

Canonical Processes in Active Reading and Hypervideo Production

Olivier Aubert

Pierre-Antoine Champin

Yannick Prié

Bertrand Richard

Université de Lyon, Lyon, F-69003, France ; Université Lyon 1, CNRS UMR5205
LIRIS, Villeurbanne, F-69622, France

first-name.lastname@liris.cnrs.fr

ABSTRACT

Active reading of audiovisual documents is an iterative activity, dedicated to the analysis of the audiovisual source through its enrichment with structured metadata and the definition of appropriate visualisation means for this metadata, producing new multimedia objects called *hypervideos*. We will describe in this article the general decomposition of active reading and how it is put into practice in the Advene framework, analysing how its activities fit into the Canonical Processes model.

Categories and Subject Descriptors

H.5.1 [Multimedia Information Systems]: Video; H.5.4 [Hypertext/Hypermedia]: Architectures, Navigation

Keywords

Annotation, Advene, Hypervideo, Document Template, Audiovisual Information Visualisation, Sharing, Time and synchronisation

General Terms

Design, Experimentation

1. INTRODUCTION

In this article, we present how the Advene¹ prototype for active reading and hypervideo creation fits into the Canonical Processes for media production as proposed by [3]. The Advene framework allows many uses that can be considered as “classical” multimedia documents production (e.g. montage) for which the processes do not really differ from those

¹Annotate Digital Video, Exchange on the NET <http://liris.cnrs.fr/advene/>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Copyright 2007 ACM X-XXXXX-XX-X/XX/XX ...\$5.00.

of other tools. We will rather focus here on non-classical processes related to audiovisual active reading, a continuous activity that takes place within the interaction between a document and a reader. The product of this interaction continuously enriches the document and induces further interaction.

The paper is organised as follows. The first part introduces the general active reading activity and the second part briefly presents the Advene model, and details the processes involved in the active reading activity in Advene. The next part is devoted to comparing Advene audiovisual active reading processes with the proposed Canonical Media Processes. Eventually, we describe some of the interoperability opportunities identified thanks to the the Canonical Media Processes.

2. ACTIVE READING

Active reading is a process where a reader assimilates and re-uses the object of his reading, as part of his knowledge work [7]. Among its goals are the discovery of a document, its enrichment or its analysis. Active reading processes generally produce *annotations* that can be re-used together with the document for searching, navigating, repurposing, etc.

In the general case, active reading involves a subject (the reader), an object (the document), and a tool (like a pen or dedicated software). It can be thought of as the combination of four major types of actions, each one being a sequence of operations possibly done through the tool.

The first action is the *inscription of marks* (annotation) on the document. These marks may contain some information, or just be placed to remember that there is something important at a specific point in the document. This first action is related to the intention of leaving a message for a future reader, who may be the same person or another one. It can rely on previous experience of active reading and pre-defined categories of marks.

The second action is the *(re-)organisation* of the marks. One way to do so is to further specify and categorise them. The marks are then structured within analysis schemas (pre-existing or specifically defined). For instance, a movie critic will often use a shot/sequence decomposition of movies, accompanied with character appearances. If he wants to focus on the narrative analyses, he will introduce annotations to document the storyline, the discourse, etc. If he wants to

focus on the cinematographic aspects of the movie, he will introduce annotations to document the camera movements, the different shot transitions, etc. He will then define several schemas that represent several points of view for the analysis of the audiovisual document. Another type of organisation is to organise the marks themselves into new documents, which, for instance, may aggregate the marks plus some personal text to serve as a result of the active reading activity.

The third action consists in *browsing the enriched document*. The active reader can experience an enriched reading of the document that builds upon the marks, allowing him to locate or search for interesting information.

The last action is the *publishing process*. When the active reader decides that his work on the document is finished, or in a stable state, he can decide to publish it. It is a snapshot of the work in progress, in order to share and/or show it (new documents and/or marks).

A characteristic of active reading is that these actions do not necessarily take place in a linear sequence, and that inscription of the marks, organisation and browsing are actions that are tightly connected during a session, while publishing often indicates the end of an active reading session.

