



HAL
open science

Le projet ASA-SHS (Atelier Sémiotique Audiovisuelle pour la constitution, description et (ré-)publication de corpus audiovisuels en ligne en Sciences Humaines et Sociales).

Peter Stockinger

► To cite this version:

Peter Stockinger. Le projet ASA-SHS (Atelier Sémiotique Audiovisuelle pour la constitution, description et (ré-)publication de corpus audiovisuels en ligne en Sciences Humaines et Sociales).. [Rapport Technique] ESCoM-AAR (Equipe Sémiotique Cognitive et Nouveaux Médias - Archives Audiovisuelles de la Recherche). 2008. hal-01562285

HAL Id: hal-01562285

<https://hal.science/hal-01562285>

Submitted on 13 Jul 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

	Programme BLANC	Réservé à l'organisme gestionnaire du programme N° de dossier : ANR-08-XXXX-00 Date de révision :
	Document scientifique associé	Edition 2008

Acronyme	ASA-SHS
Titre du projet	Atelier Sémiotique Audiovisuelle pour la constitution, description et (ré-)publication de corpus audiovisuels en ligne en Sciences Humaines et Sociales.
Proposal title <i>(en anglais)</i>	ASA-SHS – a semiotic workshop for the processing and republishing of online audiovisual corpora in social sciences and humanities.
Auteur et coordinateur scientifique du projet	Peter Stockinger , Professeur des universités Institut National des Langues et Civilisations Orientales (INALCO) Fondation Maison des Sciences de l'Homme (FMSH) – ESCoM-AAR
Début et durée du projet	Début : 15/03/2009 Durée : 36 mois
Financement ANR	354.000 euros (sur 36 mois)

*Les pages seront numérotées et l'acronyme du projet devra figurer sur toutes les pages du document en pied de page.
Un sommaire du document est bienvenu*

Table des matières

1.1	Problème posé/Rationale	3
1.2	Contexte et enjeux du projet/Background, objective, issues and hypothesis	4
1.3	Description des travaux : programme scientifique/ <i>For each specific aim: a proposed work plan should be described (including preliminary data, work packages and deliverables)</i>	11
1.4	Résultats escomptés et retombées attendues/ <i>Expected results and potential impact</i>	25
1.5	Organisation du projet/ <i>Project flow</i>	26
1.6	Organisation du partenariat/ <i>Consortium organisation</i>	29
1.6.1	Pertinence des partenaires/ <i>Consortium relevance</i>	29
1.6.2	Complémentarité et synergie des partenaires/ <i>Added value of the consortium</i>	31
1.6.3	Qualification du coordinateur du projet et des partenaires/ <i>Principal investigator and partners : résumé and CV</i>	31

1. Programme scientifique / Description du projet

Technical and scientific description of the proposal

1.1 Problème posé/Rationale

Les sciences humaines et sociales ont de plus en plus systématiquement recours à l'audiovisuel numérique pour la constitution de leurs patrimoines scientifiques (sous forme, par exemple, d'enregistrements audiovisuels de conférences, de séminaires et cours, de colloques ou encore d'études de terrain) afin de pouvoir les utiliser ultérieurement aussi bien pour la recherche elle-même que pour l'enseignement ou encore, d'une manière plus générale, pour la valorisation des connaissances scientifiques produites dans une discipline. Ainsi, en France comme ailleurs, on voit se multiplier des initiatives concrètes de réalisation d'archives, de médiathèques, d'espaces numériques de connaissances, de patrimoines scientifiques et culturels à l'aide de l'audiovisuel numérique.

Cependant, on peut constater également qu'à ces initiatives fait souvent défaut une réflexion théorique plus approfondie sur le contenu cognitif et sémiotique de ces ressources audiovisuelles, sur les méthodes ou stratégies de la production ou constitution d'un patrimoine audiovisuel et, enfin, surtout sur les différentes possibilités de son exploitation pratique. Ainsi, en France, comme ailleurs, production et diffusion (sur le web, par exemple) de corpus de ressources audiovisuelles scientifiques restent souvent enfermées dans un modèle de référence historique omniprésent qui nous vient « directement » des pratiques et traditions télévisuelles. Ce modèle de référence est cependant bien trop simple, voire simpliste, pour pouvoir satisfaire aux attentes et besoins multiples des « utilisateurs » réels ou potentiels des ressources audiovisuelles à caractère scientifique.

Par rapport à cet état de fait, l'**hypothèse centrale** du projet ASA-SHS est la suivante : seule une prise en compte sérieuse et systématique de la *structure même* du document audiovisuel, i.e. de l'*organisation interne de son contenu* et de sa « mise en scène » *audiovisuelle* ouvrira progressivement la voie à des pratiques réellement pertinentes de production, d'appropriation ou encore d'exploitation de ressources audiovisuelles réalisées en SHS par rapport à leur intérêt, à leur « utilité » pour de contextes d'usage aussi divers que sont la recherche elle-même, l'enseignement et l'éducation, l'apprentissage tout au long de la vie, l'édition multimédia ou encore le journalisme scientifique.

Le projet ASA-SHS repose sur une expérience concrète de sept ans de constitution d'un patrimoine audiovisuel numérique d'environ 5000 heures de vidéos documentant la recherche en sciences humaines et sociales sous forme d'entretiens avec des chercheurs dans les différentes disciplines, de séminaires, de colloques, mais aussi de reportages, de documentaires ou encore d'enregistrements « de terrain ». Ce patrimoine est accessible sur le site portail du Programme « Archives Audiovisuelles de la Recherche » (AAR)¹ de la Fondation Maison des Sciences de l'Homme, partenaire principal du projet ASA-SHS. Comme pratiquement toutes les autres initiatives du même type (en France, par exemple : Canal U, Diffusion du Savoir de l'ENS, Canal IRD, ...), aussi les AAR se trouvent aujourd'hui confrontées à une variété de problèmes bien concrets (difficultés de localiser des informations pertinentes, adapter des ressources audiovisuelles existantes à des contextes d'utilisation spécifiques, parcours d'exploration guidés par des principes pédagogiques, ...) qui renvoient tous à la *même* question centrale qui est celle de la nature et de l'organisation cognitives, rhétoriques et sémiotiques du contenu (i.e. des « informations ») d'un document ou un corpus de documents audiovisuels et de son expression visuelle et sonore.

Or cette problématique fait partie intégrante de toute une tradition de recherche en France (hélas trop souvent négligé), à savoir de la rhétorique, de la sémiotique (ou sémiologie) et de l'analyse du discours (filmique). En s'appuyant sur cette tradition et en utilisant le fonds audiovisuel des AAR comme contexte d'expérimentation concrète, le projet ASA-SHS se propose par conséquent de réaliser les quatre objectifs principaux suivants :

1. définir et mettre en place un **cadre explicite et opérationnel** pour la description et modélisation **cognitive, sémiotique et rhétorique** des ressources audiovisuelles véhiculant un patrimoine scientifique et culturel SHS ;
2. sur la base des analyses sémiotiques concrètes du discours audiovisuel en SHS, développer des **outils intellectuels** de « traitement » (de description, d'indexation, d'enrichissement, de publication, ...) de corpus de ressources audiovisuelles en SHS, à savoir :
 - 2.1) des **ontologies** « traduisant » les résultats de l'analyse sémiotique en un métalangage de description ;

¹ adresse du site portail des AAR: <http://www.archivesaudiovisuelles.fr/FR/>

- 2.2) une bibliothèque de modèles ou *schémas de description* de ressources audiovisuelles en SHS et
- 2.3) des *modèles de publication* adaptés à des contextes d'usages spécifiques (tels que la recherche, l'enseignement, l'édition en SHS, etc.) ;
3. la mise en place d'une *méthodologie de la chaîne du « traitement »* de ressources audiovisuelles tenant compte des intérêts et besoins des principales communautés d'utilisateurs des connaissances produites en SHS, i.e. des chercheurs eux-mêmes, des enseignants et étudiants et enfin des « professionnels » (éditeurs, journalistes, ...)
4. la réalisation d'un *pilote* utilisant trois corpus audiovisuels de travail en *communication interculturelle, héritages artistique et littéraire* et *archéologie* (et provenant en grande partie du fonds des AAR) pour exemplifier, tester et, enfin, diffuser les méthodologies et outils permettant l'exploitation personnalisée d'un fonds audiovisuel.

Ces quatre objectifs du projet se concrétisent sous le sigle d'*atelier sémiotique audiovisuel* qui se présentera comme une « boîte à outils » intellectuels intégré dans un environnement informatique déjà existant et utilisé dans le cadre du programme des AAR.

1.2 Contexte et enjeux du projet/Background, objective, issues and hypothesis

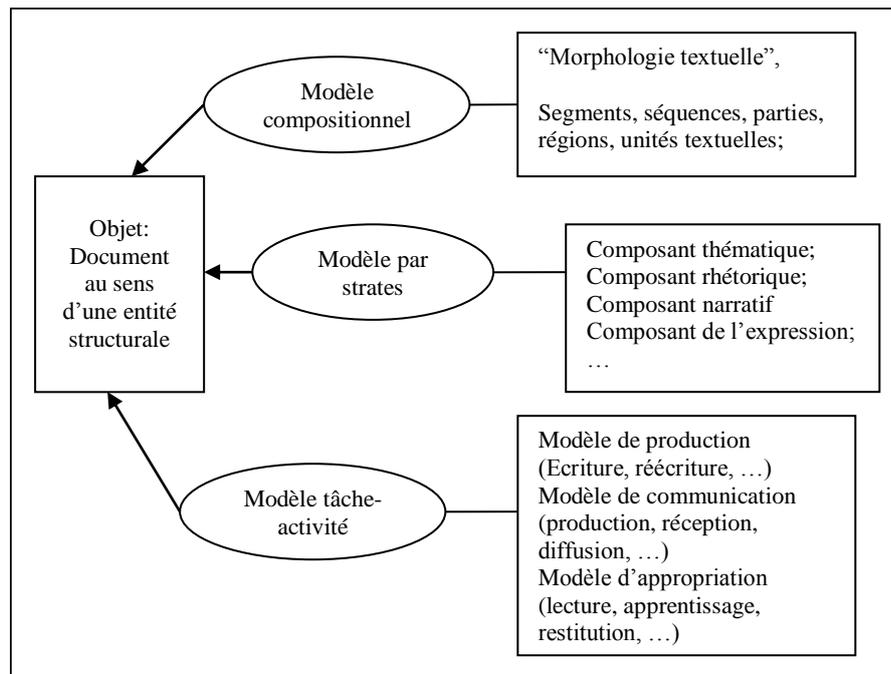
1.2.1 Le cadre théorique et la problématique générale

Le projet ASA-SHS vise donc la mise en place d'un cadre théorique et méthodologique commun à la constitution, description et plus particulièrement à la (re-)publication de corpus audiovisuels numériques en SHS. Un des enjeux principaux du cadre commun est de permettre une appréciation bien plus riche, plus adaptée aux attentes des communautés de lecteurs, du contenu intellectuel (scientifique, culturel, idéologique, ...) des ressources audiovisuelles en SHS telles que celles produites dans le cadre du programme des AAR mais aussi dans celui d'autres programmes et initiatives en France ou ailleurs. On ne peut ouvrir la voie à une telle appréciation que si la spécificité sémiotique et cognitive intrinsèque de ces ressources est prise en compte – ce qui est loin le cas, aujourd'hui, dans la très grande majorité des médiathèques ou autres bibliothèques en ligne.

Ladite nature cognitive et sémiotique des ressources audiovisuelles (comme des autres types de ressources : visuelles statiques, textuelles, ...) peut être explicitée d'une manière intuitive et simple à l'aide des sept questions typiques suivantes :

1. Quelle est la partie (séquence) de la ressource que je prends en considération ?
2. De quoi parle-t-elle (i.e. de quoi parle « ma » partie, « mon » segment) ?
3. Comment en parle-t-elle ?
4. Comment se développe-t-elle en un tout cohérent (i.e. en un segment ou une séquence) et comment s'intègre-t-elle dans un tout supérieur (i.e. en une macrostructure « finalisée ») ?
5. Comment « ma » partie, « ma » séquence s'exprime-elle (à travers des médias d'expression verbale, visuelle, sonores, ...) ?
6. En quoi ressemble-t-elle à d'autres ressources du même type (question du genre, de la tradition) ?
7. Que puis-je/dois-je faire pour conformer son contenu et/ou son expression à mes besoins ou objectifs d'usage ?

Ces sept questions possèdent à peu près le même statut que celles formant la fameuse topique journalistique (Qui ?; Quand ?; Où ?; Comment ?; Pourquoi ?), c'est-à-dire qu'elles servent à « fixer » et à orienter les idées et les habitudes aussi bien *en amont* de la production d'une information (i.e. antérieur de tout tournage) qu'*en aval* (i.e. pendant les phases de la publication à proprement dite d'une ressource : description, indexation, etc.).



