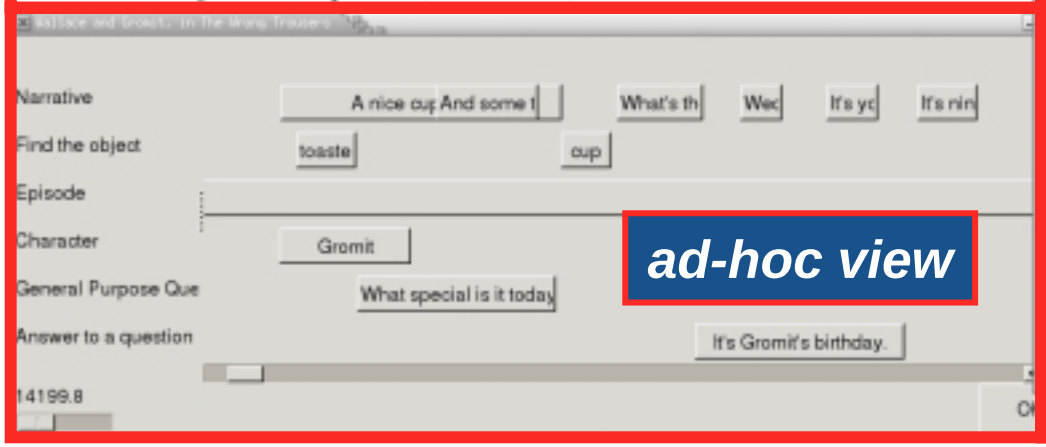
Vocabulary

Answer the questions

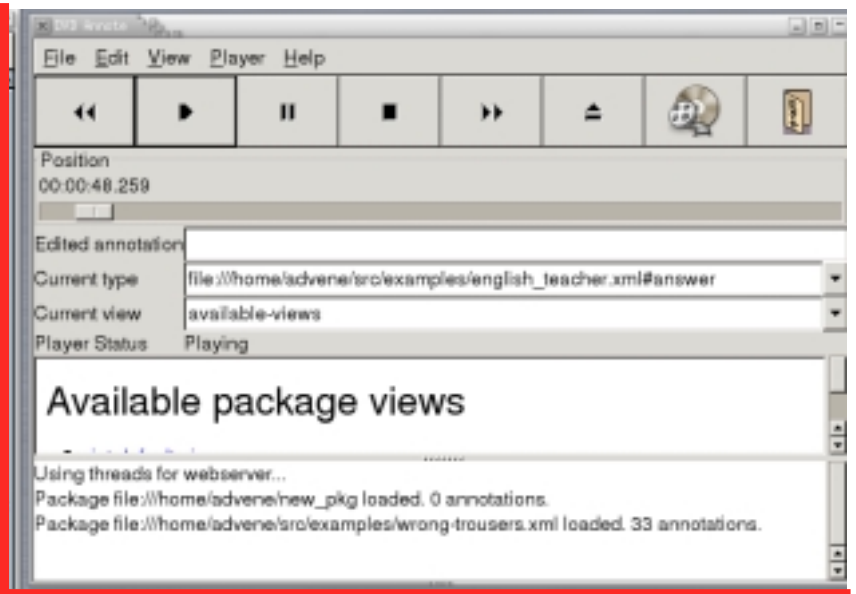
What special is it today? (00:00:23.827)

.....