3. ADVENE

The Advene project aims at empowering users with the possibility to 1/ annotate videos; 2/ visualise the annotations in user-specifiable ways; 3/ share analyses and comments on videos in the form of hypervideos. It specifies both a data model and a framework, which can be used to experiment with innovative ways of dealing with annotated audiovisual documents. General principles include sharing annotations and analyses as packages separate from video material, and aiming at simplicity of use for widespread usage.

3.1 The Advene model

Inspired by the hypervideo model [2], the Advene model aims at simplicity and genericity. It gives a central role to the notion of *annotation*, which associates a piece of data of any type (specified by a MIME type) to a fragment of an audiovisual document. *Relations* express relationships between annotations, and can also hold a content of any type.

A simple structuring of annotations and relations is offered through the notion of *annotation types* and *relation types*, which define constraints. Annotation types define the MIME-type of their annotations, while relation-types define the possible types for related annotations. For a given task in Advene, several specific types of annotations and relations may be used, so coherent sets of annotation-types and relation-types are grouped into *schemas* in order to facilitate their re-use. For instance, since shot/sequences are a common way to structure a movie, the *Shot* and *Sequence* annotation-types can be defined and grouped into a *Structure* schema (see figure 1). Schemas are not formal ontologies [5], but offer a minimal structure which can be further constrained according to the needs.

Queries offer a way to identify the elements (annotations, relations...) matching a certain condition. They are used both for building visualisations and navigating in the audiovisual document through the selection of relevant annotations.

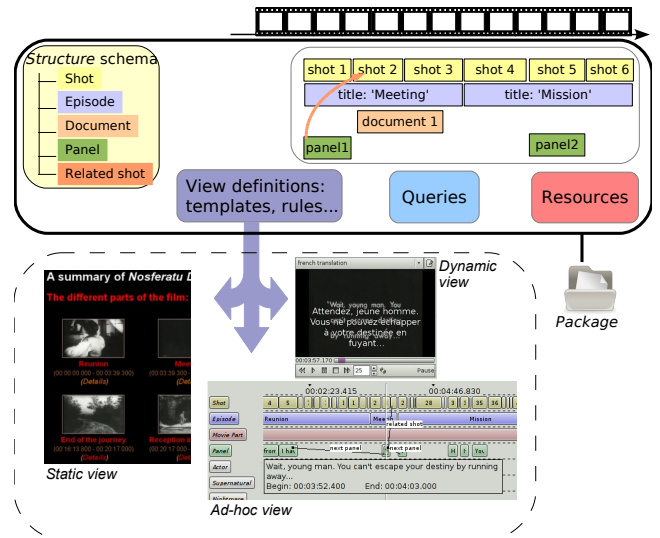


Figure 1: The Advene hypervideo model and principle

Finally, *views* allow to specify the various ways in which the annotation structure may be presented to the user. This variety of visualisations is a key feature of Advene: there are some canonical views (such as a timeline when dealing with audiovisual documents), but the variety of practices implies a variety of representations. In the current prototype, three types of views are available:

- ad-hoc (system-defined) views are specialised, GUI views such as a timeline or a dynamic transcription. They can be customized by the user through various parameters;
- static views allow to generate, through user-defined templates, X(HT)ML documents presenting elements from the annotation structure and the audiovisual document (screenshots);
- dynamic views allow to dynamically enrich the played audiovisual document by displaying information (taken from the annotation structure), navigating automatically or optionally in the document, etc. They are user-specified by a set of rules which define actions to be taken on the occurrence of certain events.

A hypervideo is then made of a number of linked views (ad-hoc, static and dynamic), more precisely of their renditions with the corresponding audiovisual documents.

As presented in figure 1, all elements of the Advene model (annotations, relations, types, schemas, queries, view definitions) can be serialized in a single documentary unit called a package, which can then be stored and exchanged, independently from the audiovisual document it is associated with. Exchanging analyses about an audiovisual document then consists in exchanging the Advene package, and applying it to a copy of the audiovisual document (for instance, two persons having their own copy of the same DVD).

3.2 The Advene framework and application

The Advene framework/application is a free (GPL license) application dedicated to the study of active-reading of au-

diophysical documents, in a collaboration perspective². It allows to quickly design and experiment new ideas in the domain of audiovisual metadata visualisation and interaction, both through the definition of views thanks to template languages, or through rapid implementation in the python language. It thus aims at providing tools for user-lead innovation [6] in multimedia visualisation.