(figure 1 : le cadre sémiotique général de la description d'un document ou d'un corpus de documents)

Les sept questions ci-dessus renvoient à l'échafaudage d'une sémiotique du document audiovisuel qui s'articule autour de trois modèles (cf. ci-après, figure 1) :

1. Le **modèle compositionnel** qui renvoie au fait qu'un document audiovisuel considéré comme un « tout cohérent » peut être décomposé en des unités (séquences, segments, parties, ...) qui, à leur tour se « comportent » comme des documents en formant des tous cohérents. Par ailleurs, il est à souligner qu'un document audiovisuel n'est pas obligatoirement limité physiquement par le début et la fin d'une « vidéo » - un document au sens d'un « signe porteur d'information » peut parfaitement se matérialiser en des « bouts » de vidéos physiquement séparés mais sémantiquement formant un tout - un document audiovisuel à support numérique doit donc être compris plutôt comme un réseau fonctionnel et hiérarchiquement intégré de « parties » qui, elles-mêmes, peuvent se présenter sous forme de documents se composant à leur tour de parties ou séquences. Pour le projet ASA-SHS, cela peut dire, par exemple, qu'un document audiovisuel au sens d'une ressource (d'informations, de connaissances) peut être un « film » en entier, i.e., dans notre cas, l'enregistrement d'un cours, ou d'un entretien ; il peut être aussi un segment particulier (i.e. un moment dans un séminaire ou cours consacré à une thématique, un sujet ; il peut être, enfin, une collection de segments appartenant à différents « films » (de cours, ou séminaires, ...) mais faisant un nouveau tout cohérent du point de vue thématique ou du point de vue d'un usage spécifique (exemple : un ensemble de segments consacrés à une langue particulière « extraits » de différents films ou enregistrements mais composant un nouveau film ré-linéarisé ou simplement un « dossier » de segments destiné à un usage pédagogique).
2. Le **modèle de strates** qui représente l'organisation interne d'un document (audiovisuel) en détaillant la dichotomie saussurienne signifié/signifiant. Ainsi, un document (audiovisuel) peut être décrit au niveau *thématique* ou *référentiel* (cf. la question « De quoi parle-t-il » ?) ; au niveau *rhétorique* et *discursif* (cf. la question « Comment en parle-t-il » ?) ; au niveau narratif (cf. la question « Comment se développe-t-il en un tout cohérent » ?) ; au niveau de l'*expression audiovisuelle* (cf. la question « Quels sont les médias qu'il utilise pour exprimer son contenu » ?) ; au niveau de l'organisation de l'expression (cf. la question « Comment les médias d'expression s'organisent-ils pour produire une mise en scène cohérente » ?). Etant donné la spécificité des ressources audiovisuelles dont le projet ASA-SHS dispose (les 5000 heures de vidéos composant le fonds des AAR), le projet travaillera surtout les plans thématique et référentiel ainsi que rhétorique et narratif, c'est-à-dire les plans qui composent le « contenu » stricto sensu (i.e. le « signifié » saussurien) d'une ressource. Mais grâce à un corpus de documentaires et de reportages dont disposent les AAR, les formes et modalités d'expression audiovisuelle formeront également l'objet d'analyse et de description.

3. Le **modèle d'activité** qui représente le fait que tout document est obligatoirement inséré en un réseau et contexte de pratiques et d'activités (aspect qui pour des raisons méthodologiques a été souvent mis entre parenthèses par les sémioticiens ...) et que ces pratiques et activités d'une part contribuent au « formatage », au « profilage » d'un document et d'autre part sont elles-mêmes tributaires de l'identité, du profil spécifique du document. Ainsi, un document audiovisuel se composant d'un ensemble d'extraits d'entretiens avec des chercheurs sur une problématique spécifique possède obligatoirement son identité, son profil spécifique qui (« à priori ») rend ledit document plus ou moins récalcitrant à des activités d'adaptation et de republication soit de l'ensemble du document soit de certaines de ses parties en vue de son exploitation dans des contextes d'usages éducatifs. Dans le cadre du projet ASA-SHS, les activités les plus importantes sont celles qui composent très typiquement la **chaîne du traitement du document audiovisuel**, à savoir :
- la production (la sélection, ...) et la préparation d'un corpus audiovisuel ;
 - la description, la segmentation, la « classification » thématique et l'indexation du corpus ;
 - l'adaptation (l'annotation, l'enrichissement, le « profilage », le « versionnement », ...) du corpus ;
 - la publication du corpus traité.

Ce cadre théorique général repose donc, comme déjà dit, sur une grande tradition de la recherche en sciences du langage mais aussi en analyse du texte littéraire. Sans vouloir remonter jusqu'à l'Antiquité, en France cette tradition est représentée par des personnalités emblématiques de la recherche sur le discours et le texte des années 60 et 70 telles que Roland Barthes, A.J. Greimas, Christian Metz, Michel Foucault ou encore Gérard Genette. En même temps, il a été très profondément renouvelé par des apports fort différents venant, par exemple, de la « nouvelle rhétorique » et de l'analyse critique du discours (notamment anglo-saxonne), des travaux sur une théorie des cultures basée sur une approche sémiotique (cf. notamment les travaux de l'anthropologue C. Geertz), des sciences cognitives (cf. notamment les travaux de D. Sperber sur une théorie de la culture, de G. Lakoff et L. Talmy sur les structures conceptuelles du langage, ...), des recherches en représentations des connaissances (cf. les travaux sur les réseaux sémantiques, les graphes conceptuels, les « topic maps », ...) mais aussi des recherches en linguistique du discours et pragmatique linguistique (cf. par exemple, la théorie systémique du langage de M.A.K. Halliday, la théorie des structures rhétoriques – R.S.T. – de B. Mann et S. Thompson, etc.).

La figure 1 représente donc le cadre théorique de référence pour le projet ASA-SHS pour la description du discours audiovisuel en SHS dont les 3 principaux enjeux, i.e. les 3 verrous à lever sont les suivants :

N°	Désignation de l'enjeu	Présentation succincte
1	<p align="center">La modélisation sémiotique du contenu scientifique des ressources audiovisuelles en SHS</p>	<ol style="list-style-type: none"> 1. comprendre et expliciter la structure thématique (les <i>topoi</i> ou encore, comme on dit en analyse littéraire, les <i>sujets</i>) qui sous-tend cette production intellectuelle et scientifique afin de pouvoir l'utiliser ensuite pour la construction d'une ontologie (d'un métalangage de description) du domaine concerné ; 2. comprendre et expliciter les principales formes rhétoriques et narratives utilisées pour thématiser une situation, un événement, un problème afin de pouvoir utiliser ces résultats dans le développement d'une ontologie rhétorique et narrative permettant de nuancer une description thématique par des critères rhétoriques forts importants pour une adaptation de ressources audiovisuelles à des usages spécifiques (il s'agit ici, par exemple, de pouvoir distinguer entre l'<i>exposé scientifique</i> ou <i>didactique</i> d'un problème ou événement, l'<i>exemplification</i> du même problème ou événement, l'<i>explication</i> du problème, la production d'<i>arguments pour</i> ou <i>contre</i>, etc.) ; 3. identifier et décrire les parties textuelles fonctionnelles dans les ressources audiovisuelles telles que les <i>introductions</i> et <i>conclusions</i>, les <i>renvois</i>, les <i>rappels</i>, les « <i>moments forts</i> », les « <i>digressions</i> » etc. afin de pouvoir réutiliser ces données dans certains contextes spécifiques d'enrichissements, d'adaptations, d'explicitations, voire d'hierarchisation de contenus scientifiques ; 4. décrire notamment dans les enregistrements audiovisuels de situations de terrain ainsi que dans les différents documentaires et reportages, l'expression audiovisuelle à proprement parler, i.e.

		<p>la « traduction » d'un contenu dans le medium audiovisuel (cadrages, plans, synchronisation de différents éléments d'expression), etc. afin de proposer une ontologie de l'expression audiovisuelle comme un outil en plus pour décrire et adapter le contenu d'un corpus audiovisuel à un usage, un contexte particulier ;</p> <p>5. expliciter, enfin, la structure typique d'un ensemble de genres de publications scientifiques et pédagogiques tels que <i>dossiers thématiques, syllabus, portefeuille de lecture, ...</i> ainsi que la structure typique des constructions ou <i>parcours narratifs</i> permettant une exploration guidée et finalisée d'un univers de connaissances (par exemple, des parcours de type « exploration », « enquête », « découverte », etc.).</p>
2	<p>(sur la base des résultats des analyses sémiotiques)</p> <p>Développement des outils intellectuels pour la chaîne du traitement de corpus audiovisuels en SHS</p>	<ol style="list-style-type: none"> 1. Conception et développement de cinq ontologies nécessaires pour la description et indexation de documents audiovisuels, i.e. une <i>ontologie textuelle</i>, une <i>ontologie du domaine ou référentielle</i>, une <i>ontologie rhétorique</i>, une <i>ontologie narrative</i> et une <i>ontologie de l'expression audiovisuelle</i>. Ces ontologies prendront comme point de départ une ontologie déjà existante et utilisée pour l'indexation de certains corpus audiovisuel en anthropologie et études culturelles, i.e. l'ontologie v6 créée à l'aide d'un éditeur d'ontologie Onto Editor et complètement intégrée dans l'environnement de description, indexation et publication utilisé à l' dans le cadre du programme AAR. Ceci dit, l'actuelle ontologie utilisée – v6 – sera profondément modifiée dans sa structure d'organisation même et acquerra un niveau d'abstraction suffisamment élevé et riche pour pouvoir être utilisée sous forme de « modules » spécifiques à un domaine de connaissances choisi. 2. Définition et développement d'une bibliothèque de modèles de description du contenu audiovisuel sous forme de schémas ou graphes conceptuels. De nouveau, cette bibliothèque s'appuiera et enrichira très considérablement une petite bibliothèque d'une soixante de modèles de description actuellement utilisés dans le cadre du déploiement des nouveaux services sur le site des AAR (nouveaux services de type vidéo-glossaire ou vidéo-encyclopédie). 3. Spécification et réalisation informatique de nouveaux modèles de publication de corpus audiovisuels sous forme d'une part de dossiers scientifiques et pédagogiques de tout genre et d'autre part de « vidéos virtuels », i.e. de parcours narratifs créant de chemins d'exploration entre des segments audiovisuels disjoints les uns les autres. 4. Intégration des ontologies ASA-SHS, de la bibliothèque des modèles de description du contenu audiovisuel et des modèles et formulaires de publication dans l'environnement numérique actuel de travail appelé <i>Interview</i> (et co-développé avec l'INA Recherche dans le cadre d'une convention entre INA et FMSH).
3	<p>Exploitations pratiques des recherches sémiotiques et démonstration de leur utilité en vue d'une appropriation de ressources audiovisuelles</p>	<ol style="list-style-type: none"> 1. Travail concret de description, indexation republication ... de corpus audiovisuels choisis (dans les trois domaines : histoire, littérature et archéologie) à l'aide des outils sémiotiques développés dans le cadre du projet ASA-SHS. 2. Tests et validations des travaux par différents groupes d'utilisateurs composés de chercheurs d'une part et d'enseignants, étudiants d'autre part. Un quatrième groupe d'utilisateurs sera composé de « professionnels » d'utilisation de connaissances en SHS : éditeur (à commencer par les éditions FMSH qui viennent de mettre en place un programme de production de DVD à partir du fonds audiovisuels des AAR),

	<p>en ligne qui est conforme aux attentes, besoins et désirs des différentes communautés travaillant en SHS et/ou à partir des connaissances produites en SHS</p>	<p>journalistes et médiateurs socioculturels (travaillant notamment dans les deux domaines de la communication sociale et de la communication au service de la médiation linguistique et interculturelle).</p> <p>3. Documentation technique, méthodologique et didactique de l'ensemble des procédés nécessaires pour une « bonne » réutilisation des résultats du projet en dehors de la FMSH.</p> <p>4. Généralisation de l'approche sémiotique expérimentée sur trois corpus de travail à l'ensemble du fonds des AAR représentant quelques 5000 heures de vidéos et, au-delà des AAR aux patrimoines d'un ensemble de partenaires de la FMSH. Il s'agit notamment d'enregistrements audiovisuels provenant de fouilles archéologiques, d'enregistrements sociolinguistiques (écoles de langue, dialogues et interactions, enregistrées, ...), de corpus audiovisuels de type « road movies » (documentations audiovisuelles de villes, de pays, ...) et, enfin, de corpus composés de documentaires et de reportages à caractère scientifique ou, plus généralement culturel.</p> <p>5.</p>
--	--	--

A chacun de ces trois enjeux est consacrée une tâche précise du projet (cf. le chapitre suivant consacré à la description du programme de travail).

1.2.2 L'expérience du programme « Archives Audiovisuelles de la Recherche » (AAR)

Le projet ASA-SHS s'appuie sur deux expériences concrètes étayant son hypothèse principale de travail. La première expérience consiste en le développement, depuis sept ans à la Fondation Maison des Sciences de l'Homme (FMSH), du programme « Archives Audiovisuelles de la recherche » (AAR) ayant comme objectif la constitution d'un patrimoine multimédia en ligne en sciences humaines et sociales ; la seconde consiste en la participation, depuis 2000, de la FMSH à plusieurs projets de R&D français et Européens dédiés plus précisément aux questions à la *chaîne du traitement* de corpus audiovisuels afin de les mieux adapter aux attentes variées de leurs utilisateurs (lecteurs, ...) et aux contextes d'usage spécifiques (tels que la recherche elle-même, l'enseignement, le journalisme scientifique, etc.).