Has Wallace bought Gromit a present? (00:00:49.237)

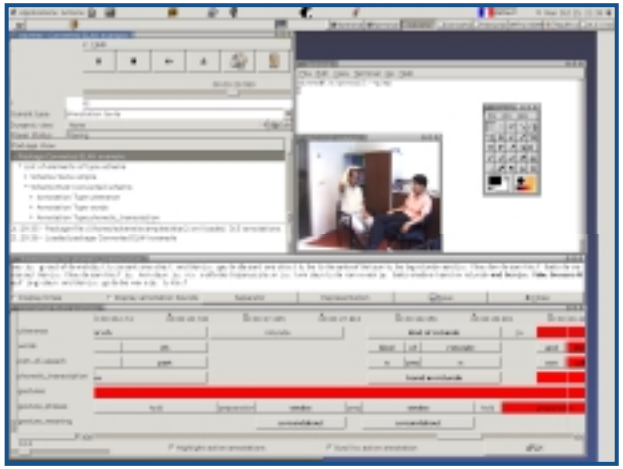


ad-hoc view

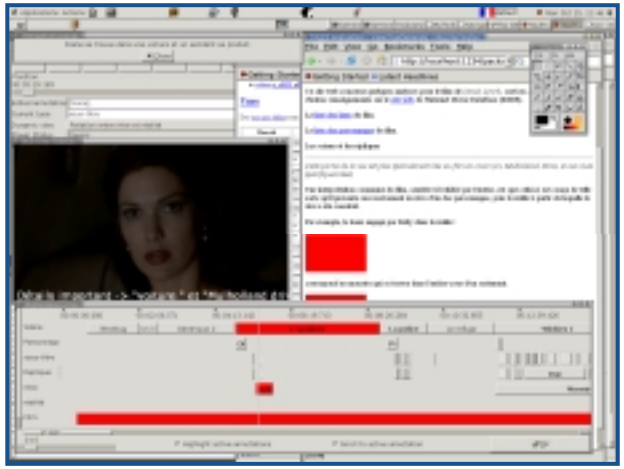


Augmented video player view

More examples

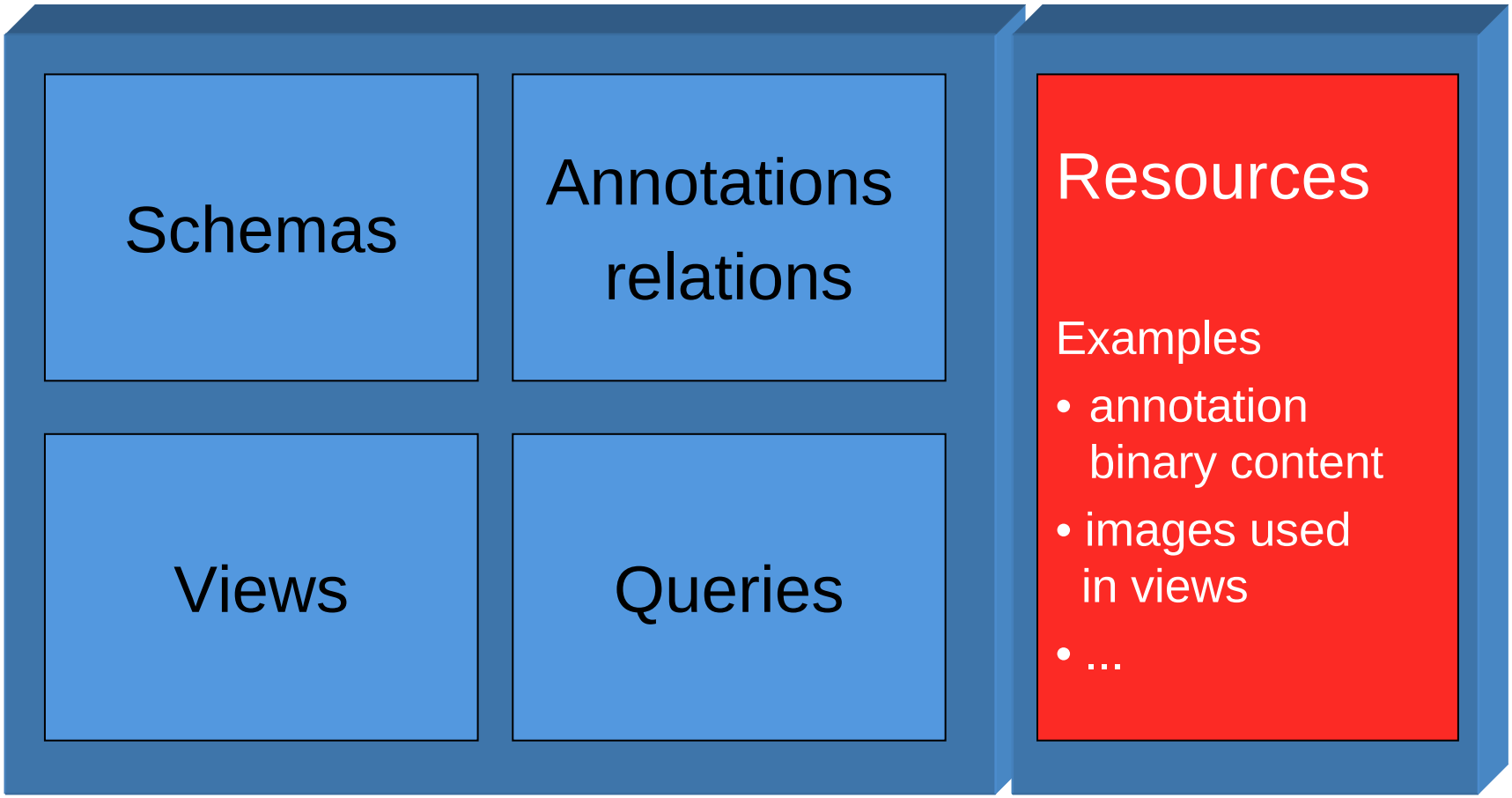


ELAN

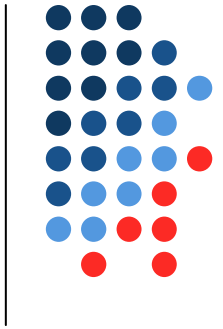


Mulholland drive

Advene
package related resources

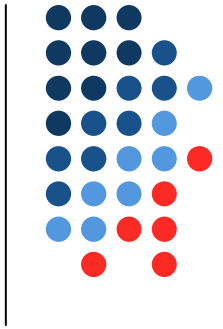


Conclusion



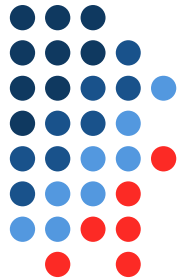
- AVIS : not only retrieval, but video usage
- Hypervideo is a concept for considering new usages
- Video usage is based on descriptors
- We need to build integrated, descriptor-based systems that facilitate the emergence of new usages/descriptors/interfaces
- The Advene project aims at it, with
  - Minimal modeling
  - Simplicity / extensibility
  - Easiness to integrate exterior knowledge (ontology, thesaurus)
  - Interaction with existing video retrieval tools.