The Advene prototype is cross-platform, and relies on standard software components, most notably the versatile VLC video player, enabling the use of almost any video format as source. The resulting views are rendered either in the application itself (for enriched audiovisual contents), or in a standard web browser through an embedded web server.

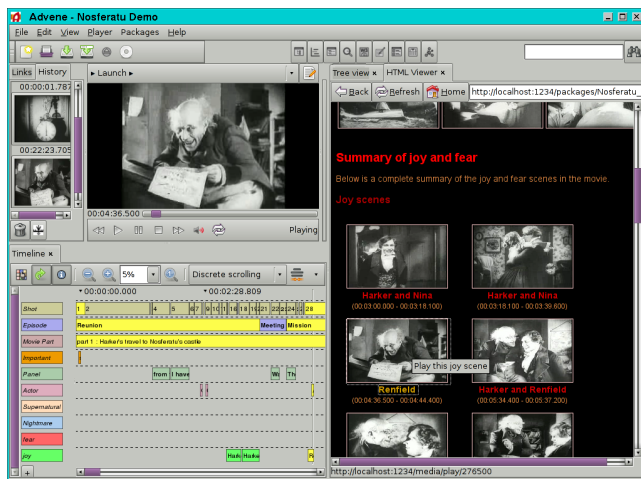


Figure 2: An example Advene session. An HTML view, generated from annotations, at the right hand side gives access to the corresponding movie parts. The timeline offers another view of the annotations. Some of them were automatically extracted (Shots), others are the result of an active reading process.

4. ADVENE PROCESSES

Let us describe how the four above-identified steps of active reading are carried out in the Advene framework. In order to illustrate the processes, we base our description on a scenario in which a teacher intends to give a course on Murnau's movie *Nosferatu*. A real Advene package, resulting from such an analysis of *Nosferatu*, can be downloaded from the Advene website, along with the Advene prototype. The narration of the teacher activity will be displayed in *italic characters*. Figure 2 presents the different processes involved during the four steps.

Mr Jones, teacher in humanities, wants to give a course about the expression of mood in movies. He bases his work on Murnau's movie Nosferatu, and more specifically on how the movie's nightmarish mood is built. He starts from scratch, having seen the movie only once and read some articles about it. We will see in the following parts how he annotates the movie in order to navigate in it and build his own argumentation, and how he defines representations for his annotations. In the end, after having completed his analysis,

²The prototype and examples are freely downloadable from the web site <http://liris.cnrs.fr/advene/>

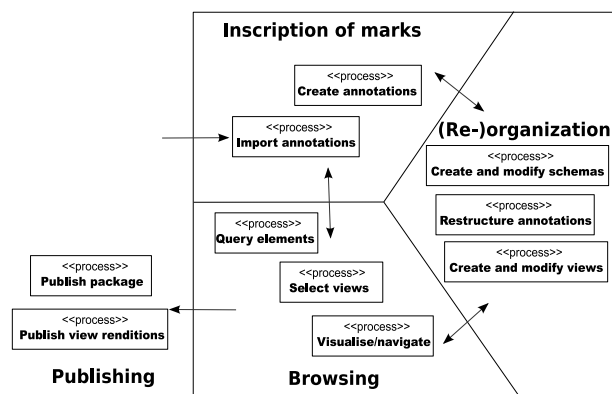


Figure 3: A session of active reading is dedicated to the study and the enrichment of one AV document in the Advene prototype. The session is composed of various actions related to the four active reading phases.

he produces both a text commenting some of the scenes, and an Advene package that he is able to distribute to his students, so that they can experiment with various interpretation possibilities by themselves.

4.1 First phase of analysis

In the first step of analysis, the teacher plays the movie and **annotates** it, both to mark sequences of interest in the movie, and to keep a track of the ideas that come to his mind.

Inscription of marks (annotation). Using the note-taking editor of Advene, Mr Jones types timestamped notes in a textual form, that he will later convert to annotations (Create annotations). He also uses an external application that generates a shot segmentation, and imports the resulting data in Advene, generating annotations corresponding to the shots (Import annotations).

The annotation process can be either manual (note-taking), automatic (shot-segmentation, speech recognition, import from other applications, etc.), or hybrid (assistance is offered to create annotations). In the end, annotations are pieces of data anchored to an audiovisual stream. Their origin can be stored as a meta-data but is not really relevant.