Le programme AAR recouvre actuellement un fonds d'environ 5000 heures de vidéos documentant la recherche en sciences humaines et sociales sous forme d'entretiens avec des chercheurs dans les différentes disciplines, de séminaires, de colloques, mais aussi de reportages, de documentaires ou encore d'enregistrements « de terrain ». Ce patrimoine est publiquement accessible sur le site portail du programme AAR de la Fondation Maison des Sciences de l'Homme (FMSH). Un « événement » scientifique (séminaire, entretien, conférence, colloque, etc.) est publié sur le portail des AAR sous forme d'un site dynamique (généralisé automatiquement à partir des données collectées au préalable par les réalisateurs auprès des auteurs et des métadonnées issues d'un travail de segmentation des vidéos, de la classification thématique des segments vidéo ou encore de leur indexation et annotation). La consultation elle-même se fait sous forme d'un « feuilletage » de différents « chapitres » (i.e. segments audiovisuels) qui composent le *vidéo-livre interactif* de l'événement filmé dans le cadre des AAR.

Cependant, comme pratiquement toutes les autres initiatives du même type (en France, par exemple : Canal U, Diffusion du Savoir de l'ENS, Canal IRD, ...), aussi les AAR se trouvent confrontées aux trois grands problèmes suivants :

1. la qualité et la richesse de son contenu est en quelque sorte écrasée par la quantité, le volume d'heures offertes à la communauté intéressée (composée de chercheurs, enseignants, étudiants, mais aussi journalistes et professionnels travaillant dans des ONG, des associations ou des agences gouvernementales) ;
2. le contenu (i.e. le ou les propos) véhiculé par les vidéos des AAR possède son identité propre lié à l'habitus aussi bien *individuel* de l'auteur (du chercheur, de l'expert, ... interviewé, de l'enseignant-chercheur responsable d'un séminaire, etc.) que *collectif* des usages et traditions disciplinaires et institutionnels (comme on le sait bien, deux linguistes, pour prendre cet exemple, travaillant sur la grammaire d'une langue donnée peuvent en faire un objet de connaissance fort différent) ;

3. le contenu scientifique est presque entièrement *monolingue* (i.e. soit uniquement en français – ce qui est le cas majoritairement –, soit uniquement en une des 15 autres langues actuellement représentées dans les AAR).

Le premier problème concerne avant tout la question de la description, classification et indexation de corpus audiovisuels ou plus largement multimédias. Le deuxième problème concerne plus particulièrement l'explicitation du contenu d'un document en tenant compte à la fois de son identité d'origine spécifique et du « profil » culturel et cognitif de son destinataire (projeté ou réel). Le troisième problème est traditionnellement associé à celui de la traduction au sens habituel du terme, mais dans le contexte des AAR il s'agit plutôt de la mise en place de services simples d'aides à la compréhension linguistique du contenu (sous forme, par exemple, de traduction de titres thématiquement saillants, de mots clés dénommant la structure à la fois thématique et rhétorique d'un segment, de thèmes faisant partie de l'ontologie du domaine, etc.). Ces trois problèmes constituent des vrais enjeux aussi bien pour une meilleure diffusion d'un patrimoine scientifique sur un « marché » numérique intrinsèquement multilingue et multiculturel que pour une appropriation plus adaptée aux attentes et besoins des communautés d'utilisateurs concernés.

Par ailleurs, les analyses statistiques régulières de la fréquentation du site AAR, des enquêtes menées auprès du public des AAR via un questionnaire en ligne sur le site portail² et enfin les retours réguliers des utilisateurs (enseignants, chercheurs, étudiants, journalistes, ...) du fonds audiovisuel des AAR, montrent bien les limites manifestent d'une « simple » médiathèque se contentant d'un ensemble d'accès plus ou moins « standards » à son fonds : grandes difficultés de localisation et de sélection d'une information pertinente dans une large base de données audiovisuelles ; linéarité temporelle du flux audiovisuel empêchant des formes d'exploration plus souple du type « feuilletage » ; absence d'aide d'exploration et d'appropriation des fonds de ressources audiovisuelles en soi vraisemblablement très pertinentes ; absence de terminologies usuelles qui permettraient de mieux comprendre la structure d'un fonds et, par conséquent, de mieux l'explorer ; trop de difficultés non seulement techniques mais intellectuelles d'utiliser, voire de réutiliser de ressources audiovisuelles pour des activités spécifiques dans la recherche mais aussi dans l'enseignement, voire la communication de vulgarisation, etc.

1.2.3 L'expérience de projets de R&D français et Européens

Les trois problèmes identifiés ci-dessus ont en effet motivé la participation de la FMSH à une série de projets de R&D français et Européens visant plus particulièrement une meilleure prise en compte des pratiques de traitement de corpus audiovisuels – traitement au sens de description, d'indexation, d'enrichissement, d'adaptation linguistique ou encore de publication plus diversifiée. Contrairement au projet ASA-SHS, il s'agissait de projets à visée surtout technologiques avec comme objectif concret la mise en place d'outils et d'environnements informatiques suffisamment sophistiqués pour rendre compte de la grande diversité et richesse des pratiques liées au traitement de corpus audiovisuels. Ainsi, la FMSH a participé au projet de R&D français OPALES, financé dans le cadre du programme RIAM et coordonné par l'INA³, et participe actuellement à un deuxième projet RIAM (ANR-CNC) de nom SAPHIR⁴, également coordonné par l'INA et devant se terminer fin 2008.

Ces deux projets technologiques de R&D ont donc permis la mise en place d'une infrastructure technologique (d'un environnement de travail) fort complexe ouvrant la possibilité à des activités de description, d'indexation et de republication très riches de corpus audiovisuels – activités qui permettent de mieux prendre en considération les attentes et besoins des utilisateurs.

En même temps, la FMSH a participé à plusieurs projets de R&D Européens (dans le cadre des programmes Leonardo, eContent et IST du 6^{ème} PCRD) tel que, par exemple, le projet LOGOS⁵, coordonné par la télévision hongroise Antenna Hungarica et réunissant une quinzaine de partenaires de dix pays de la CE. Ce projet qui doit se terminer en janvier 2009, porte sur la création d'un environnement technologique de publication de ressources audiovisuelles et visuelles statiques pour l'enseignement (formel et informel) via toute forme de support numérique (Internet, téléphone mobile, ipod, i-TV, etc.).

² Cf. <http://www.archivesaudiovisuelles.fr/FR/questionnaire.asp>

³ cf. Antoine Isaac, Un système d'annotation multiforme et communautaire de documents AV : Opales. http://liris.cnrs.fr/~yprie/Projets/SDN04/Presentations/MDAV04_isaac.pdf

⁴ cf. <http://www.ina.fr/entreprise/activites/recherche-audiovisuelle/projets-recherche/saphir.html>

⁵ cf. <http://www.logosproject.com/>

Des expérimentations sur des corpus restreints ont été réalisées et ont donné des premiers résultats encourageants publiés sur un site portail spécialisé du programme des AAR – site portail qui s’intitule « Diversité Linguistique et Culturelle » (DLC)⁶. Ce site donne un accès à un corpus audiovisuel (d’environ 150 heures) *re-traité*, i.e. réécrit, ré-indexé et ré-publié à l’aide d’une ontologie du domaine et en tenant compte des usages possibles de ces ressources dans le cadre d’un enseignement supérieur (notamment niveau L2/L3 et M1/2). Le site DLC est divisé en quatre grandes rubriques :

- une rubrique (de base) appelée « *médiathèque DLC* » qui est, comme son nom l’indique, une bibliothèque de ressources audiovisuelles (et visuelle statique) en ligne ;
- une deuxième rubrique « *encyclopédie hypermédia* » qui est une véritable encyclopédie permettant l’exploration de l’ensemble du fonds audiovisuel selon une quarantaine de grands sujets ou topiques du domaine en question (et se trouvant très régulièrement sur tous les plans d’enseignement en sciences du langage et études des langues/civilisations) ;
- une troisième rubrique composée de *dossiers hypermédiatiques et pédagogiques* (sous forme de syllabus, de portefeuilles de lecture, ...) pour enseignants et étudiants ;
- une quatrième rubrique composé d’une *bibliothèque de dossiers bilingues*, i.e. de vidéos en langue originale (surtout en français et espagnol) adaptées à des utilisateurs possédant des connaissances passives du français mais probablement incapables de suivre un discours scientifique en français (exemple de paires de langues : français/russe ; français/anglais ; espagnol/français ; français/turc ; français/arabe ; ...).

1.2.4 Conséquences et enjeux

L’intérêt de ces projets et de pilotes tels que le pilote DLC est d’avoir montré la *faisabilité conceptuelle et technique* d’une approche très sophistiquée du traitement du document audiovisuel, i.e. de la segmentation, description, classification, indexation, adaptation, publication selon les besoins et désirs des différents utilisateurs et en prenant en considération les contraintes socioculturelles spécifiques des usages d’un document audiovisuel donné. Ce résultat n’a pas été très évident. En citant encore une dernière fois le site « Diversité Linguistique et Culturelle » (DLC) qui a bénéficié de l’input technologique et informatique des projets cités, celui-ci propose des prestations d’explorations et d’appropriations des contenus audiovisuels qui vont bien loin au-delà des prestations habituelles : une encyclopédie hypermédia en ligne complétée d’index thématiques pour explorer des notions et problématiques telles que la politique linguistique dans le monde, le rapport entre langue et littérature, les différentes formes sociales d’une langue naturelle, etc. ; des dossiers pédagogiques aussi bien pour l’enseignement formel que pour des apprentissages informels du genre « sensibilisation culturelle » ; des versions bilingues de certains documents audiovisuels, etc.

Or, ce type de prestations qui diversifie énormément les utilisations possibles **d’un même document audiovisuel** « source » à travers des pratiques de « retraitement » (de « réécriture » ou encore, en anglais de « *repurposing* ») ne se rencontre certainement pas (encore) sur les principaux sites diffusant des ressources audiovisuelles en SHS.

Par contre, certains sites qui ne sont pas des sites d’institutions de recherche ou d’enseignement commencent à inclure progressivement des prestations bien plus sophistiquées que la simple diffusion de vidéos en ligne. En France, ce sont, par exemple, le site de TV5 et lesite.tv⁷ de la chaîne France 5 et du SCEREN/CNDP qui proposent des utilisations pédagogiques de certaines de leurs ressources audiovisuelles – tout comme RFI qui montre, sur son site⁸, d’une manière fort convaincante comment utiliser des (ses) ressources sonores pour un enseignement du français. En Allemagne, un site très intéressant est celui de la Deutschen Welle⁹ qui propose, à partir d’une republication de corpus de leurs ressources audiovisuelles (des reportages, des nouvelles télé, etc.), des programmes d’apprentissage de langue (du français, de l’allemand, de l’anglais, etc.). Enfin, citons encore l’excellent site i-Mark de la FAO¹⁰ qui propose non seulement des prestations pédagogiques, sur la base d’un re-traitement de documents audiovisuels « source » (notamment dans le domaine de l’agriculture) mais qui, en plus, offre tout un kit à ceux ou celles qui souhaitent produire des prestations pédagogiques en ligne.

La FMSH considère, en effet, que les médiathèques et autres vidéothèques proposant des ressources audiovisuelles en SHS devraient s’inspirer de ce type de démarche pour les transformer en des véritables ressources et outils pour

⁶ cf ; <http://semioweb.msh-paris.fr/corpus/dlc/FR/>

⁷ cf. <http://www.lesite.tv/>

⁸ cf; http://www.rfi.fr/lfr/statiques/accueil_apprendre.asp

⁹ cf. <http://www.dw-world.de/dw/0.2142.2617.00.html>

¹⁰ cf. http://www.imarkgroup.org/index_fr.asp

les différentes communautés travaillant en SHS et/ou à partir des connaissances produites par les SHS. Cependant ce que les deux expériences citées ont montré sans ambiguïté, c'est non seulement la pertinence scientifique mais tout simplement la nécessité de prendre en compte la *nature sémiotique du document audiovisuel* si on veut s'engager dans une telle voie.

Et c'est cela le principal enjeu du projet ASA-SHS – un enjeu qui n'est pas ou pas prioritairement technologique ! En effet, tout en s'appuyant sur un environnement technologique très adapté aux besoins spécifiques du traitement sophistiqué de corpus audiovisuels (production, description, enrichissement, publication, ...) déjà réalisé en coopération avec l'INA et/ou dans le cadre de ses différents projets de R&D français et Européens, le projet ASA-SHS ne fait « que » utiliser ces technologies tout en mettant l'accent sur les apports des sciences humaines (notamment de l'analyse du discours et de la sémiotique) pour « solutionner » le problème de l'appropriation d'un document audiovisuel existant par son utilisateur.

1.3 Description des travaux : programme scientifique/For each specific aim: a proposed work plan should be described (including preliminary data, work packages and deliverables)

1.3.1 Les 5 grandes tâches du projet ASA-SHS

Le projet ASA-SHS vise la mise en place d'un cadre sémiotique et cognitif commun pour la constitution, la description et surtout la (re-)publication de corpus audiovisuels numériques en SHS. Ce cadre porte le nom d'« atelier sémiotique » qui fait allusion d'une part au rapport intrinsèque, dans ce projet, entre la théorie (sémiotique) et la pratique (de « traitement » de corpus audiovisuels en SHS), et d'autre part à la notion de « bricolage » employée si merveilleusement par Claude Lévi-Strauss pour caractériser et expliciter l'activité humaine de la création intellectuelle.

