

Using Personal Annotations to Enrich Interactions with Videos

contact@olivieraubert.net
@Olivier_Aubert

Context

- Audiovisual annotation/metadata
 - How to bridge the semantic gap?
 - What do we need annotations for?
- 2 dimensions
 - Personal annotations
 - Document (hypervideo) building from annotations

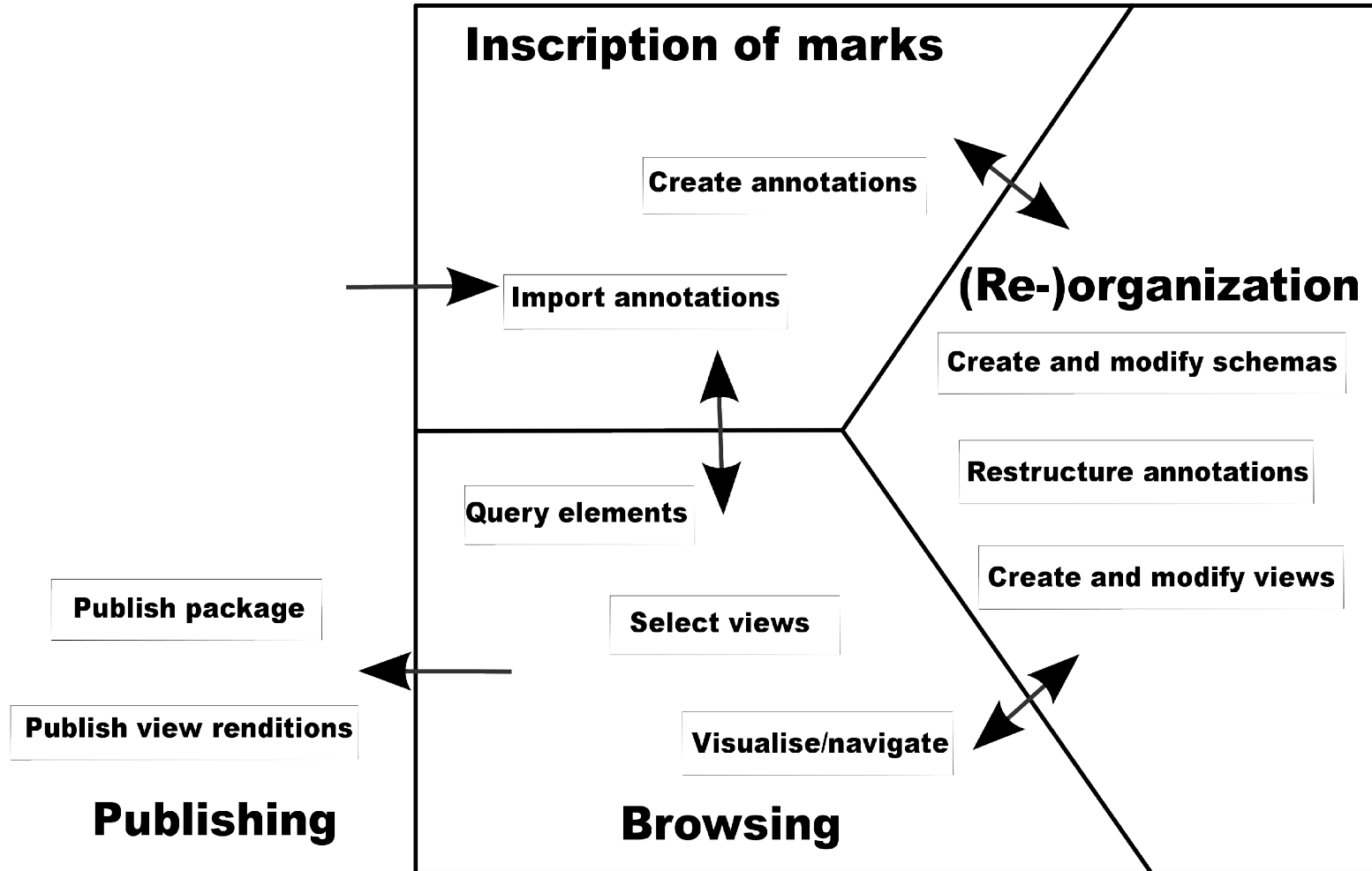
Summary

- Context – active reading
- Advene project
 - Use case: museographie.fr
- Visu project
- Conclusion