(Re)-Organisation/Browsing. Now that the teacher has created a first set of annotations, and thought of some ideas while watching the movie, he **organises** them in order to facilitate the navigation in the movie and present the information that he considers as relevant.

From the shot annotations and his own notes, Mr Jones identifies the shots containing nightmarish elements (Visualise/Navigate), creates a new Nightmare annotation type (Create structure), copies the annotations into it and describes them further (Create/Restructure annotations). By creating annotations of type Episode containing a short textual description of each part, he moreover defines his own table of contents of the movie (Create views), to ease the navigation in the movie. This organisation activity involves the creation of new annotation types, of views presenting

them as well as the creation of new annotations based on the previous ones.

From his new set of annotations, the teacher wants to dig into some ideas about the occurrence of specific characters or animals in the movie, by identifying specific annotations from the annotation structure. He can select them manually by identifying them in a view (*Visualise/Navigate*), or use a query (*Query*). Queries in Advene select elements from a package based on specified constraints. For instance, given a movie annotated with the appearances of characters, a query can return the appearances of a specific character in the movie. Queries can also be used cross-package, in order to gather information from a number of different packages.

Using queries, the teacher identifies a number of shots featuring animals (spiders, hyenas...) that contribute to the dark mood of the movie. While some queries can be generic, they are also meant to be user-defined, in order to match the specific needs of the user for a given task.

Document generation. Once elements are identified (manually or by a query), they can be used to create new documents, through Advene *views*. A view mixes selected elements from the annotation structure with the media assets or fragments thereof.

In order to ease navigation in the movie, the teacher defines a table of contents static-view (Create view), generated from the annotations of type Episode, illustrated by screenshots extracted from the movie, with hyperlinks allowing to play the corresponding part of the movie.

He also creates a dynamic view that displays the title of the episode as a caption over the video, in order to always know which part of the movie he is in when he navigates through the annotations.

Both types of view can be hyperlinked, creating a hypermedia document.

4.2 Second phase of analysis

Using the views he has just created, the teacher quickly navigates in the movie (Select views/Visualise/Navigate), and identifies shots which are relevant for his analysis of the nightmarish mood of the movie. While browsing the movie, he creates new annotation types (*Create annotations*) and new types (*Create and modify schemas*), identifying the occurrences of supernatural events. In the active-reading analysis, we find here a quick succession of browsing-inscription-organisation activities, when the user decides to enrich the annotations while watching the movie. Inscription occurrences may last a couple of seconds, and should generally not be obtrusive to the browsing process.

In order to have an overview of the nightmarish elements, he decides to create a dynamic view that generates a dynamic montage, showing only the shots related to nightmarish elements (Create view). This allows him to more precisely feel and analyse their relevance and relationships. This organisation activity brings the user to create more specific views, dedicated to the message/analysis that he wishes to carry.

Watching his new montage (Select view/Visualise/Navigate) corroborates the ideas that he wishes to convey, and allows Mr Jones to have a clearer view of how he will present them to his students. After the inscription/organisation phase, the browsing activity allows to visualise the resulting data, and evaluate its relevance.

4.3 Publishing

Now that Mr Jones has identified the relevant items and precised his analysis, he can write it down as a critic, in a static view illustrated by screenshots of the movie linked to the corresponding video fragments (Create view). He prints the rendition of this static view from a standard web browser, in order to distribute it to his students in class (*Publish view renditions*). In this last phase of analysis, the publishing activity is generally preceded by an organisation activity, in order to build the visualisation, and a browsing activity, used to validate the resulting view. Once the view has been defined and validated, it can be published.

He also cleans up the package containing his annotations and views, in order to remove his intermediate remarks and notes. He uploads the package to his homepage on a web-server (*Publish package*), so that his students can download it and use it to navigate in the movie, build new ideas and analyses from the same raw data, through the same active reading cycle. Publication and distribution of hypervideos is mainly intended to be in the form of Advene packages, thus needing specific software to render it. Indeed, one of the principles of Advene is to preserve the structured information (annotations) as long as possible in the process, thus at best from its production to its visualisation by the end-user. An Advene package, being a single file independent from the audiovisual material, can be distributed through any standard means (transmitted by e-mail, stored on a webserver) while preserving structured information, as in datuments [4]. It is also possible to distribute the result of the application of views to the package. For static views, it commonly consists of a generated XHTML document, which can be saved and distributed independently from its source package. Dynamic views may also be rendered in SMIL for instance.