Un des enjeux principaux du cadre sémiotique et cognitif commun est de permettre une prise en compte bien plus riche, plus adaptée aux attentes des communautés de lecteurs, du contenu intellectuel (scientifique, culturel, idéologique, ...) des ressources audiovisuelles en SHS telles que celles des AAR mais aussi d'autres programmes et initiatives en France ou ailleurs. Or, cela ne peut se faire que par l'utilisation des différents paramètres qui concourent à la définition de la spécificité sémiotique et cognitive intrinsèque de ces ressources (cf. à ce propos le cadre théorique général exposé succinctement en 1.2.1) – ce qui est bien loin le cas, aujourd'hui, dans la très grande majorité des médiathèques ou autres bibliothèques audiovisuelles ou multimédias en ligne.

Afin de pouvoir réaliser ses objectifs, le projet ASA-SHS s'appuie sur cinq tâches principales (figure 2) décomposées, chacune en plusieurs sous-tâches (cf. ci-après, les chapitres 1.3.2 à 1.3.6).

N°	Tâche	Intitulé	Durée
1	Tâche 1	Spécification et explicitation du contexte « utilisateur » du projet	T-0 – T-12
2	Tâche 2	Description et modélisation sémiotiques du discours audiovisuel en SHS	T-6 – T-18
3	Tâche 3	Spécification et réalisation des langages de description (i.e. des ontologies du projet) et des modèles de (re-)publication de corpus audiovisuels en SHS	T-18 – T-30
4	Tâche 4	Expérimentations, développements des pilotes, tests et validation et généralisation de l'approche sémiotique sur tout type de corpus audiovisuels	T-24 – T-48
5	Tâche 5	Gestion du projet, communication et valorisation de ses résultats	T-0 – T-48

(figure 2 : les cinq grandes tâches composant le projet ASA-SHS)

La **première tâche** réunit l'ensemble des activités consacrées à l'explicitation du contexte opérationnel du projet ASA-SHS qui prend comme point de départ l'état actuel des services en ligne proposés par le portail du programme des Archives Audiovisuelles de la recherche (AAR)¹¹:

1. analyse détaillée des attentes et besoins des communautés d'utilisateurs de ressources audiovisuelles en SHS afin de pouvoir orienter les travaux d'analyse sémiotique de corpus audiovisuels choisis ;
2. localisation, description et évaluation comparatives de l'existant (en termes de traitement de corpus audiovisuels et de diffusion de corpus traités sous forme de genres de publication spécifiques) ;
3. mise en place d'un référentiel de « bonnes pratiques » pour les différents travaux de R&D dans le cadre du projet AAR (référentiel qui inclura également des réalisations propres au programme des AAR telle que celle du site déjà cité consacré à la diversité linguistique et culturelle¹²) ;
4. ensemble avec trois groupes d'utilisateurs différents (enseignants, étudiants et « professionnels » des SHS), spécification d'un ensemble scénarios d'expérimentation permettant de tester concrètement les avancées du projet (cf. tâche 4) ;
5. constitution de trois corpus de travail en *histoire*, *littérature* et *archéologie* se nourrissant du fonds audiovisuel du programme des AAR (qui est « ouvert » dans ce sens qu'il s'enrichit d'environ 40 à 50 heures de nouvelles productions audiovisuelles par mois)
6. mise en place, enfin, d'un ensemble de critères permettant d'évaluer d'une manière critique le progrès du projet ASA-SHS.

La **deuxième tâche** est consacrée à la description sémiotique des trois corpus de ressources audiovisuelles en *histoire*, *littérature* et *archéologie* - ressources produites depuis fin 2001 dans le cadre du programme des AAR et composées d'enregistrements d'entretiens avec de chercheurs dans pratiquement toutes les disciplines en SHS, de conférences, de séminaires, de cours etc. mais aussi de documentaires, de reportages, voire de prises de vue « de terrain » (ethnographique, sociolinguistique, etc.). Les résultats de cette tâche seront ensuite utilisés, dans la tâche 3, pour le développement des instruments et outils méthodologiques et techniques nécessaires pour procéder à des descriptions, indexation, annotations et publications *personnalisées* (i.e. dépendant des besoins et intérêts des auteurs et aussi des objectifs d'une publication). En s'appuyant sur le cadre théorique général esquissé dans la figure 1, cette deuxième tâche embrasse les activités principales suivantes :

1. comprendre et expliciter la **structure thématique** (les *topoi* ou encore, comme on dit en analyse littéraire, les *sujets*) qui sous-tend cette production intellectuelle et scientifique afin de pouvoir l'utiliser ensuite pour la construction d'une ontologie (d'un vocabulaire de description) du domaine concerné ;
2. comprendre et expliciter les principales **formes rhétoriques** et narratives utilisées pour thématiser une situation, un événement, un problème afin de pouvoir utiliser ces résultats dans le développement d'une ontologie rhétorique et narrative permettant de nuancer une description thématique par des critères rhétoriques forts importants pour une adaptation de ressources audiovisuelles à des usages spécifiques (il s'agit ici, par exemple, de pouvoir distinguer entre l'*exposé scientifique* ou *didactique* d'un problème ou événement, l'*exemplification* du même problème ou événement, l'*explication* du problème, la production d'*arguments pour ou contre*, etc.) ;
3. identifier et décrire les parties **textuelles fonctionnelles** dans les ressources audiovisuelles telles que les *introductions* et *conclusions*, les *renvois*, les *rappels*, les « *moments forts* », les « *digressions* » etc. afin de pouvoir réutiliser ces données dans certains contextes spécifiques d'enrichissements, d'adaptations, d'explicitations, voire d'hierarchisation de contenus scientifiques ;
4. décrire notamment dans les enregistrements audiovisuels de situations de terrain et dans les différents documentaires et reportages, l'**expression audiovisuelle** à proprement parler, i.e. la « traduction » d'un contenu dans le medium audiovisuel (cadres, plans, synchronisation de différents éléments d'expression), etc. afin de proposer une ontologie de l'expression audiovisuelle comme un outil en plus pour décrire et adapter le contenu d'un corpus audiovisuel à un usage, un contexte particulier ;
5. expliciter, enfin, la structure typique d'un ensemble de **genres de publications** notamment scientifiques et pédagogiques tels que dossiers thématiques, syllabus, portefeuille de lecture, ... ainsi que celle de constructions ou parcours narratifs permettant une exploration guidée et finalisée d'un univers de connaissances (en prenant comme référence analogique, par exemple, les parcours de type « exploration », « enquête », « découverte », etc.).

¹¹ cf. <http://www.archivesaudiovisuelles.fr/FR/>

¹² cf. <http://www.languescultures.eu/>

La **troisième tâche** dépend d'une manière critique, comme déjà dit, des résultats de la deuxième, i.e. des résultats de l'analyse sémiotique des trois corpus audiovisuels de travail en histoire, littérature et archéologie. Elle est consacrée à la mise en place des outils et instruments nécessaires pour procéder, ensuite (cf. tâche 4) aux expérimentations devant montrer la pertinence et l'intérêt de l'approche sémiotique pour les différentes activités de la constitution, de la production, de la publication et de l'appropriation de contenus scientifiques (ou de nature plus largement culturelle) en SHS. Les outils et instruments relèvent de trois catégories : 1) les **ontologies**, 2) les **modèles de description** du contenu audiovisuel et 3) les **modèles et formulaires de publication** de corpus audiovisuels

Les ontologies sont des sortes de vocabulaires de description de domaines et forment une partie du métalangage de description des corpus audiovisuels. Dans le cadre du projet ASA-SHS et conformément au cadre théorique sémiotique, cinq ontologies spécifiques doivent être développées : une **ontologie textuelle** (i.e. des types de segments composant un document audiovisuel), une **ontologie du domaine**, une **ontologie rhétorique**, une **ontologie narrative** et une **ontologie de l'expression audiovisuelle**.

Elles – ces cinq ontologies – concourent ensemble à la spécification de **modèles de description** de contenu audiovisuel sous forme de schémas ou graphes conceptuels. Ces schémas ou graphes conceptuels représentent des structures de contenu audiovisuel typiques ou récurrentes de niveau plus ou moins spécialisés (et donc peu utilisable en dehors d'un corpus donné) ou, au contraire assez, voire très généraux (et donc utilisables tels quels ou moyennant certaines modifications pour la description de toutes sortes de ressources audiovisuelles). Ensemble, ils forment une bibliothèque de modèles dont se sert l'utilisateur pour procéder, selon ses besoins et désirs, à des description plus ou exhaustives, plus ou moins partielles, voire partiales, plus ou moins générales ou, au contraire, spécialisées de « son » corpus audiovisuel.

Les **modèles de publication** représentent la structure essentielle de tel ou tel genre de publication (scientifique, pédagogique, ...) et les **formulaires** des sortes de protocoles à respecter lors du traitement d'un corpus audiovisuel en vue de sa (re-)publication selon un modèle de publication choisi.

Ensemble, ontologies de description et modèles et formulaires de publication composent donc les **instruments intellectuels et techniques centraux** de l'atelier sémiotique à l'aide duquel un utilisateur ou un groupe d'utilisateurs pourra travailler sur des corpus audiovisuels selon ses besoins et intérêts. La troisième tâche est donc par conséquent consacrée aux quatre activités principales suivantes :

1. La conception et au développement des ontologies du projet ASA-SHS. Il s'agit notamment des cinq ontologies : textuelle, du domaine, rhétorique, narrative et de l'expression audiovisuelle. Ces ontologies prendront comme point de départ une ontologie déjà existante et utilisée pour l'indexation de certains corpus audiovisuel en anthropologie et études culturelles, i.e. l'ontologie v_6 créée à l'aide d'un éditeur d'ontologie Onto Editor et complètement intégrée dans l'environnement de description, indexation et publication utilisé à l' dans le cadre du programme AAR. Ceci dit, l'actuelle ontologie utilisée – v_6 – sera profondément modifiée dans sa structure d'organisation même et acquerra un niveau d'abstraction suffisamment élevé et riche pour pouvoir être utilisée sous forme de « modules » spécifiques à un domaine de connaissances choisi.
2. La définition de modèles de description du contenu audiovisuel sous forme d'une bibliothèque de schémas ou graphes conceptuels qui sera mise à la disposition des utilisateurs de l'atelier sémiotique. De nouveau, cette bibliothèque s'appuiera et enrichira très considérablement une petite bibliothèque d'une soixante de modèles de description actuellement utilisés dans le cadre du déploiement des nouveaux services sur le site des AAR (nouveaux services de type vidéo-glossaire ou vidéo-encyclopédie).
3. La spécification et la réalisation informatique de nouveaux modèles de publication de corpus audiovisuels sous forme d'une part de dossiers scientifiques et pédagogiques de tout genre et d'autre part de « vidéos virtuels », i.e. de parcours narratifs créant de chemins d'exploration entre des segments audiovisuels disjoints les uns les autres.
4. L'intégration des ontologies ASA-SHS, de la bibliothèque des modèles de description du contenu audiovisuel et des modèles et formulaires de publication dans l'environnement numérique actuel de travail appelé *Interview* (et co-développé avec l'INA Recherche dans le cadre d'une convention entre INA et FMSH) avec comme but de pouvoir l'utiliser dans les expérimentations (tâche 4).

La **quatrième tâche** est réservée aux expérimentations de description, d'indexation, d'enrichissement et d'adaptation de corpus audiovisuels provenant du fonds AAR, et à leurs publications adaptées à des usages scientifiques et pédagogiques, voire journalistiques et « professionnels »

La quatrième tâche est consacrée plus particulièrement au développement du pilote du projet ASA-SHS se basant sur les trois corpus de travail en histoire, littérature et archéologie. Les ressources audiovisuelles des trois corpus seront traitées, i.e. décrites, indexées et annotées à l'aide de la bibliothèque des modèles de description développée dans la tâche 3 et publiées ensuite pour des usages de recherche, des usages éducatifs et des usages professionnels. Cette démarche fera l'objet de procédures d'évaluation répétées par les quatre groupes d'utilisateurs principaux du projet ASA-SHS, i.e. les chercheurs en SHS, les enseignants, les étudiants et les « professionnels » dépendant crucialement de la recherche en SHS (médiateurs socioculturels, travailleurs sociaux, éditeurs, journalistes et auteurs de vulgarisation, ...).

Enfin un des enjeux principaux de cette quatrième tâche concerne celle de la généralisation de l'approche sémiotique à l'ensemble du corpus des 5000 heures de vidéos des AAR. En effet, outre les expérimentations « à grandeur réelle », un des buts essentiels de cette tâche est de démontrer la pertinence et l'intérêt de cette approche sémiotique pour le traitement de corpus audiovisuels de tout genre – pas seulement de corpus de type « discours scientifique », mais aussi, par exemple, de corpus audiovisuels tels qu'on peut les trouver en archéologie (enregistrements de fouilles), en anthropologie et ethnologie (enregistrements de terrain), dans le domaine de la préservation du patrimoine culturel (enregistrements des objets possédant un intérêt patrimonial).