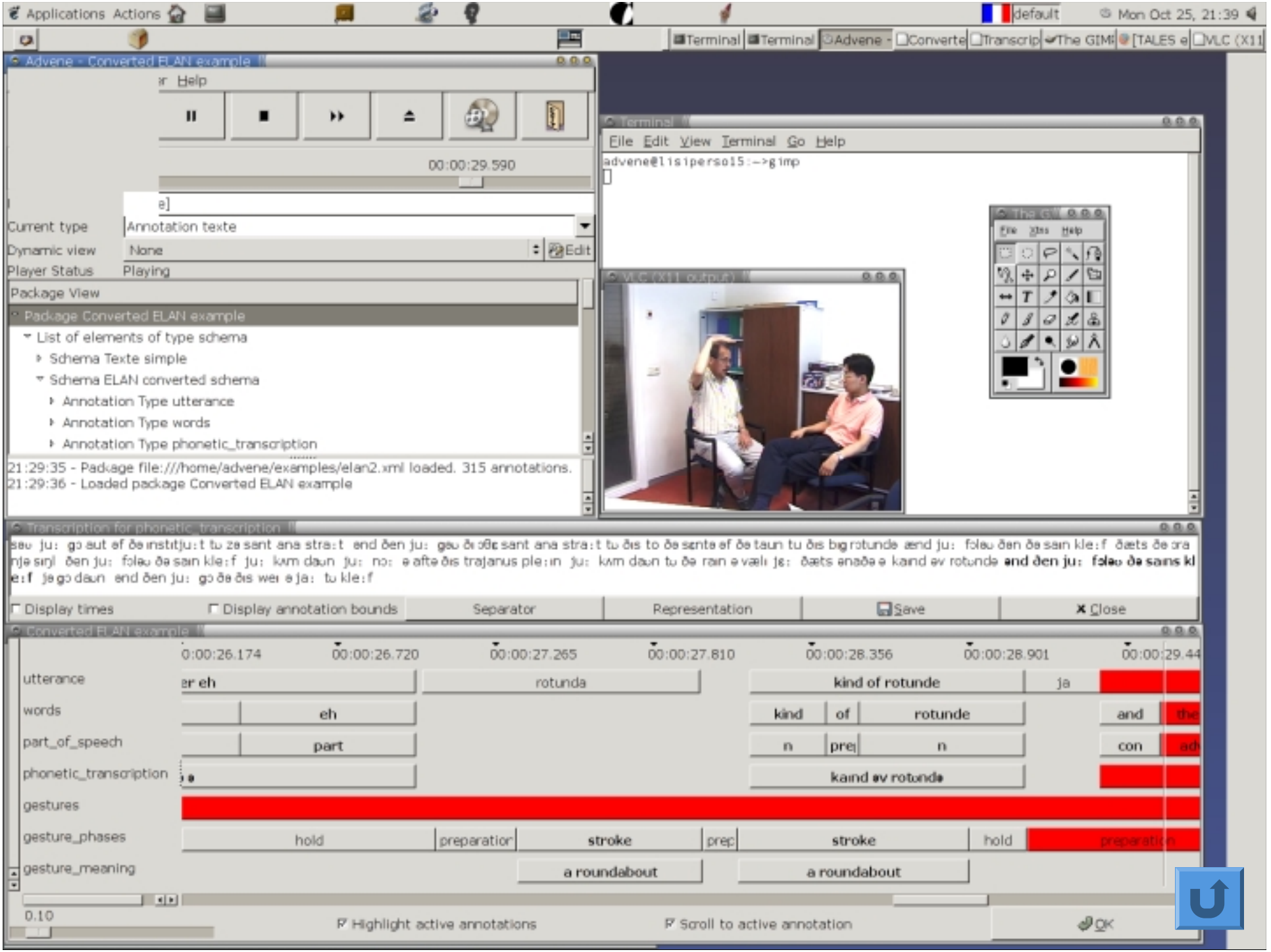
Ongoing work

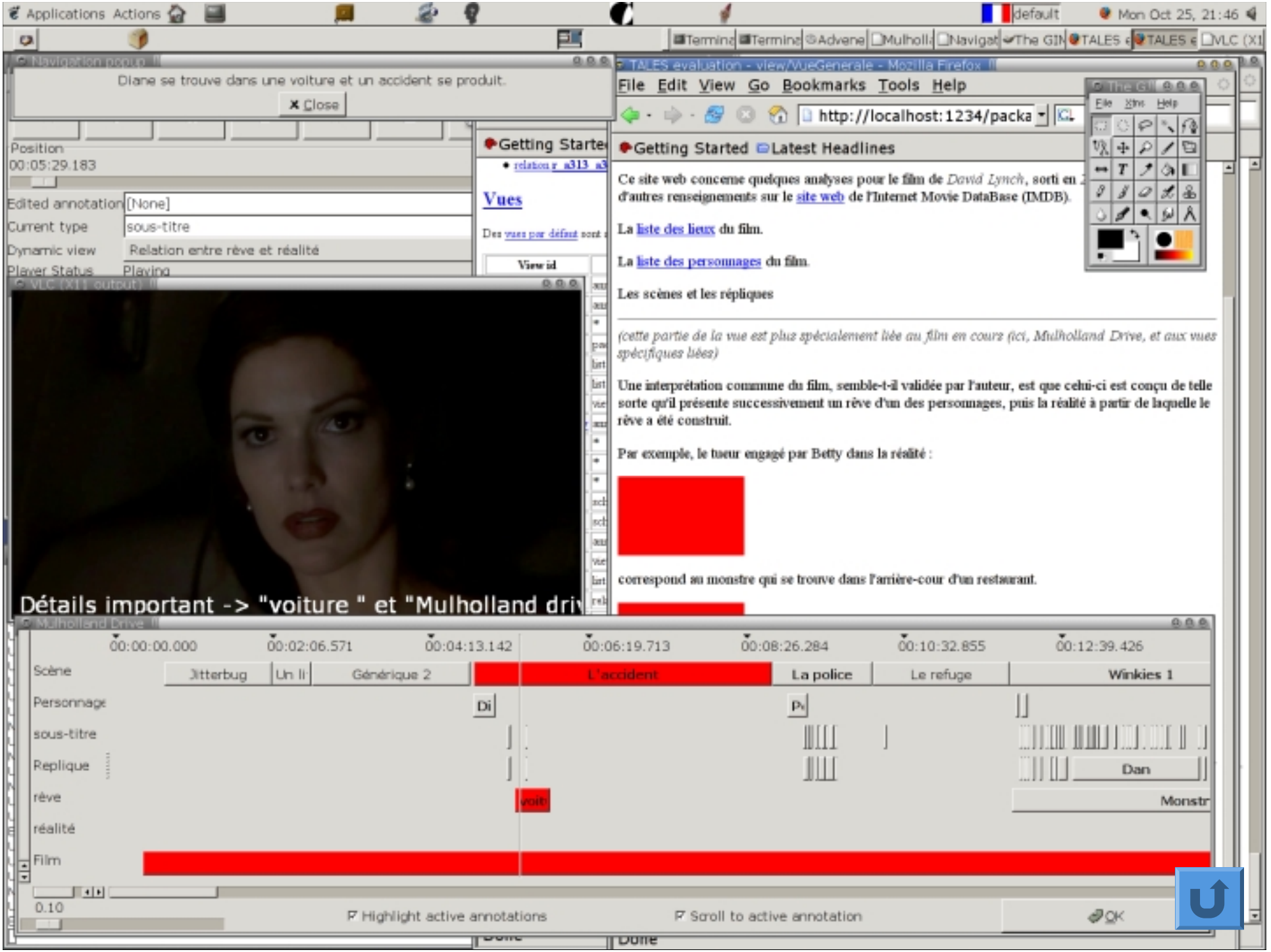


- Theoretical analysis
  - Audiovisual temporalities, detemporalisations
  - Hypervideo perception: cognitive studies and interfaces
  - Document theory (package, film references, etc.)
- Prototype
  - Finalization for open-source release
  - Various developpements
- Projects
  - DVD oriented project (cinema teaching)
  - Humanities oriented project (interaction analysis)
- We need money!
  - piloting the Open Source development