5. CORRESPONDANCE WITH CANONICAL MEDIA PROCESSES

In order to avoid ambiguities, since our vocabularies are close but may express different notions, we will prefix the various Canonical Processes with the *CP*-prefix: *CP-annotate*, *CP-package*, *CP-query*...

As presented in figure 2, the four categories of activity are not used to transcribe a linear process, but are rather combined in an iterative process, which may go from one activity to another. Moreover, each activity may have variable durations, from a couple of seconds – for instance, the time used to take a note (*Create annotations*) or to take a look at a table of contents (*Visualise/Navigate*) – to several minutes or hours – for instance the time used to make a complete description of shots (*Create annotations*) or to watch the whole movie (*Visualise/Navigate*).

The *CP-create media assets* process is not directly present in the active reading activity. Active reading aims at exploiting existing media assets, issued from other applications, possibly also reusing existing annotations that may be issued from a *CP-annotate* process in the media asset production application(s). However, some of the results (views) of the active reading activity might be considered as new media assets.

Given the iterative nature of the active reading cycle, the *CP-premeditate* process is also continuously involved, be it in the creation of annotations, in their restructuring, in the definition of adequate views, or even in the view selection

and visualisation.

Since users always manipulate structured data, the *CP-package* process is in fact continuously occurring in an Advene session. Its result can be materialised at any time by saving the metadata structure to an Advene *package*. Queries and views are also stored in the Advene package, so that a package with its accompanying media files constitutes a documentary unit, keeping all necessary data to render the hypervideo. An Advene package is then both a container for all elements (metadata, queries, views) and a document generator.

The *CP-annotate* process is clearly mapped to the *Create annotations*. The *CP-semantic annotate* process can be identified in two aspects: in the Advene data structure expression which provides a simple categorization of elements (annotation- and relation-types), and, in a more generic way [1], in the capability of using any type of content in the annotations, RDF snippets for instance, in order to integrate semantic data.

The *CP-query* process also directly maps to the *Query elements* process.

The (Re)-organisation phase of the active reading activity involves multiple Canonical Processes: the input to the *CP-construct message* process is made of the *CP-premeditate* outputs, which are themselves elaborated in the browsing phase (*Select views* and *Visualise/Navigate*). The resulting message is then fed into a *CP-organise* process and materialised as Advene view templates (*Create views*), which transcribe the intention of the author. The *Create and modify structure* and *Restructure annotation* Advene processes may be part of a *CP-package* process, organising the logical articulation of the annotations.

The Browsing phase, like the Publishing phase, consists in selecting appropriate Advene views and getting their rendition. It thus involves the *CP-publish* (for generating documents from the view templates) and *CP-distribute* (for rendering the generated documents).

5.1 Interoperability opportunities

Analysing the active reading activity through the Canonical Media Processes allows to identify many interoperability opportunities with other applications.

Advene can be used on any audiovisual document. It can thus be used to annotate media assets produced by other applications. The generic nature of Advene annotations' contents makes it possible to also integrate annotations (output from the *CP-annotate* process) from any application.

Thanks to its XML template system and embedded web server, Advene can provide to external applications XML representations for its annotations (output of *CP-annotate*), its query results (output of *CP-query*) or its view definitions (output of *CP-organise*).

Eventually, the aim of Advene is to provide flexible tools to explore audiovisual documents and build analyses. It thus helps to build the necessary input for the *CP-premeditate* and *CP-construct message* processes.

6. DISCUSSION

Advene being focused on metadata exchange and visualisation, we have identified two related issues: distribution and interaction.

The *CP-distribute* process is indeed not as canonical (in the sens of atomic) as it could be: it merges the distribu-

tion (transmitting the published document to the user) with the actual rendition (using a device or software to render the published document). Apart from the misunderstanding problems it may encourage, we think it lacks an opportunity to encourage interoperability at the exchange level. Indeed, in Advene, structured metadata is preserved as far as possible, which means that what is transmitted (between users wishing to share analyses about a movie for instance) is annotations, queries and visualisation specifications, and not necessarily published documents (view renditions).