Mais, même au-delà de ces références qui sont celles du projet ASA-SHS, l'hypothèse ici est que l'approche développée ici, pourrait parfaitement bien servir (tout en l'adaptant, bien sûr) à des productions audiovisuelles « amateurs » telles qu'on les trouve sur les sites emblématiques tels que *Daily Motion* ou *YouTube* ... En effet, les modèles de description construits sur la base des différentes ontologies représentent des structures de contenu audiovisuel typique. Or, rien n'empêche de développer des mini-bibliothèques de modèles stéréotypés de description simples et généraux qui pourrait être utilisé par n'importe qui, peu importe son niveau de connaissances.

Ceci étant, il va sans dire que cette quatrième tâche dépend crucialement des résultats de la troisième tâche qui, elle, présuppose les résultats de la deuxième tâche. Les principales activités composant cette tâche sont les suivantes :

1. Mise en place des sites pilotes du projet qui, en s'appuyant sur les acquis d'un site tel que celui du DLC, documentera et exemplifiera les avancées concrètes du projet, l'intérêt de ses avancées par rapport à l'existant et leurs conformités par rapport indicateurs de succès. Mais ils constitueront à terme des sites à part entier avec leurs communautés d'utilisateurs, services et ressources.
2. Travail concret de description, indexation republication ... de corpus audiovisuels choisis (dans les trois domaines : histoire, littérature et archéologie) à l'aide des outils sémiotiques développés dans le cadre du projet ASA-SHS.
3. Tests et validations des travaux par différents groupes d'utilisateurs composés de chercheurs d'une part et d'enseignants, étudiants d'autre part. Un quatrième groupe d'utilisateurs sera composé de « professionnels » d'utilisation de connaissances en SHS : éditeur (à commencer par les éditions FMSH qui viennent de mettre en place un programme de production de DVD à partir du fonds audiovisuels des AAR), journalistes et médiateurs socioculturels (travaillant notamment dans les deux domaines de la communication sociale et de la communication au service de la médiation linguistique et interculturelle).
4. Documentation technique, méthodologique et didactique de l'ensemble des procédés nécessaires pour une « bonne » réutilisation des résultats du projet en dehors de la FMSH.
5. Généralisation de l'approche sur l'ensemble du fonds des AAR, notamment sur les enregistrements audiovisuels provenant de fouilles archéologiques, d'enregistrements sociolinguistiques (écoles de langue, dialogues et interactions, enregistrées, ...), sur les « road movies » (documentations audiovisuelles de villes, de pays, ...) et, enfin, sur l'ensemble des documentaires. Par ailleurs, cette activité inclura des travaux de coopération avec d'autres entités institutionnelles. Parmi les partenaires pressentis, on compte l'INRAP (l'Institut National de Recherche en Archéologie Préventive) avec son fonds fort intéressant composé d'enregistrements audiovisuels de fouilles en France, le réseau des MSH en France, un certain nombre de laboratoires d'anthropologie et d'ethnologie, etc.

Enfin, la **cinquième tâche** réunit l'ensemble des activités nécessaires à la bonne conduite du projet lui-même ainsi qu'à la communication et dissémination de ses résultats.

1. Une première activité est concernée par la coordination technique et scientifique du projet. Cette activité, somme toute assez peu compliquée dans le cas concret de ce projet qui ne compte qu'un seul partenaire contractuel avec l'ANR, aura à sa charge la coordination et le suivi aussi exact que possible des différentes activités du projet selon l'échéancier spécifié ci-après ; les rapports avec l'ANR d'une part et les différents services administratifs de la FMSH (budget, comptabilité, ressources humaines, service informatique, ...)

d'autre part. Elle veillera également, bien évidemment, à la production, à temps, de toutes les fournitures identifiées ci-après.

2. Une deuxième activité réunit l'ensemble des actions de communication scientifique des avancées et des résultats du projet sous forme de participation à des colloques, workshops, etc. en France et ailleurs et aussi sous forme de l'organisation de deux workshops à la FMSH à Paris. Des publications dans des revues et une monographie en langue française sont également prévues.
3. Une troisième activité très importante réunit d'une part les actions de sensibilisation et de formation à destination d'utilisateurs potentiels des résultats du projet ASA-SHS et d'autre part les actions de communication de valorisation à proprement parler sous forme de communiqué de presse, de réalisation de brochures et de dépliants ou encore sous forme d'un film institutionnel. Enfin, sont prévues également des journées « porte ouverte » à la FMSH.

Dans les cinq sous-chapitres ci-après, chacune des cinq tâches est présentée plus en détail.

	Programme BLANC	Réservé à l'organisme gestionnaire du programme N° de dossier : ANR-08-XXXX-00 Date de révision :
	Document scientifique associé	Edition 2008

1.3.2 Tâche 1 : Spécification et explicitation du contexte opérationnel du projet

La tâche 1 est décomposée en 7 activités plus spécifiques qui, comme déjà dit, doivent expliciter le contexte opérationnel du projet (cf. 1.3.1).

Sous-Tâche	Objectifs	Résumé des activités	Fourniture	Durée
Sous-Tâche 1.1 <i>Analyses des attentes des principaux groupes d'utilisateurs de contenus audiovisuels en SHS</i>	1. Expliciter les contraintes culturel et professionnel dans la mise en place du cadre sémiotique et des outils de « traitement » de corpus audiovisuels (cf. T 2 et T3). 2. Disposer d'un moyen de comparaison critique entre les réalisations du projet et un ensemble d'attentes données	<ul style="list-style-type: none"> ▪ Constitution formelle de quatre groupes de travail composé de chercheurs ; d'enseignants ; d'étudiants et de « professionnels » et définition d'un agenda de travail pour toute la durée du projet. ▪ Enquêtes sous forme d'un questionnaire en ligne et de réunions « qualitatives » auprès d'un panel représentatif de personnes faisant partie de ces quatre groupes d'utilisateurs. 	Rapport technique « Contexte du Projet » (RT1)	T-0 – T-4
Sous-Tâche 1.2 <i>Spécification des scénarios d'usage pour le projet ASA-SHS</i>	1. Définition et validation de scénarios qui accompagneront et exemplifieront le travail d'expérimentations (T 4)	Avec les groupes de travail, définition d'un petit ensemble de 4 à 6 scénarios d'utilisation de ressources audiovisuelles SHS pour : <ul style="list-style-type: none"> ▪ la recherche académique ▪ l'enseignement formel (supérieur, secondaire) ▪ l'enseignement informel (« sensibilisation culturelle ») ▪ l'usage professionnel (édition ; travail sur terrain) 	Rapport technique « Contexte du Projet » (RT1)	T-0 – T-5
Sous-Tâche 1.3 <i>Spécification des pilotes du projet</i>	1. Définition du résultat démontrable du projet ASA-SHS (T 4)	<ul style="list-style-type: none"> ▪ Spécification de la structure conceptuelle du pilote du projet (structure du site, services, interfaces, ...) 	Rapport technique « Contexte du Projet » (RT1)	T-1 – T-5
Sous-Tâche 1.4 <i>Constitution des corpus de travail</i>	1. Réunir l'ensemble des ressources audiovisuelles qui doivent être traitées (décrites, indexées, annotées, ...) par	<ul style="list-style-type: none"> ▪ Identification des trois corpus de travail en histoire, littérature et archéologie. ▪ Localisation des ressources pertinentes dans le fonds des AAR 	Rapport technique « Contexte du Projet » (RT1)	T-2 – T-10

	l'atelier sémiotique (T 4)	<ul style="list-style-type: none"> ▪ Eventuellement production de nouvelles ressources (entretiens, prises de vue, ...) ou collecte de ressources appartenant à d'autres « fournisseurs de contenu » ▪ Publication des trois corpus sous forme de médiathèques en ligne. 	2/ Corpus (OR1) accessibles en ligne sous forme d'une médiathèque simple.	
Sous-Tâche 1.5 <i>Analyse critique et comparative de l'existant</i>	1. Disposer d'un référentiel de « best practices » (bonnes pratiques) pour les activités du projet ASA-SHS (T2, T3, T4)	<ul style="list-style-type: none"> ▪ Analyse comparative des prestations de services « contenu audiovisuel » fournis par les principaux fournisseurs de contenu en France (Canal U, Canal IRD, TV5, Cité des Sciences, ...), en Europe (Deutsche Welle, ...) et en Amérique du Nord (Research Channel, ...) ▪ Identification et analyse comparative des principales pratiques de traitement de corpus audiovisuels ▪ Identification des outils de traitement et de publication existants 	Rapport technique « Contexte du Projet » (RT1)	T-6 – T-10
Sous-Tâche 1.6 <i>Localisation d'outils métalinguistiques déjà existants</i>	1. Identification des outils métalinguistiques déjà existants pouvant être utilisés au moins partiellement dans l'élaboration des ontologies et modèles de description (T 3)	<ul style="list-style-type: none"> ▪ Thesaurus et terminologies en sciences humaines et sociales (thesaurus de l'UNESCO, ...) ▪ Thesaurus et terminologies en histoire, littérature et archéologie ▪ Thesaurus et terminologies en analyse du discours, du texte, du document filmique, ... ▪ Ontologies existantes 	Rapport technique « Contexte du Projet » (RT1)	T-6 – T-12
Sous-Tâche 1.7 <i>Spécification d'un ensemble d'indicateurs de réussite et mise en place de procédés de test et d'évaluation</i>	1. Permettre un « monitoring », une supervision opérationnelle de l'avancement du travail dans le cadre du projet ASA-SHS (toutes tâches concernées)	<ul style="list-style-type: none"> ▪ <i>Indicateurs quantitatifs</i> : nombre d'heures indexées/heure ; nombre de modèles de description utilisés/nombre total de modèles ; nombre de visites sur le site du pilote ; ... ▪ <i>Indicateurs qualitatifs</i> : opinions des membres des quatre groupes relatives aux ontologies, modèles de description ; modèles de publication ; satisfaction des utilisateurs des nouveaux services de contenu audiovisuel ; ▪ <i>Procédés de test et de validation</i> des résultats de l'analyse sémiotique, des ontologies, modèles de description et modèles de publication ; des corpus audiovisuels décrits, indexés et republiés, 	Rapport technique « Contexte du Projet » (RT1)	T-6 – T-7

1.3.3 Tâche 2 : Description et modélisation sémiotiques du discours audiovisuel en SHS

Tâche 2 est décomposée en 4 activités plus spécifiques qui, comme déjà dit (1.3.1) visent à une description sémiotique systématique du discours scientifique audiovisuel dans les trois corpus de travail choisis (histoire – littérature – archéologie) et dont les résultats serviront à l'élaboration des ontologies, modèles de description et modèles et formulaires de publication (T3).

Sous-Tâche	Objectifs	Résumé des activités	Fourniture	Durée
Sous-Tâche 2.1 <i>Description de la structure thématique</i>	<ol style="list-style-type: none"> 1. Fournir une liste structurée et commentée des principaux <i>sujets</i> ou <i>topoi</i> organisant d'une manière typique les discours scientifiques dans les trois corpus choisis. 2. Identifier les thèmes servant d'input à la construction de l'ontologie du domaine (T3) 	<ul style="list-style-type: none"> ▪ En utilisant un canevas méthodologique déjà existant et utilisé pour la description de productions audiovisuelles dans le cadre du programme AAR, identification et description des structures thématiques organisant le discours dans les corpus choisis au préalable (cf. T. 1.4). ▪ Produire un document listant, nommant et explicitant l'ensemble des thèmes supposés nécessaires pour rendre compte du contenu énoncé dans les trois corpus. 	Dossier Scientifique « Sémiotique » (DS1)	T-6 – T-18
Sous-Tâche 2.2 <i>Description narrative et rhétorique</i>	<ol style="list-style-type: none"> 1. Fournir une liste structurée et commentée des principaux genres de discours utilisés dans le corpus de travail. 2. Fournir une liste structurée et commentée de principales structures d'intégration syntagmatique caractérisant le développement du discours dans le corpus concerné 3. Identifier les éléments rhétoriques et narratifs servant d'input à la construction des deux ontologies narratives et rhétoriques (T3) 	<ul style="list-style-type: none"> ▪ Description des structures syntagmatiques (i.e. narratives) et rhétoriques typiques organisant le discours dans les corpus choisis. ▪ Production d'un document listant, nommant et explicitant les genres de discours (tels que <i>exposé scientifique, chronologie, argumentaire, ...</i>) ainsi que les structures d'intégration syntagmatique (de type : <i>tel segment présuppose tel autre segment ; tel segment est un développement de tel autre segment, tel segment est en contradiction avec un tel autre segment, ...</i>) 	Dossier Scientifique « Sémiotique » (DS1)	T-6 – T-18
Sous-Tâche 2.3 <i>Description textuelle et audiovisuelle</i>	<ol style="list-style-type: none"> 1. Fournir une liste structurée et commentée des principaux types fonctionnels de segments composant les 	<ul style="list-style-type: none"> ▪ Description des formes typiques de la mise en scène audiovisuelle des informations produites dans les corpus choisis ▪ Identification et description des principaux types 	Dossier Scientifique « Sémiotique » (DS1)	T-6 – T-18

	<p>vidéos qui font partie du corpus de travail.</p> <ol style="list-style-type: none"> Fournir une liste structurée et commentée des principaux éléments de mise en scène audiovisuelle d'un contenu scientifique. Identifier les types de segments et les éléments de mise en scène audiovisuelle servant d'input à la construction des deux ontologies textuelle et d'expression audiovisuelle (T3) 	<p>fonctionnels de segments audiovisuels composant une vidéo dans le corpus concerné (types de segments tels que <i>introduction, conclusion, renvois, citation, rappel, digression, etc.</i>)</p> <ul style="list-style-type: none"> Production d'un document listant, nommant et explicitant l'ensemble des éléments de mise en scène audiovisuelle ainsi que l'ensemble des types fonctionnels de segments audiovisuels 		
<p>Sous-Tâche 2.4</p> <p><i>Description de genres de publication</i></p>	<ol style="list-style-type: none"> Décrire les scénarios typiques d'un ensemble de genres de publication scientifique et pédagogique Produire des « maquettes » ou schémas logiques pour leur implémentation sous forme de <i>templates</i> ou modèles informatiques (T3). 	<ul style="list-style-type: none"> Description sémiotique et modélisation de la structure typique de genres de publication à caractère scientifique et surtout pédagogique (syllabus, portefeuille, parcours narratifs à caractère pédagogiques, etc.). Production d'un document contenant les maquettes ou schémas logiques de ces genres de publication. 	Dossier Scientifique « Genres de Publication » (DS2)	T-12 – T-18

1.3.4 Tâche 3 : Spécification et réalisation des langages de description et de modèles de (re-)publication

Tâche 3 est décomposée en 4 activités plus spécifiques qui, comme déjà dit, s'appuient sur les résultats des travaux réalisés en T2 afin de développer concrètement les ontologies, la bibliothèque de modèles de description de corpus audiovisuels et modèles de (re-)publication.

