

La transformation numérique des documents



Bruno Bachimont Université de technologie de Compiègne, France

- Le document, c'est technique :
 - La grammatisation
- Le document, c'est la mémoire de l'événement :
 - La grammatisation de la mémoire

Prologue 1 : grammatisation !



Exergue

- Ce sont également des outils linguistiques externes aux sujets parlants qui modifient les espaces de communication — standardisation et stabilisation. Les premiers grammairiens, confrontés à des espaces d'oralité et par conséquent de variations, traduisaient bien souvent ce fait en se vantant d'avoir “réduit à des règles” la langue dont ils s'occupaient. »
 - Auroux, Grammatisation
- La formalisation machinique des gestes du travailleur résulte d'une analyse puis d'une synthèse — réalisée comme artefact par la technoscience. Notons au passage qu'une telle formalisation est une sorte de grammatisation [...] comme analyse en tant que discrétisation du continu. »
 - Stiegler, Individuation et grammatisation.
- Le dernier stade de la grammatisation est le nôtre : son stade numérique, qui est aussi celui de la société hyperindustrielle où l'extériorisation des fonctions de lecture et de computation semble dissociée de l'intériorisation qui accompagnait autrefois calcul et lecture. [...] l'informatique ne fait pas disparaître la lecture et l'écriture. C'est au contraire une archi-lecture qui change les conditions de la lecture et de l'écriture. »
 - Ars industrialis.

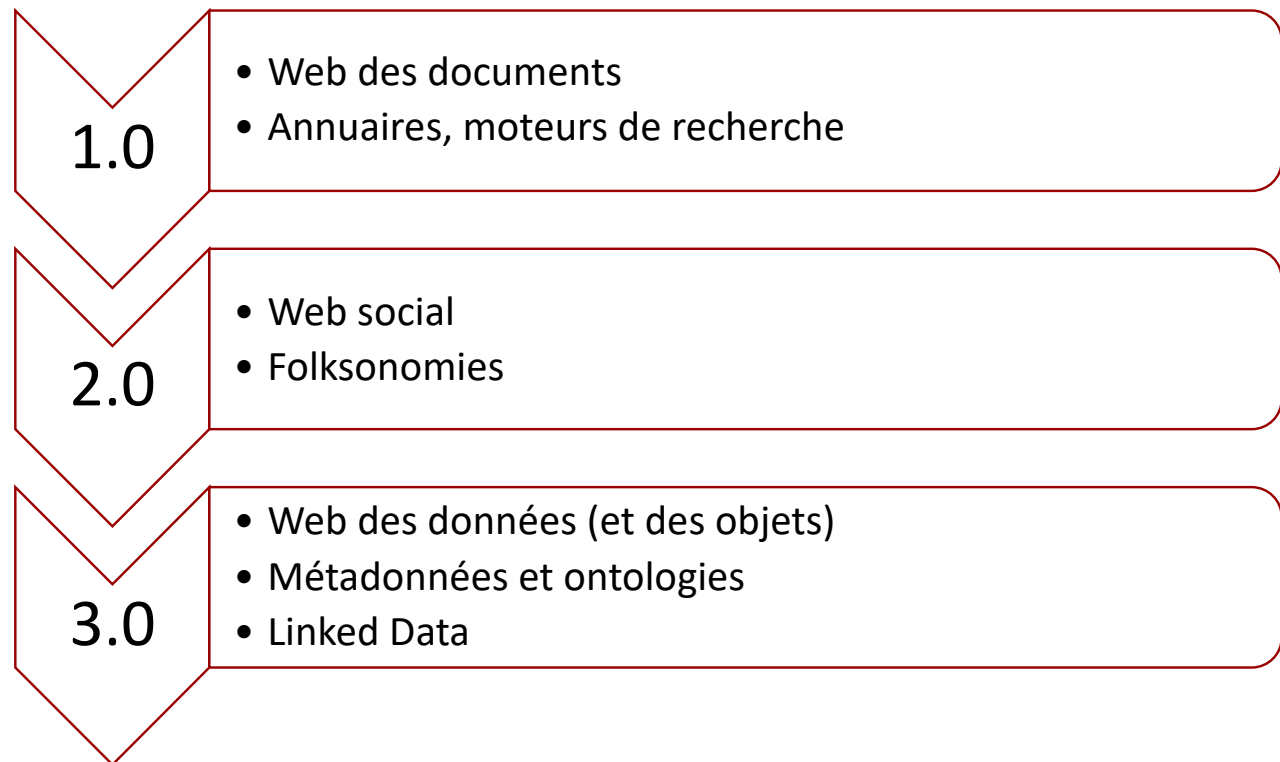
L'idée

- Enregistrer et écrire, ça transforme et ça normalise ce qu'on écrit, ce n'est pas une simple description ou fixation ;
- Or, écrire, c'est discrétiser.
- Et, discrétiser, c'est numériser.
- Ben alors : si on numérise tout, comme on le fait actuellement,
 - Quelles transformations sont impliquées ?
 - Quelles conséquences ?