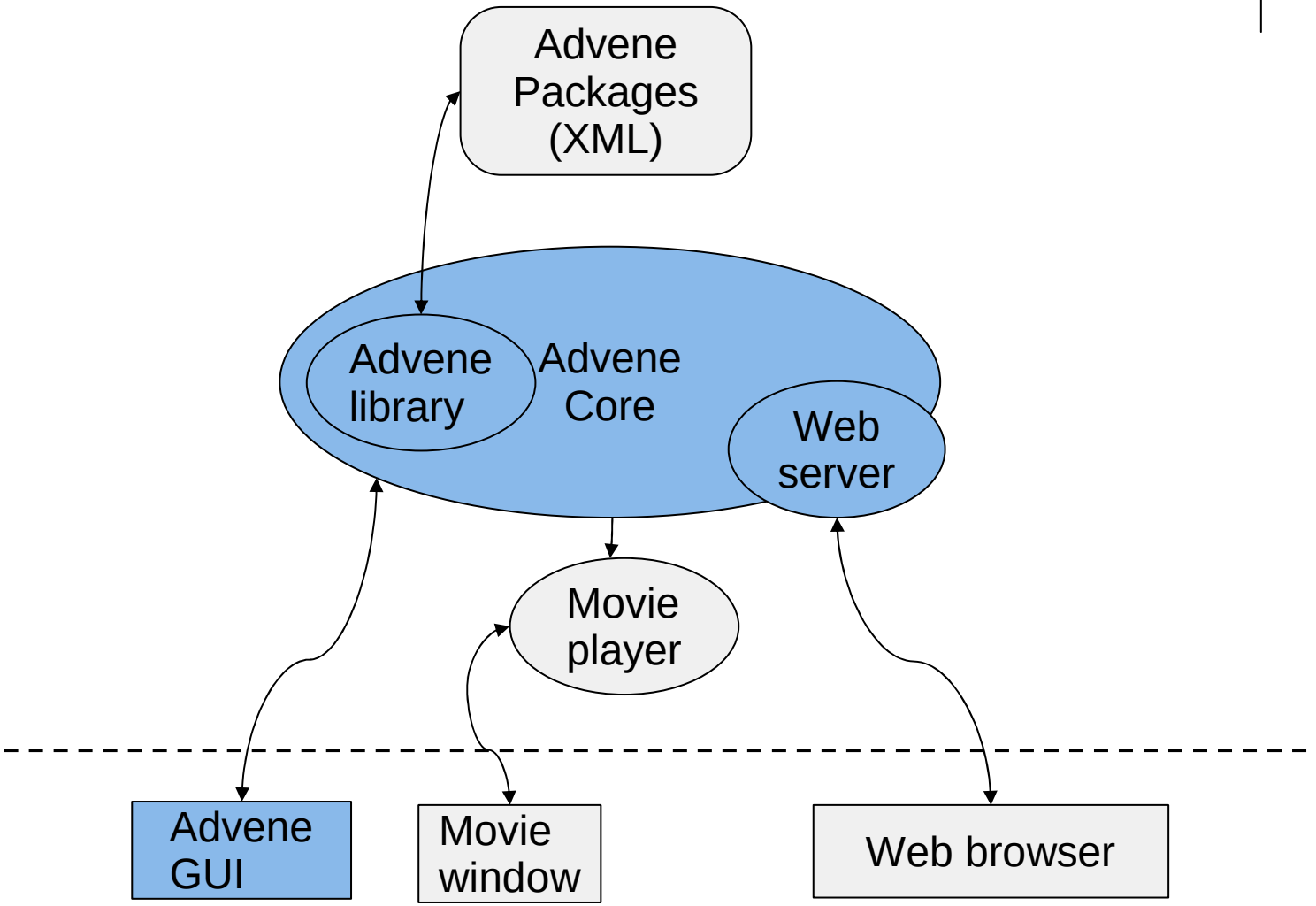
Thank you...