Sous-Tâche	Objectifs	Résumé	Fourniture	Durée
<p>Sous-Tâche 3.1</p> <p><i>Développement des ontologies du projet</i></p>	<ol style="list-style-type: none"> Développer les cinq ontologies du projet à l'aide de l'outil Onto Editor tout en se servant, comme point de départ, d'une 	<ul style="list-style-type: none"> Sur la base des descriptions sémiotiques précédentes (T 2.1) et en tenant compte de l'existant (ontologie v6 déjà utilisée) spécification du vocabulaire et de la structure du vocabulaire des cinq ontologies du 	Vocabulaires d'ontologies « ASA-SHS » sous forme de fichiers xml (OR2)	T-18 – T-24

	ontologie déjà existante (v6) et utilisée pour l'indexation de corpus audiovisuels dans le cadre du programme AAR	<p>projet : textuelle, thématique, narrative, rhétorique et d'expression audiovisuelle.</p> <ul style="list-style-type: none"> ▪ Réalisation des cinq ontologies à l'aide de l'outil d'édition d'ontologies Onto Editeur (outil développé par l' dans le cadre du projet de R&D Saphir pour ses besoins en termes de production d'ontologies). 		
Sous-Tâche 3.2 <i>Développement des ontologies du projet</i>	1. Spécifier la bibliothèque de modèles (i.e. schémas ou graphes) de description de corpus audiovisuels	<ul style="list-style-type: none"> ▪ Sur la base des résultats de la description sémiotique des trois corpus de travail (T2) et en s'appuyant sur les vocabulaires des cinq ontologies ASA-SHS développées dans la tâche 3.1 – spécification d'une liste (ouverte) de schémas ou modèles de description sous forme de petits graphes représentant des configurations de contenu typiques et pouvant être utilisés tels que ou d'une manière personnalisée par les utilisateurs pour décrire et indexer les contenus audiovisuels 	Schémas ou graphes de description « ASA-SHS » sous forme de fichiers xml (OR3)	T-22 – T-26
Sous-Tâche 3.3 <i>Implémentation des modèles de (re-)publication</i>	1. Réaliser d'un point de vue informatique les modèles de publication	<ul style="list-style-type: none"> ▪ Réalisation informatique des modèles de (re-)publication définis et décrits dans T2.4 	« Templates » de publication (fichiers xml/xslt représentant des structures de publication adaptables) (OR4)	T-18 – T-24
Sous-Tâche 3.4 <i>Intégration informatique</i>	1. Intégrer, dans l'environnement de travail Interview, les ontologies, modèles de description, modèles et formulaires de publication	<ul style="list-style-type: none"> ▪ Intégration des nouveaux outils de description (i.e. les ontologies du projet, la bibliothèque des modèles de description ainsi que les modèles et formulaires de publication) dans l'environnement de traitement de corpus audiovisuels (description, indexation, enrichissement, traduction, ...) appelé Interview et utilisé dans le cadre du programme AAR. 	Mise à jour de l'environnement de travail de traitement de corpus audiovisuels appelé Interview (OR5)	T-23 – T-26

1.3.5 Tâche 4 : Expérimentations, développement des pilotes, tests et validation et généralisation de l'approche sur l'ensemble du fonds des AAR

Tâche 4 est décomposée en 6 activités plus spécifiques qui, comme déjà dit, doivent expliciter le contexte opérationnel du projet en tenant compte à la fois des spécificités du programme AAR, de ses utilisateurs réels et potentiels ainsi que d'une manière plus générale, de l'existant.

Sous-Tâche	Objectifs	Résumé	Fourniture	Durée
Sous-Tâche 4.1 <i>Préparation des sites pilotes</i>	1. Mise en place des pilotes du projet servant à démontrer et valider ses avancements et résultats	<ul style="list-style-type: none"> ▪ En tenant compte des inputs de T1 (notamment T1.3, T1.2, T1.4) préparation des pilotes du projet sous forme de sites « médiathèque en ligne » servant à la fois à la diffusion des corpus audiovisuels choisis, à la publication des avancements du projet et aux interactions avec les communautés d'utilisateurs (questionnaires enligne, listes de diffusion,) 	Pilote du projet (OR6)	T-24 – T-27
Sous-Tâche 4.2 <i>Premier traitement expérimental des corpus de travail</i>	1. Tester et valider les outils développés en T3 sur une partie des corpus de travail défini en T1.3	<ul style="list-style-type: none"> ▪ En tenant compte des inputs de T1 (notamment T1.2 et T1.4) et en utilisant une première sélection de modèles de description (T 3.2) premier traitement (description, indexation, enrichissement et republication) de test sur un petit sous-ensemble de vidéos faisant partie des corpus audiovisuels choisis en T1.4 avec comme but de tester les outils métalinguistiques et leurs réalisations informatiques (T3). 	Rapport technique « Expérimentations » (RT2) Pilote du projet (OR6)	T-25 – T-28
Sous-Tâche 4.3 <i>Traitement à grandeur réelle du corpus de travail</i>	1. Réalisation « à grandeur réelle » des pilotes du projet sur la base des inputs en T.1 et T.3	<ul style="list-style-type: none"> ▪ Traitement (description, indexation, enrichissement et republication) extensif des corpus audiovisuels choisis en T1.4 et diffusion sur les sites pilotes du projet sous forme de publications enligne adaptées à des usages spécifiques dans la recherche et l'enseignement (y inclut l'apprentissage dit informel) en respectant les scénarios définis en T 1.2 	Rapport technique « Expérimentations » (RT2) Pilote du projet (OR6)	T-28 – T-48
Sous-Tâche 4.4 <i>Tests et validation du pilote</i>	1. Détecter et évaluer les écarts entre les attentes des principaux groupes	A l'aide des quatre groupes de travail : <ul style="list-style-type: none"> ▪ Procéder à une évaluation minutieuse tout au long de la réalisation des pilotes du projet ASA-SHS. 	Rapport technique « Expérimentations » (RT2)	T-28 – T-36

	<p>d'utilisateurs et les réalisations effectives du projet.</p> <p>2. Le cas échéant, procéder aux modifications et améliorations nécessaires et possibles dans le cadre du projet ASA-SHS.</p>	<ul style="list-style-type: none"> ▪ Procéder, autant que possible, aux modifications et/ou améliorations souhaitées. ▪ Documenter les évaluations des pilotes. ▪ Mettre en place des évaluations des pilotes à l'aide de questionnaires en ligne visant des panels représentatifs des quatre principaux groupes d'utilisateurs. 	Pilote du projet (OR6)	
<p>Sous-Tâche 4.5</p> <p><i>Travail d'expérimentation et de publication avec des groupes d'utilisateurs choisis</i></p>	<p>1. Expérimenter le travail sémiotique sur des corpus choisis avec des utilisateurs à profils et intérêts différents.</p> <p>2. Evaluer l'intérêt général de l'approche sémiotique pour le traitement des corpus audiovisuels visant des utilisations spécifiques et adaptées à des contextes d'usage particuliers.</p>	<ul style="list-style-type: none"> ▪ Constituer, avec l'aide des quatre groupes de travail, des nouveaux groupes d'utilisateurs « recrutés » parmi les chercheurs, les enseignants, les étudiants et les « professionnels » en SHS. ▪ Former ces groupes dans le travail sémiotique portant sur des corpus audiovisuels et leur réutilisation dans des contextes d'usage spécifiques. ▪ Choisir avec ces groupes de petits corpus de travail. ▪ Identifier des contextes d'usages spécifiques. ▪ « Faire travailler » les groupes d'utilisateurs avec l'atelier sémiotique (i.e. l'environnement Interview augmenté des ontologies ASA-SHS, des modèles de description et des modèles de publication). ▪ Publier ces travaux sur le site ASA-SHS diffusant les pilotes du projet. ▪ Evaluer avec les utilisateurs le travail entrepris et les résultats. ▪ Produire un rapport technique. 	<p>Rapport technique « Expérimentations » (RT2)</p> <p>Pilote du projet (OR6)</p>	T-36 – T-48
<p>Sous-Tâche 4.6</p> <p><i>Généralisation de l'approche sémiotique</i></p>	<p>1. Appliquer l'approche sémiotique à l'enrichissement et la valorisation de l'ensemble du fonds audiovisuel des AAR.</p> <p>2. Montrer que cette approche peut servir au traitement de toute sorte de corpus – et pas seulement au traitement de corpus SHS.</p>	<ul style="list-style-type: none"> ▪ En utilisant l'atelier sémiotique, produire des republications d'autres corpus faisant partie du programme des AAR. ▪ Améliorer les prestations du site portail AAR en s'appuyant sur les résultats du projet ASA-SHS. ▪ Travailler avec d'autres « fournisseurs de contenu » sur des corpus de leurs choix à la fois pour exemplifier l'approche sémiotique et pour montrer sa pertinence et son intérêt en dehors du programme AAR et de son fonds audiovisuel. 	<p>Rapport technique « Expérimentations » (RT2)</p> <p>Pilote du projet (OR6)</p>	T-36 – T-48

1.3.6 Tâche 5 : Gestion, communication et valorisation du projet

Tâche 5 est décomposée en 5 activités plus spécifiques qui, comme déjà dit, doivent expliciter le contexte opérationnel du projet en tenant compte à la fois des spécificités du programme AAR, de ses utilisateurs réels et potentiels ainsi que d'une manière plus générale, de l'existant.

Sous-Tâche	Intitulé	Résumé	Fourniture	Durée
Sous-Tâche 5.1 <i>Gestion technique et scientifique du projet</i>	<ol style="list-style-type: none"> Assurer une coordination scientifique et technique efficace du projet, conformément aux objectifs développés dans la description technique et scientifique du projet Assurer la progression du projet suivant l'échéancier établi. Assumer le rôle d'intermédiaire entre les services administratifs de la FMSH concernés par la gestion du projet et l'ANR. 	<ul style="list-style-type: none"> Constitution d'un comité de projet composé de représentants de la FMSH, des groupes de travail et des institutions associées au projet ASA-SHS. Organisation, coordination et suivi scientifique du travail scientifique et technique. Organisation de réunions de travail, techniques et plénières. Organisation des revues avec l'ANR. Rédaction de tous les rapports techniques et scientifiques. Réalisation des rapports de gestion de projet : rapports semestriels, rapport final de synthèse et récapitulatif des dépenses. Comptabilité des ressources financières du projet. 	<p>Rapport d'avancement N°1 – N°8 (DM1)</p> <p>Rapport Final (DM2)</p>	T-0 – T-48
Sous-Tâche 5.2 <i>Constitution de groupes de travail et d'un réseau de partenaires</i>	<ol style="list-style-type: none"> Faire en sorte que le projet ASA-SHS soit co-piloté et évalué en permanence par des partenaires du projet extérieur à la FMSH 	<ul style="list-style-type: none"> Constitution de quatre groupes de travail permanents composés de chercheurs, enseignants, étudiants et « professionnels » en SHS (éditeurs, médiateurs socioprofessionnels, journalistes, ...). Réunions régulières avec ces quatre groupes de travail. Constitution d'un réseau de partenaires susceptibles d'adopter la démarche sémiotique développée dans le cadre du projet ASA-SHS. 	<p>Rapport d'avancement N°1 – N°8 (DM1)</p>	T-0 – T-48
Sous-Tâche 5.3 <i>Communication scientifique</i>	<ol style="list-style-type: none"> Faire connaître les travaux du projet ASA-SHS aux communautés scientifiques concernées 	<ul style="list-style-type: none"> Présentation des travaux du projet ASA-SHS dans des workshops nationaux et internationaux. Publication d'articles dans des revues scientifiques nationales et internationales Préparation d'une monographie en français pour l'éditeur <i>ISTE - Hermes Science Publication</i>. 	<p>Rapport d'avancement N°1 – N°8 (DM1)</p> <p>Publications scientifiques (PS)</p>	T-0 – T-48

<p>Sous-Tâche 5.4</p> <p><i>Campagnes de sensibilisation et de valorisation</i></p>	<p>1. Faire connaître les résultats du projet ASA-SHS aux personnes, institutions et médias concernés</p>	<ul style="list-style-type: none"> ▪ Mises en place de moyens de communication adéquats pour la valorisation du projet. ▪ Réalisation de ressources de communication (sites web, brochures, communiqués de presse, ...). ▪ Réalisation, chaque année d'un événement de type « journée porte ouverte ». 	<p>Supports de communication du projet (SC)</p>	<p>T-12 – T-48</p>
<p>Sous-Tâche 5.5</p> <p><i>Intégration pratique dans activités AAR</i></p>	<p>1. Assurer la durabilité des résultats du projet ASA-SHS. 2. Montrer l'utilisation pratique « à grandeur réelle » des résultats du projet ASA-SHS.</p>	<ul style="list-style-type: none"> ▪ Intégration progressive des outils sémiotiques (T3) dans le work flow actuel définissant les activités de production et publication de contenus audiovisuels dans le cadre du programme des AAR. 	<p>Rapport d'avancement N°1 – N°8 (DM1)</p>	<p>T-0 – T-48</p>

1.4 Résultats escomptés et retombées attendues/*Expected results and potential impact*

1.4.1 Les résultats visés

Le projet ASA-SHS vise deux résultats concrets :

Premier résultat :

Le traitement sémiotique et la (re-)publication en ligne de trois corpus audiovisuel de travail (histoire – littérature – archéologie) d'environ 120 heures, chacun, provenant du fonds audiovisuel des AAR et, puis, d'une manière progressive de l'ensemble du fonds des AAR (d'environ 5000 heures de vidéos).