Intended audience

- Targets: scholars, movie critics...
- Willing to invest time to precisely annotate video documents
- Specific analyses
- Small-scale corpuses

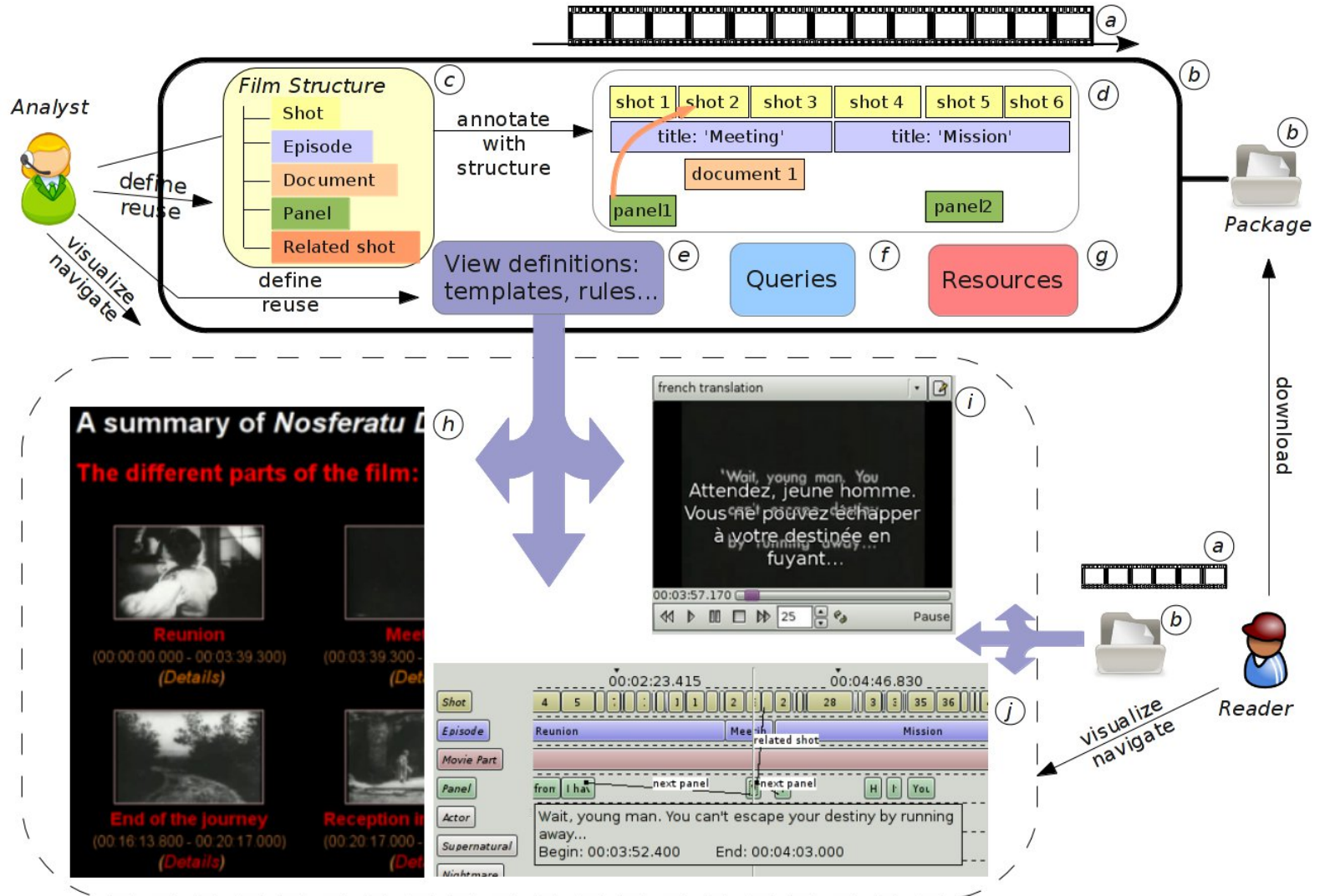
Active reading



Advene project

- Annotate Digital Video, Exchange on the Net :
Create, use, exchange one's own analyses of audiovisual documents as hypervideos
- Since 2002 with Yannick Prié and P.A. Champin
- Free software (GPL) cross-platform desktop app.
www.advene.org

Advene principle



Advene interface

The screenshot displays the Advene web interface for a video titled "Tim Berners-Lee at TED: The next Web of open, linked data". The interface is divided into several panels:

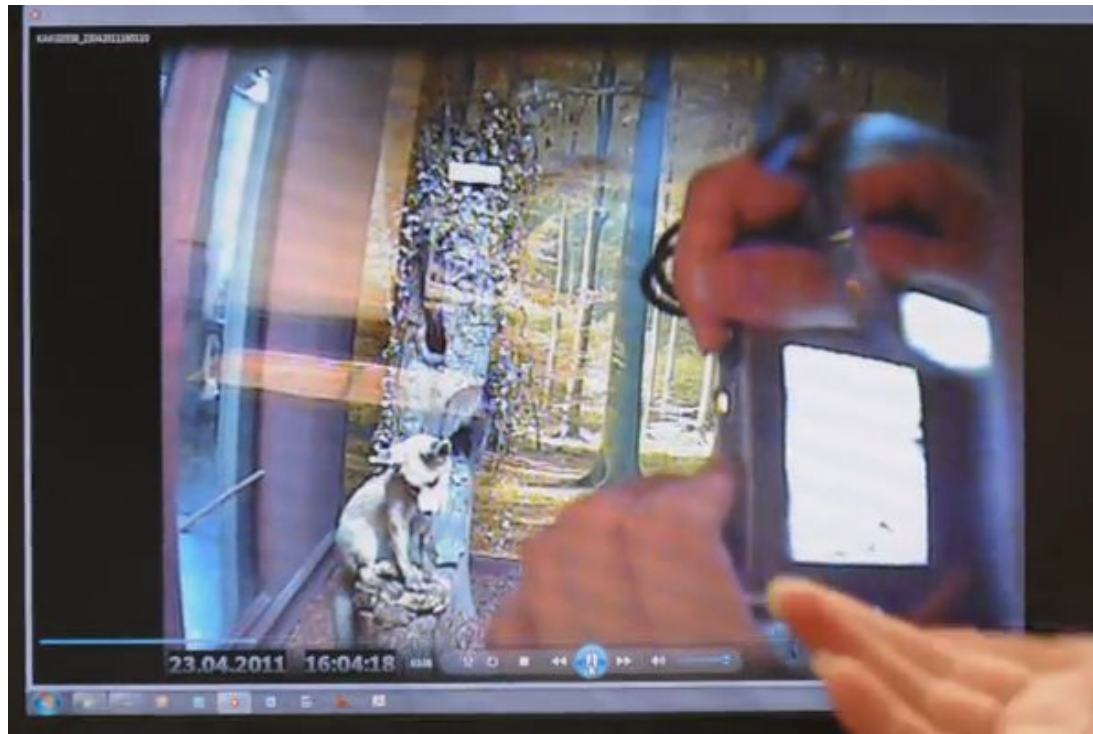
- Video Player:** Shows a video of Tim Berners-Lee on stage. The timeline at the bottom indicates the current position at 00:02:29.203.
- Transcript:** A scrollable text area showing the video's transcript. The visible text includes: "systems. So that, in all that diversity, if I wanted to figure out how to build something out of one little bit of this and a bit of this, everything I looked into, I had to connect to some new machine, I had to learn to run some new program. I had to - I would find the data may be, the information I wanted, in some new data format and they were all com- incompatible. It was just very frustrating, the frustration was on this- all this unlocked potential. In fact on all these disks, there were documents. So, if you just imagine they all being part of some big virtual documentation system in the sky, then- say, on the internet, then life would be so much easier. Well, once you have an idea like that, it kinds of gets under your skin, and even if people don't read your memo (actually he did, it was found after he died, his copy, it was found and he'd written 'vague but exciting' in pencil".
- Table of contents:** A list of links for different parts of the video:
 - Origin of the Web: frustration
 - A grassroots movement
 - The importance of data
 - The principles of Linked Data
 - It's working: DBpedia
 - Government data
 - Demand raw data now
 - Scientific data
 - Personal data
 - OpenStreetMap
 - What it's all about
- Timeline:** A detailed view of the video's structure, showing segments like "Text annotation", "Transcript", "Links", "Parts", "Summary", "Shots", and "Subtitles from TimBerners-Lee_2009.1.srt". The "Parts" segment is currently selected, showing "Origin of the Web: frustration" and "A grassroo".
- Inspector:** A panel showing the current annotation "a1231" with its duration (00:01:27.641) and a screenshot of the video frame.
- HTML Viewer:** A panel showing the HTML content of the selected annotation, which includes the text from the transcript and a small video thumbnail.