The related issue deals with interaction: in active reading, a cycle of annotation/visualisation constantly modifies the annotations, their structure or their visualisation. An essential step in this cycle is the visualisation of metadata. It is for the moment mapped to the *CP-distribute* process, while involving no transmission of data. However, if we split *CP-distribute* into *CP-transmit* and *CP-render*, it can be more precisely mapped to the *CP-render* process.

7. CONCLUSION AND FUTURE WORK

In this article, we have presented the notion of active reading, and detailed the various activities that it involves. We have seen how it is put into practice for audiovisual documents in the Advene framework for hypervideo production. Through a running example, we have described the active reading activities, and how they relate to Canonical Media Processes.

The analysis through Canonical Media Processes has allowed us to identify interoperability opportunities with other applications. It is all the more important that one goal of Advene is the integration and visualisation of audiovisual metadata. The Canonical Media Process thus constitute a excellent opportunity to foster interactions between projects.

8. REFERENCES

- [1] O. Aubert, P.-A. Champin, and Y. Prié. Integration of semantic web technology in an annotation-based hypervideo system. In *SWAMM 2006, First International Workshop on Semantic Web Annotations for Multimedia, 15th World Wide Web Conference*, May 2006.
- [2] O. Aubert and Y. Prié. Advene: active reading through hypervideo. In *ACM Hypertext'05*, pages 235–244, Salzburg, Austria, Sep 2005.
- [3] L. Hardman, F. Nack, Z. Obrenovic, B. Kerherve, and K. Piersol. Canonical process of media production. *Multimedia Systems Journal, Special Issue*, 2007.
- [4] P. Murray-Rust and H. Rzepa. The next big thing: From hypermedia to datuments. *Journal of Digital Information*, 5(1), 2006.
- [5] R. Troncy. Integrating Structure and Semantics into Audio-visual Documents. In *Second International Semantic Web Conference (ISWC2003)*, pages 566–581, Sanibel Island, Florida, USA, 2003. Springer.
- [6] E. von Hippel. *Democratizing innovation*. MIT Press, Cambridge, Apr 2005.
- [7] R. Waller. Functionality in digital annotation: Imitating and supporting real-world annotation. *Ariadne online review*, Mar 2003.

Canonical processes	Active reading in Advene
Premeditate (1)	<i>Inscription of marks/organisation/browsing</i> Due to the iterative and continuous nature of thought building in active reading, the premeditate process cannot really be isolated. It rather takes place in every step of the activity. Input: thoughts of the author Output: necessary schemas, annotations, queries or views
Create (2)	N/A, the active reading activity does not produce media assets per se, but exploits existing audiovisual documents.
Package (3)	<i>Inscription of marks/organisation/browsing</i> In every step of active reading, the metadata structure and accompanying queries and views are present, and can be materialised into an Advene package.
Annotate (4)	<i>Inscription of marks</i> Creation of the annotations, with spatio-temporal links to the media assets. Input: Media sources Output: annotation structure
Query (5)	<i>Organisation</i> Advene queries allow to select appropriate annotations. Input: Advene elements. Output: Advene elements matching a specify query.
Construct message (6)	<i>Organisation</i> Structuration of the presentation of data. Input: the ideas from the premeditate process, the annotation structure, queries. Output: draft of views.
Organise (7)	<i>Organisation</i> Definition of views to render the selected annotations. Input: Advene elements. Output: View definitions.
Publish (8)	<i>Browsing, Publishing</i> <i>CP-publishing</i> , meaning here generation of documents from the templates, occurs in the Browsing phase and also in the Publishing phase. Input: Advene elements. Output: an Advene package and/or rendered views.
Distribute (9)	<i>Browsing, Publishing</i> We are considering that the <i>CP-Distribute</i> is made of two parts: the transmission of the necessary data, and the rendition of documents through appropriate software. Advene is not concerned per-se by transmission. We rather rely on standard means of distribution, thanks to the creation of Advene packages as files that can be stored and exchanged. The rendition of view is currently done through a standard web browser, or the instrumented video player integrated into the prototype.

Table 1: Relationships between active reading phases and Canonical Media Processes.