Le but pratique de cette opération est double :

- a) pourvoir le site portail des AAR d'un ensemble de services et de prestations de contenu SHS avancés (sous forme de dossiers thématiques, d'encyclopédies en ligne, d'index thématiques de segments vidéos ou encore de parcours narratifs guidés d'exploration.) permettant une meilleure prise en compte de la diversité de la demande en « contenu SHS » et une appropriation plus adaptée aux attentes des communautés des utilisateurs (i.e. des chercheurs, des enseignants, des étudiants, des « professionnels », etc.) ;
- b) en se servant d'un des rares portails audiovisuels en France diffusant déjà depuis plusieurs années des vidéos produites en SHS, montrer aux autres initiatives comparables en France et dans le monde francophone, l'intérêt, voire la nécessité d'adopter la démarche proposée par le projet ASA-SHS en vue de permettre aux communautés concernées en vrai travail d'appropriation des ressources audiovisuelles (au lieu de se contenter d'une diffusion « de bout à bout » sans grand intérêt ...)

Deuxième résultat :

Le développement, sous le sigle d'*atelier sémiotique audiovisuel*, d'un ensemble de quatre outils d'aide pour les activités de traitement de corpus audiovisuels en SHS. Voici les quatre outils :

1/ *Des ontologies* (i.e. de « vocabulaires » structurés analogiquement comparables à des thésaurus ou terminologies) nécessaires à la description, classification, indexation, adaptation, ... de ressources audiovisuelles « source » en vue de leurs réutilisation dans des contextes d'usage spécifiques et adaptés aux besoins et intérêts des acteurs concernés.

Etant donné l'hypothèse sémiotique de départ du projet ASA-SHS, cinq ontologies seront développées, testées et proposées aux utilisateurs intéressés :

1. une ontologie thématique ou référentielle (i.e. du domaine de connaissances), une ontologie rhétorique (i.e. des actes et genres de discours) ;
2. une ontologie narrative (i.e. de l'intégration syntagmatique des segments composant un document audiovisuel) ;
3. une ontologie textuelle (i.e. d'identification de types fonctionnels récurrents de segments composant un document audiovisuel) ;
4. une ontologie de l'expression audiovisuelle (i.e. des différents médias d'expression d'un contenu stricto sensu ; des méthodes et stratégies d'expression, etc.).

2/ *Une bibliothèque de modèles de description* (représentant des lieux récurrents, des configurations générales, des structures typiques, etc.) et dont les utilisateurs peuvent se servir pour traiter des ressources audiovisuelles et de les adapter à des contextes d'usages spécifiques. Parmi ces modèles certains concernent l'identification et la description des thèmes ou sujets (la topique) d'un discours scientifique, d'autres concernent l'intégration narrative de plusieurs thèmes en un tout (sous forme de parcours narratifs), d'autres encore concernent davantage les enrichissements d'un segment sous forme de commentaires, de références, de liens, de propositions de tout genre, etc.

3/ *Des modèles de publication* à l'aide desquels des corpus audiovisuels « source » pourront être (re-)publiés sous forme, par exemple, de dossiers thématiques, de dossiers pédagogiques, etc.

4/ Une méthodologie de la chaîne du « traitement » de ressources audiovisuelles sous forme d'un guide pédagogique expliquant en détail les différentes possibilités de l'appropriation « personnalisée » de corpus audiovisuels existants.

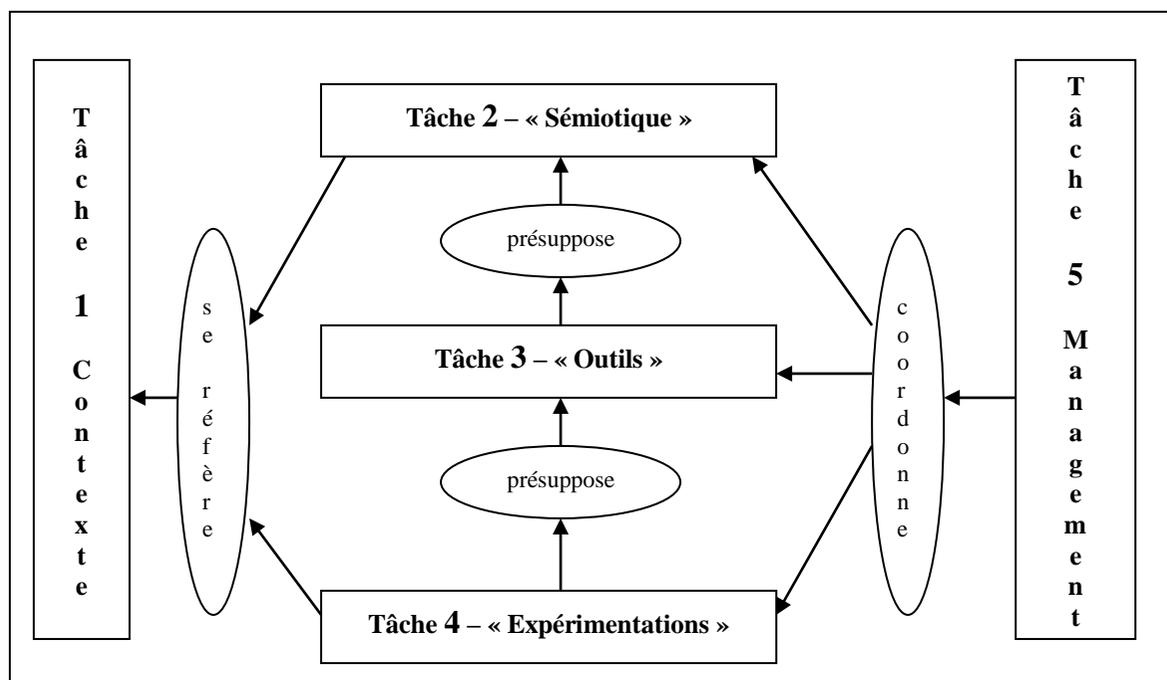
Il faut noter que l'ensemble de ces outils « intellectuels » se présenteront comme des composants intégrés dans un environnement informatique de travail déjà existant qui s'appelle Interview et qui a été développé par l'INA et la FMSH, dans le cadre d'une convention liant ces deux institutions.

1.4.2 Retombées

Les retombées scientifiques pratiques, voire politiques de ces deux résultats semblent être importants dans la mesure que sont concernées par ceux-ci toutes les institutions de recherche et d'enseignement en sciences humaines et sociales désireuses non seulement de « mettre en ligne » leurs patrimoines mais d'en faire des véritables ressources actives pour la recherche, pour l'enseignement et aussi pour les différentes pratiques professionnelles dans lesquelles les connaissances SHS sont demandées, voire nécessaires.

1.5 Organisation du projet/Project flow

1.5.1 Les dépendances fonctionnelles entre les 5 tâches du projet ASA-SHS



(figure 3 : rapports et dépendances fonctionnelles entre les cinq grandes tâches composant le projet ASA-SHS)

La figure 3 représente les différentes relations entre les cinq tâches du projet ASA-SHS. Elle montre plus particulièrement, les rapports de dépendances critiques entre les tâches 4 (qui utilisera les résultats de la troisième tâche), 3 (qui utilisera les résultats de la deuxième tâche) et 2 (qui fournira l'input théorique et intellectuel nécessaire pour les deux tâches 3 et 4). Par ailleurs, le fait qu'aussi bien la tâche 2 que la tâche 4 se réfère à la tâche 1 (définissant le contexte du projet) montre que le projet ASA-SHS possède ses « gardes fous » qui devraient l'aider à bien accomplir ses missions. Parmi ces gardes fous, on compte plus particulièrement :

- des expériences du partenaire du projet ASA-SHS assez importantes aussi bien en termes de R&D dans les domaines qui intéressent le projet qu'en terme de production – diffusion professionnelles de patrimoines scientifiques et culturels ;
- dès le départ du projet, le travail concerté avec des communautés d'utilisateurs composés de chercheurs, d'enseignants (du supérieur mais aussi du secondaire), d'étudiants (du supérieur mais aussi du secondaire) et des professionnels (formateurs et médiateurs socioculturels, éditeurs et journalistes) ;
- la constitution d'un référentiel de « benchmark » (de bonnes pratiques) dont pourra s'inspirer le projet ASA-SHS et qui lui servira en même temps comme outil d'évaluation comparative critique ;
- des enquêtes sur les attentes des communautés concernés, la définition de scénarios d'usage guidant notamment l'élaboration du pilote du projet et, enfin, la définition de procédures de tests et de validation de l'avancement des travaux du projet.

1.5.2 Echancier

La figure 4 montre le planning temporel (l'échéancier) des cinq tâches composant le projet ASA-SHS. Il est très simple et dicté par les dépendances fonctionnelles entre les 5 tâches. La tâche 5 « Management » possède obligatoirement la même durée de vie que le projet lui-même. La définition du contexte opérationnel du projet précède obligatoirement les autres tâches – elle occupe donc toute la première année. La tâche du traitement sémiotique – la tâche 2 « Sémiotique » - des corpus de travail peut intervenir à partir du sixième mois, une fois les trois corpus bien délimité et rendus accessible en ligne. La tâche 3 « Outils » dépend crucialement des résultats de la deuxième tâche en ne pourra donc commencer qu'une fois mes analyses sémiotiques terminées, c'est-à-dire à partir du 18^{ème} mois. Par contre, les premières expérimentations pourront commencer même si le développement des outils n'est pas encore complètement terminés – au contraire, elles pourront aider à gagner du temps s'il faut modifier tel ou tel caractéristiques, telle ou telle partie dans les outils.

La valorisation du projet et la dissémination des résultats de celui-ci pourront démarrer dès la première année (cf. tâche 5 « Management »). Enfin, comme le projet ASA-SHS ne connaît qu'un seul partenaire contractuel avec l'ANR (i.e. la FMSH), un effort particulier sera entrepris pour la constitution très rapide de quatre groupes de travail composés de profils d'utilisateurs spécifiques : chercheurs, enseignants, étudiants et « professionnels » (i.e. éditeurs, journalistes et médiateurs socioculturels). Cette activité fait également partie de la tâche 1 « Management ».

Tâche/Tasks	Partenaire	Année 1 Year 1		Année 2 Year 2		Année 3 Year 3		Année 4 Year 4	
		6	12	18	24	30	36	42	48
1. Tâche 1 « Contexte », FMSH	FMSH								
2. Tâche 2 « Sémiotique », FMSH	FMSH								
3. Tâche 3 « Outils », FMSH	FMSH								
4. Tâche 4 « Expérimentations », FMSH	FMSH								
5. Tâche 5 « Management », FMSH	FMSH								
Rapports d'avancement semestriel Progress report/expenses	FMSH	☺	☺	☺	☺	☺	☺	☺	☺

	Dossier Scientifique « Sémiotique » (DS1)	18	FMSH
	Dossier Scientifique « Sémiotique » (DS2)	18	FMSH
3. « Outils »			
	Vocabulaires d'ontologies « ASA-SHS » sous forme de fichiers xml (OR2)	30	FMSH
	Schémas ou graphes de description « ASA-SHS » sous forme de fichiers xml (OR3)	30	FMSH
	« Templates » de publication (fichiers xml/xslt représentant des structures de publication adaptables) (OR4)	30	FMSH
	Mise à jour de l'environnement de travail de traitement de corpus audiovisuels appelé Interview (OR5)		
4. « Expérimentations »			
	Rapport technique « Expérimentations » (RT2)	48	FMSH
	Pilote du projet (OR6)	48	FMSH
5. « Management »			
	Rapport d'avancement N°1 – N°8 (DM1)	6, 12, 18, 24, 30, 36, 42, 48	FMSH
	Rapport de synthèse et récapitulatif des dépenses (DM2)	48	FMSH
	Publications scientifiques (PS)	A partir de 12	FMSH
	Supports de communication du projet (SC)	A partir de 12	FMSH

(figure 5 : liste de livrables)

1.6 Organisation du partenariat/*Consortium organisation*

1.6.1 Pertinence des partenaires/*Consortium relevance*

Le projet ASA-SHS est proposé par un seul partenaire, la Fondation de la Maison des Sciences de l'Homme (FMSH) qui, cependant, s'efforcera, dès le début du projet, à mettre en place un réseau de partenaires ainsi que quatre groupes de travail (composés de chercheurs, d'enseignants, d'étudiants et de « professionnels ») afin de permettre un processus de validation de la démarche elle-même, des outils et des expérimentations menées dans le cadre du projet.