Different levels of flexibility

- Accompanying the exploratory analysis
- Variety of annotation creation interfaces
 - Manual annotation with possible assistants
 - Using automatic results as base
- Flexibility in annotation structure
 - Dynamic (re)structuring
- Flexibility in annotation visualisation

Actual use: museographie

Objective: study visitor's perception of museum exhibits through analysis of visitor video interviews



Technical setup



Analysis process

- Transcribe interview (note-taking view)
- Identify hexadic signs (from course-of-action methodology)
- Use relations to express courses of experience
- Generate visualizations through templates
- Write analyzes
- Publish on the web: www.museographie.fr

Advene interface

The screenshot displays the Advene interface for a video titled "Advene - MZS_yuki.azp - Cours d'expérience, Yuki, Musée zoologique, Strasbourg, avril 2011". The interface includes a menu bar (File, Edit, View, Player, Packages, Help), a toolbar with various icons, and a main content area. The main content area is divided into three sections: a video player on the left, a transcript on the right, and a timeline at the bottom.

The video player shows a scene of birds in a museum display case. The transcript on the right contains the following text:

V: mh I think it's a little bit bored it's not + it's just really a kind of collection for me + it's like there are collections + we have this kind of collections of birds this: huge amount + it's really huge amount <C: mh> but we didn't really show out like + were they live <C: mh> and where like + where they come from + what kind of environment do they live <C: mh> and also for this kind of the little notes not really clear because like + at least when you write where do get this bird this-this exact one what's the species name in latin and maybe in english or in germany I don't know but generally in french it's not really that good for-to people who doesn't really-doesn't know anything about knowledge to understand what they are <C: mh> they are just collections like-for a professor animal behaviour or animal or zoo knowledge + they will understand but for no more people it's a little bit bored

C: so you just pass
V: I just pass yea I couldn't even-I couldn't remember all the birds I can't
C: so sometime you look at <V: mh> we look at one of y...