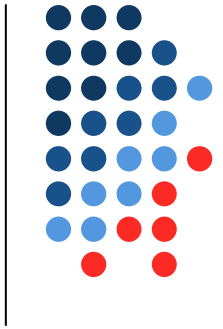

Applications Actions
default
Mon Oct 25, 21:39
Terminal
Terminal
Advene
Converte
Transcrip
The GIMP
[TALES e
VLC (X11
Advene - Converted ELAN example
Help
00:00:29.590
Current type
Annotation texte
Dynamic view
None
Edit
Player Status
Playing
Package View
Package Converted ELAN example
List of elements of type schema
Schema Texte simple
Schema ELAN converted schema
Annotation Type utterance
Annotation Type words
Annotation Type phonetic_transcription
21:29:35 - Package file:///home/advene/examples/elan2.xml loaded. 315 annotations.
21:29:36 - Loaded package Converted ELAN example
Terminal
File Edit View Terminal Go Help
advene@lisiperso15:~>gimp
VLC (X11 output)
Transcription for phonetic_transcription
Display times
Display annotation bounds
Separator
Representation
Save
Close
Converted ELAN example
00:00:26.174
00:00:26.720
00:00:27.265
00:00:27.810
00:00:28.356
00:00:28.901
00:00:29.44
utterance
rotunda
kind of rotunde
ja
words
eh
kind
of
rotunde
and
part_of_speech
part
n
con
phonetic_transcription
kaind ev rotunde
gestures
gesture_phases
hold
preparation
stroke
stroke
hold
gesture_meaning
a roundabout
a roundabout
0.10
Highlight active annotations
Scroll to active annotation
OK


Applications Actions
Mon Oct 25, 21:46
Diane se trouve dans une voiture et un accident se produit.
Close
Position
00:05:29.183
Edited annotation
[None]
Current type
sous-titre
Dynamic view
Relation entre rêve et réalité
Player Status
Playing
Vues
File Edit View Go Bookmarks Tools Help
http://localhost:1234/packa
Getting Started
Latest Headlines
Ce site web concerne quelques analyses pour le film de David Lynch, sorti en
d'autres renseignements sur le site web de l'Internet Movie DataBase (IMDB).
La liste des lieux du film.
La liste des personnages du film.
Les scènes et les répliques
(cette partie de la vue est plus spécialement liée au film en cours (ici, Mulholland Drive, et aux vues spécifiques liées)
Une interprétation commune du film, semble-t-il validée par l'auteur, est que celui-ci est conçu de telle sorte qu'il présente successivement un rêve d'un des personnages, puis la réalité à partir de laquelle le rêve a été construit.
Par exemple, le tueur engagé par Betty dans la réalité :
correspond au monstre qui se trouve dans l'arrière-cour d'un restaurant.
Détails important -> "voiture " et "Mulholland dri
00:00:00.000
00:02:06.571
00:04:13.142
00:06:19.713
00:08:26.284
00:10:32.855
00:12:39.426
Scène
Jitterbug
Générique 2
L'accident
La police
Le refuge
Winkies 1
Personnage
sous-titre
Replique
Dan
rêve
réalité
Film
0.10
Highlight active annotations
Scroll to active annotation
OK

Advene architecture

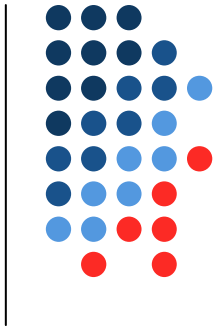


TAL / TALES
generalities



- Developped for the Zope application server
- TAL : Template Attribute Language
  - Objective : to describe view in X(HT)ML + iteration and replace instructions int the attributes
- TALES : TAL Expression Syntax
  - Web requests for accessing elements
  - Path-like expression (cf. URL)

TAL / TALES
example



TAL Template

```
<h1>Studying <em  tal:content='here/title'>The title</em></h1>
Annotations:
<ul>
  <li tal:repeat='annot here/annotations'>
    <strong tal:content='annot/content/data'>The text</strong>
    (duration:
     <span tal:replace='annot/fragment/formated/duration'>xxx</span>
     )  </li> </ul>
```

Template pre-visualisation

Studying *The title*

Annotations:
The text (duration : xxx)

Visualisation

Studying *The Wrong Trousers*

Annotations :
Episode one: Gromit's Birthday (duration: 00:00:02.123)
A nice cup of tea (duration: 00:00:01.453)
...