Enfin, grâce à sa position très singulière dans le paysage des établissements de recherche SHS en France, la FMSH pourra s'appuyer sur un réseau très étendu de partenaires académiques et institutionnels pour diffuser aussi largement que possible les résultats du projet ASA-SHS et pour faire coopter par un nombre significatif d'institutions de recherche et d'enseignement en France et dans le monde francophone la démarche développée dans ledit projet.

La Fondation Maison des Sciences de l'Homme (FMSH) est un établissement d'utilité publique, personne morale de droit privé, ayant pour but, selon le décret du 4 janvier 1963 modifié : « la promotion de l'étude des sociétés humaines considérées avant tout dans leurs réalités actuelles et à partir d'elles ».

La Fondation Maison des Sciences de l'Homme exerce ses activités sous la tutelle du ministère chargé de l'enseignement supérieur et de la recherche. L'objectif de ses fondateurs était de créer un lieu où les différentes institutions pourraient installer côte à côte leurs centres de recherche en sciences humaines et sociales, et où chacun conserverait son indépendance, condition de toute créativité intellectuelle. Elle est destinée à servir de lieu d'expérimentation et de mise au point pour des idées, des théories nouvelles et des formes d'organisation, afin d'accompagner les progrès de la recherche avancée et les transformations des sociétés. Elle vise à faciliter le dialogue entre chercheurs spécialistes des différentes disciplines et des différentes aires culturelles pour une coopération intellectuelle et une réflexion en commun.

Chargée d'exercer auprès de la recherche française en sciences sociales une mission de soutien et d'accompagnement, la FMSH constitue une plate-forme de collaboration scientifique nationale et internationale entre institutions, disciplines et chercheurs. Son action, qui complète celle des grands établissements d'enseignement supérieur et de recherche, s'exerce :

1. Par le regroupement, à l'intérieur d'un même immeuble, autour d'instruments de travail communs de formations de recherche et de documentations relevant d'institutions différentes ;
2. Par l'organisation de réseaux nationaux et internationaux de coopération scientifique ;
3. Par la mise en place d'équipes expérimentales susceptibles de déterminer des orientations et des méthodes nouvelles de recherche dans les sciences sociales ;
4. Par la collecte, la publication et la diffusion des informations scientifiques.

Ainsi, la FMSH s'appuie sur un vaste réseau de partenariats avec des institutions prestigieuses tant nationales qu'internationales au service de la production et de la diffusion de la connaissance, fruit du travail des chercheurs en sciences humaines et sociales du monde entier.

En 2002 a été mis en place, à la Fondation Maison des Sciences de l'Homme, le programme des Archives Audiovisuelles de la Recherche (AAR). Les cinq missions du Programme AAR sont les suivants :

1. Collecte et constitution de patrimoines scientifiques et culturels numériques à l'aide notamment des technologies audiovisuelles (tournage d'événements scientifiques et culturels, numérisation de documents audiovisuels, postproduction et publication en ligne).
2. Ingénierie sémiotique et informatique des fonds audiovisuels et multimédias du programme des AAR : création et gestion d'ontologies de description ; description, indexation et classification de corpus spécifique ; annotation, adaptation et traduction de corpus spécifiques.
3. Publication et/ou republication d'enregistrements audiovisuels de patrimoines culturels et scientifiques sous forme de sites thématiques multimédias, de genres de publication spécifiques (vidéo-livres interactifs, dossiers thématiques à usage éducatif, vidéo-lexiques, encyclopédies hypermédias évolutifs, parcours narratifs ré-linéarisés, ...) ou encore pour un public multilingue et multiculturel.
4. Conception et développement de ressources scientifiques et d'outils techniques appropriés d'une part pour la production et la gestion d'archives multimédias, et d'autre part pour la description, indexation et (re-)publication en ligne de corpus audiovisuels.
5. Mise en place de partenariats avec des organismes publics et privés pour la valorisation et la communication en ligne de connaissances scientifiques et culturelles à l'aide des technologies de l'audiovisuel numérique.

D'un coût annuel direct d'environ 520 000 euros, le programme des AAR est financé a) en partie sur le budget propre de la Fondation, b) depuis 2006 par le MENESR via un PPF dans le cadre du contrat quadriennal entre la Fondation et son Ministère tutelle, ainsi que c) d'une manière très importante via des projets de R&D français (notamment dans le cadre du programme ANR-RIAM) et Européens (notamment dans le cadre du 6^{ème} PCRD).

12 personnes travaillent d'une manière directe et exclusive pour le programme AAR ; plusieurs services de la FMSH sont plus ou moins directement impliqués dans l'évolution du programme des AAR : outre le service de R&D ESCoM, il s'agit notamment du service informatique, des éditions de la FMSH, de la bibliothèque, des différents programmes de bourses et

d'accueil de chercheurs, des programmes de coordination des colloques, de certains programmes de recherche « fournissant » un contenu de qualité et encore du tout nouveau Institut d'Études Avancées.

Pour plus d'informations :

- Le site officiel de la FMSH : <http://www.msh-paris.fr/>
- Le site du portail AAR donnant l'accès à, actuellement, environ 5000 heures de vidéos dans toutes les disciplines en SHS de presque 1300 chercheurs d'une trentaine de pays du monde : <http://www.archivesaudiovisuelles.fr/FR/>

1.6.2 Complémentarité et synergie des partenaires/Added value of the consortium

Montrer la complémentarité et la valeur ajoutée des coopérations entre les différents partenaires. L'interdisciplinarité et l'ouverture à diverses collaborations seront à justifier en accord avec les orientations du projet.

Le projet ASA-SHS est proposé par un seul partenaire, la Fondation de la Maison des Sciences de l'Homme (FMSH) qui, cependant, s'efforcera, dès le début du projet, à mettre en place un réseau de partenaires ainsi que quatre groupes de travail (composés de chercheurs, d'enseignants, d'étudiants et de « professionnels ») afin de permettre un processus de validation de la démarche elle-même, des outils et des expérimentations menées dans le cadre du projet.

1.6.3 Qualification du coordinateur du projet et des partenaires/Principal investigator and partners : résumé and CV

CV de Peter Stockinger

1/ Profession :

- Depuis octobre 1990: professeur titulaire des universités en sciences du langage (spécialisation : sémantique cognitive ; sémiotique du discours, NTIC) à l'Institut National des Langues et Civilisations Orientales (I.Na.L.C.O.) à Paris
- Travail depuis 1991 régulièrement pour la FMSH où il est, plus particulièrement, le responsable scientifique du programme des AAR

2/ Formation :

Sciences du langage (sémiotique et analyse du discours), linguistique et psychologie cognitive

- 1987 : Thèse d'état es lettres et sciences humaines à l'Université de Paris III - La Sorbonne Nouvelle. Mention: très honorable avec les félicitations à l'unanimité du jury.
- 1984 : Thèse d'état es lettres et sciences humaines à l'Université de Paris III - La Sorbonne Nouvelle. Mention: très honorable avec les félicitations à l'unanimité du jury.
- 1980 : Ph.D. (Doktor der Philosophie) à l'université de Salzburg (Autriche). Mention: très honorable avec les félicitations à l'unanimité du jury.

3/ Distinction :

- 1985 : Médaille de Bronze du CNRS pour les recherches en sémiotique et linguistique du discours.

4/ Responsabilités de R & D :

- Depuis 1994, responsable scientifique d'une quinzaine de projets de R&DT français et Européens à la FMSH.

2006 – 2009 :

1. Projet Européen LOGOS (6^{ème} PCRD) coordonné par Antenna Hungarica et réunissant un consortium de 15 partenaires de la France, du Royaume Uni, de l'Italie, de la Grèce, de l'Hongrie, de la Finlande, de la Bulgarie et de la Slovaquie. Ce projet vise la diffusion de corpus audiovisuels numériques pour tous supports de visualisation (PC, TV, téléphone mobile, psp, ipod, ...).
2. Projet Européen DIVAS (6^{ème} PCRD) coordonné par Archetypon, société grecque et réunissant un consortium de partenaires de la France, d'Allemagne, d'Autriche, de la Grèce, d'Israël et de Russie. Ce projet vise le développement d'un moteur de recherche de contenus audiovisuels.
3. Projet français de R&D SAPHIR (programme RIAM) coordonné par l'INA. Ce projet vise la republication de corpus audiovisuels selon des genres de publication différents : films virtuels, dossiers pédagogiques, dossiers bilingues, etc.

2004 – 2006 :

1. Projet Européen CHIRON (programme Leonardo da Vinci) coordonné par la FMSH dans le domaine de l'u-learning (« ubiquitous learning »).
2. Projet Européen JASON (programme Leonardo da Vinci) coordonné par l'Université de Coimbra, Portugal) visant la conception et la publication d'un centre de ressources numériques pour l'enseignement dans la production, gestion et exploitation d'archives multimédias.
3. Projet bilatéral franco-argentin entre la FMSH et l'Université de Cuyo à Mendoza visant la mise en place d'un portail multimédia commun d'accès et de diffusion de connaissances scientifiques produites (sous forme de documents audiovisuels, de revues, de données structurées, ...) par les deux institutions partenaire.

5/ Domaines de Recherche

- 1) Sémiotique et (nouveaux) médias
 - 1.1) Approche sémiotique et discursive de l'information et exploitation technologique ;
 - 1.2) Sémiotique du document audiovisuel numérique ;
 - 1.3) Description comparative de l'information développée et mise en scène dans les sites web
- 2) Sémiotique, Culture et Communication
 - 2.1) Théorie sémio-linguistique du discours ;
 - 2.2) Etudes sémiotique des médias ;
 - 2.3) Sémiotique des cultures et communication interculturelle : cultures, langues et langages.

6/ Publications (seulement ouvrages – depuis 1999) :

1. 2005 *Les sites web*. Description, évaluation et production. Paris, Hermes Science Publications (220 p).
2. 2003a) *Le document audiovisuel. Description et exploitations pratiques*. Paris, Hermes Science Publications (240 p).
3. 2003b) (eds.) *Portails et collaboratoires pour l'enseignement et la recherche*. Paris, Hermes Science Publications (230 p).
4. 2001 *Traitement et contrôle de l'information*. Perspectives sémiotiques et textuelles. Paris, Hermes Science Publications (195 p).
5. 2000 (avec Charlotte Nikitenko; eds.) *La publication en ligne*. Paris, Hermes Science Publications (225 p).
6. 1999 *Les nouveaux produits d'information*. Conception et sémiotique du document. . Paris, Hermes Science Publications (370 p).

CV de Elisabeth de Pablo

1/ Profession :

- Depuis septembre 1995: ingénieur – chercheur à la FMSH spécialisée dans les nouveaux médias, l'infographie et le traitement (production, description, indexation, publication, ...) de ressources multimédias et, plus spécifiquement de ressources audiovisuelles

2/ Formation :

Etudes de langues, sciences du langage et ingénierie multilingue à l'INaLCO (Institut National des Langues et Civilisations Orientales, diplômée d'un DEA et d'un DESS

3/ Expériences de R & D :

- Depuis 1995, travail dans plusieurs projets de R&D Européens et français portant sur le déploiement de services et prestations de contenus numériques pour la recherche et l'éducation (exemples : projet ESPRIT IV « Virtual Hypermedia Factory », de 1996 à 1999, portant sur la mise en place de services de publication web pour des centres de recherche et des éditeurs spécialisés en SHS ; projet Leonardo ARCHIMED, de 1999 à 2002, portant sur la spécification d'un « campus numérique » pour l'enseignement à distance ; projet Leonardo JASON, de 2004 à 2006, portant sur la création d'une formation à distance dans la gestion d'archives audiovisuelles ; projet 6^{ème} PCRD LOGOS portant sur la mise en place d'un pilote audiovisuel en ligne en anthropologie et études de civilisations, ...).

5/ Domaines de Recherche

- 1) Production et traitement de corpus audiovisuels
 - 1.1) Description, annotation et enrichissement de corpus audiovisuels ;
 - 1.2) Traduction et adaptation de corpus audiovisuels ;
 - 1.3) Publications et éditions de corpus audiovisuels traités pour le web et tout support numérique.
- 2) Infographie et multimédias
 - 2.1) Sémiotique des interfaces multimodales ;
 - 2.2) Développement de sites web.