The timeline at the bottom shows a scale from 00:07:56.080 to 00:13:53.140. It features a video thumbnail strip and several annotation tracks:

- Verbalisations**: V: I know it | V: mh I think it's a | V: CV: ok so | V: oh I say- V: I t | V: eve | V: n | CV: I'm | V: ye | V: | V: sl | C: V: I t | V: I d | V:
- Représentamen R**: La gale | La gale | Les oiseaux et es | Les mâles | Les
- Engagement E**: Eviter | Essaye | Essayer de trou | Essayer de trouver des | Env
- Anticipation A**: Attent | Attentes liées à d | Attentes liées à la prése | Atte
- Référentiel S**: Ce mus | Signe hexadique n°3 | Chez certa | Nou
- Unité de cours d'expérien**: Essayer de comprendre les intentions du musée sans y arriver | Cor

The timeline also shows a video player thumbnail strip and a video player control bar with a play button, a progress bar, and a volume icon. The video player control bar shows the video is playing at 42% scale with discrete scrolling.

Timeline

Synchronized transcription

Annotation types

Annotations + relations

Published hypervideo

Cours d'expérience, Yuki, Musée zoologique, Strasbourg, avril 2011 - Cours d'expérience - Google Chrome

Cours d'expérience, Yuki, | x

www.museographie.fr/exemple/MZS_Yuki/MZS_yuki_rss.mp4

Cours d'expérience, Yuki, Musée zoologique, Strasbourg, avril 2011

| [Signe hexadique n°1](#) | [Signe hexadique n°2](#) | [Signe hexadique n°3](#) | [Signe hexadique n°4](#) |

V: the crocodile I-I- that's what I just said + I said that-that big huge crocodile <C: yea> I saw it from the river Nile because the crocodile in the river Nile is one of the biggest crocodile in the world that's what I just say in chinese in this video <C: mh> yea +

C: so you know-you know it and euh

V: I know it yea I do + this kind of stuff I do-I have lots of memory euh I saw it's euh we didn't go inside because that's the museum for the insects <C: and> and were really both afraid of insects so we passed it <C: mh mh> (rires) yea + here or I don't have they are all birds + we saw + I tried all for birds + we just go inside because read an article + introduction and saw biggest collection of birds <C: yea> trying to see different kinds of birds

C: so how-do-did you find this-those birds+

V: mh I think it's a little bit bored collection for me + it's like-there are collections of birds. This huge museum

Signe hexadique n°3

Représentamen R

- La galerie des oiseaux
- Les oiseaux et es cartels associés aux oiseaux
- Les mâles et les femelles

Engagement E

- Essayer de voir différentes espèces d'oiseaux
- Essayer de trouver des informations
- Essayer de trouver des relations / des justifications sans en être très sûre



Visu project

- With Yannick Prié, Nicolas Guichon
- Web-based videoconferencing tool for language teaching
 - Session planning and management
 - Trace-based user activity monitoring
 - Document production from traces
- Free software (LGPL) until end 2011
github.com/ithaca/visu

Traces in Visu

- 3 types of activity traces
 - Interaction trace (automatically recorded *during* the interaction)
 - Marker trace (set by users *during* the interaction)
 - Comment trace (set by users *after* the interaction in the retrospection room)

Interaction room

The screenshot displays a web-based interface for an online interaction room. At the top, navigation tabs include 'Accueil', 'Utilisateurs', 'Séances', 'Salon synchrone', and 'Bilans'. The current session is titled 'Séance : Les loisirs' with a duration of 12 minutes. The main video area shows three participants: Vincent, A. Bein, and V. Caroline. A 'Votre marqueur' (Your marker) window is open over the video, showing a dropdown menu with 'A. Bein' selected and the text 'confond foot'. Below the video, a 'Résumé de l'activité' (Activity summary) timeline shows a blue bar representing the session duration from 00:00:00 to 00:03:44. The chat window on the right lists messages from participants, including 'ping pong', 'tennis', and 'rugby'. The bottom left corner features a list of sports: football, tennis, danse, gymnastique, ping pong, volley-ball, and handball. The interface is in French, with the language set to 'Français'.

Retrospection room

The screenshot displays a video conference interface for a session titled "Les loisirs", recorded on 20-10-2011 at 15:14. The interface includes a top navigation bar with tabs for "VISU", "Accueil", "Utilisateurs", "Séances", "Salon synchrone", and "Bilans". On the right, it shows the language set to "Français", a "Chat" button, the user name "S. Serguei", and a "Déconnexion" button.

The main area features three video feeds: V. Caroline (left), S. Serguei (top right), and A. Belin (bottom right). A timeline at the bottom indicates the current time is 00:05:17. A red box highlights this time on the timeline, with a red line extending down to a comment box containing the text "vous voyer / voier".

On the left side, there is a "Comments" section and a sidebar for "S. Serguei" with options for "Marqueurs", "Messages", and "Documents". The "Marqueurs" section includes "Marker sur la séance", "V. Caroline", and "Messages". The "Documents" section includes "Images" and "Videos".

On the right side, there is a "Liste des bilans" section.

Trace-based review editing

The screenshot displays a software interface for reviewing and editing video traces. The main window is titled "Les loisirs", enregistrée le 7-11-2011 à 11:10. It features a video player with three participants: P. Yannick, V. Caroline, and S. Serguel. Below the video is a timeline with a play button and a progress indicator at 00:03:05. The editing panel on the right, titled "Edition de bilan", shows a list of corrections for "Bilan Correction Serguel". The corrections include:

- T Bloc titre
- I Bloc texte
- Bloc vidéo ...
- Bloc comm...

The editing panel also displays the duration of the video (00:17) and a comment: "Tu as dit le 'tennis à table' mais on dit le 'tennis de table'".

At the bottom, there are three panels for comments and actions:

- Comments: A pink panel with a green plus icon.
- P. Yannick: A yellow panel with icons for Markers, Messages, and Documents.
- S. Serguel: A blue panel with icons for Markers, Messages, and Documents.

Review visualisation

VISU Accueil Utilisateurs Séances Bilans Langue : Français Chat V. Caroline Déconnexion

Explorateur

Filtrer par séance

Date prévue	Titre de séance	Propriétaire de séar
Tue Jan 25 18:0	Introduction, faisons c	D. Marie
Tue Feb 1 18:10	L'identité nationale: vt	S. Morgane
Tue Feb 22 18:1	L'art: architecture, peir	D. Marie
Thu Oct 20 15:0	Les loisirs	V. Caroline
Mon Nov 7 11:10	Les loisirs	V. Caroline

Tous les bilans
 Mes bilans (créés par moi)
 Autres bilans (créés par d'autres) et partagés avec moi

Filter:

Bilan Correction Yannick - V. Caroline -

À propos de la séance: Les loisirs
De V. Caroline, le 07-11-2011 11:10

Bilan Correction Serguei - V. Caroline -


À propos de la séance: Les loisirs
De V. Caroline, le 07-11-2011 11:10

Bilan de "Bilan Correction Serguei"

Bilan Correction Serguei

Orthographe
dans les marqueurs, tu as mal orthographié deux mots:
- "sky", s'écrit correctement: "ski".
- "paraplante" s'écrit correctement "parapente".

Prononciation
Durée 00:17
Tu as dit le "tennis à table" mais on dit le "tennis de table".
Durée 00:04



Pause Couper le son

Conclusion

- Annotations are not only for searching
 - Navigating/structuring
 - Hypervideo building based on annotations
- Automation cannot cover every need
 - Need for appropriate (possibly assisted) manual annotation tools...
 - ...that will be used by motivated users

Conclusion 2/2

- Lessons learned
 - Video annotation is a complex task
 - User motivation
 - Genericity vs. simplicity
 - Automation vs. specificity
 - Hypervideo building still requires assistance from experts
 - but we can gather standard practices and templates

Thanks

Thanks for your attention

www.advene.org
www.museographie.fr
www.visu-tutorat.org