Quelques repères

- La technique de l'écriture et la constitution de la grammaire
 - cf. Sylvain Auroux : la révolution technologique de la grammatisation
- La technique grammaticale (latine) et la linguistique
 - Ibidem
- L'imprimerie et la circulation du savoir
 - Cf. Elizabeth Eisenstein : la révolution de l'imprimé dans l'Europe des premiers temps modernes
- L'IA et la documentation ?
 - C'est aujourd'hui !

Plusieurs Web



Des contenus aux données

➤ Du document indexé...

- Recherche documentaire
 - Paradigme: bibliothèque
 - Outils : html + http

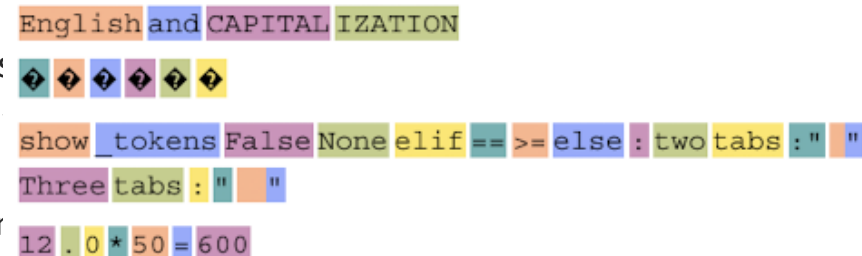
➤ À la ressource annotée...

- Recherche d'information, agrégation et publication
 - Paradigme : web sémantique
 - Outils : XML + RDF + Ontologies
 - Données : format + formel ; la donnée affirme un fait sur le monde (e.g. Auteur = Balzac).



➤ À la donnée manipulée

- Données collectées : manipulées leurs corrélations:
- Paradigme : données (les IA statistiques/générati
- Outils : ML + DL + IA g
- Données : format (sans formel : la donnée n'affirr qu'elle-même, rien sur sur le réel dont elle est la donnée).



Un ch'tiot exemple : le Web et son archivage

Dépôt Légal du Web

Fichier Marques-pages Connection Affichage Préférences

Date de consultation : 2005-11-26 19:50:49

Versions disponibles

- Version actuellement en ligne
- Dernière version enregistrée
- Versions antérieures
 - le 14-10-2005 à 20:45:24
 - le 24-11-2005 à 19:03:33
 - le 25-11-2005 à 16:54:13
 - le 26-11-2005 à 19:50:49
 - le 27-11-2005 à 12:24:24
 - le 28-11-2005 à 11:37:52
 - le 12-12-2005 à 02:56:52

France 5 - Chaîne de télévision sur la découverte, l'emploi, la santé : documentaires, émissions, reportages, programmes

Précédent Suivant Recharger Stop http://www.france5.fr/ Go

france 5.fr

France 5 est une chaîne du groupe francetélévisions

Mercredi 14 décembre

ACCUEIL PROGRAMME LES SITES FRANCE 5 EMISSIONS THEMATIQUES LA CHAÎNE

SAMEDI 26
19:00 DOSSIER SCHEFFER
AVEC LE COMMISSAIRE DIVISIONNAIRE
FRÉDÉRIC PÉCHENARD

Pour en savoir plus >>

RECHERCHE

☐ Dans France 5

☐ Dans la grille des programmes

OK

TOUT FRANCE 5

Actu société

Arts et culture

Découverte nature

Emploi éco

Famille jeunesse

Histoire

Santé

Sciences

Vie pratique

PARTICIPEZ
Réactions, forums, contributions, témoignages...

FRANCE 5 EDUCATION
Les nouveaux services
Educatifs Multimédia de France 5

FRANCE 5 EMPLOI
BEN VIVRE LE MONDE
DU TRAVAIL

LES ZOULOUS

LES RENDEZ-VOUS DU JOUR

09:10 "L'oeil et la main" : De la discrimination à la création d'entreprise.

09:40 "Cas d'école" : Ados et sexualité.

10:35 "L'atelier de la mode" : La tendance vintage.

11:10 "Question maison" : L'art nouveau, le viager...

12:00 "Silence, ça pousse !" : Le site de Siguyria, au Sri Lanka.

19:55 "Le journal du blogue".

20:05 "Ubik" : Marie Gillain, Tommy Lee Jones, Clarika...

>> Bandes annonces | Tout le programme

09:40 - "Cas d'école"
La sexualité des ados. Reportages en ligne en vidéo.
[En savoir plus](#)

19:55 - "Le journal du blogue"
La parole est aux blogueurs ! Retrouvez la vidéo de l'émission en ligne.
[En savoir plus](#)

LES EMISSIONS

A vous de voir

Arrêt sur images

Avis de sorties

Brigade nature

C dans l'air

Le journal du blogue

Les amphip de France 5

Les maternelles

Les Zouzous

L'odyssée de l'espèce

bandes annonces
sur France Télévisions

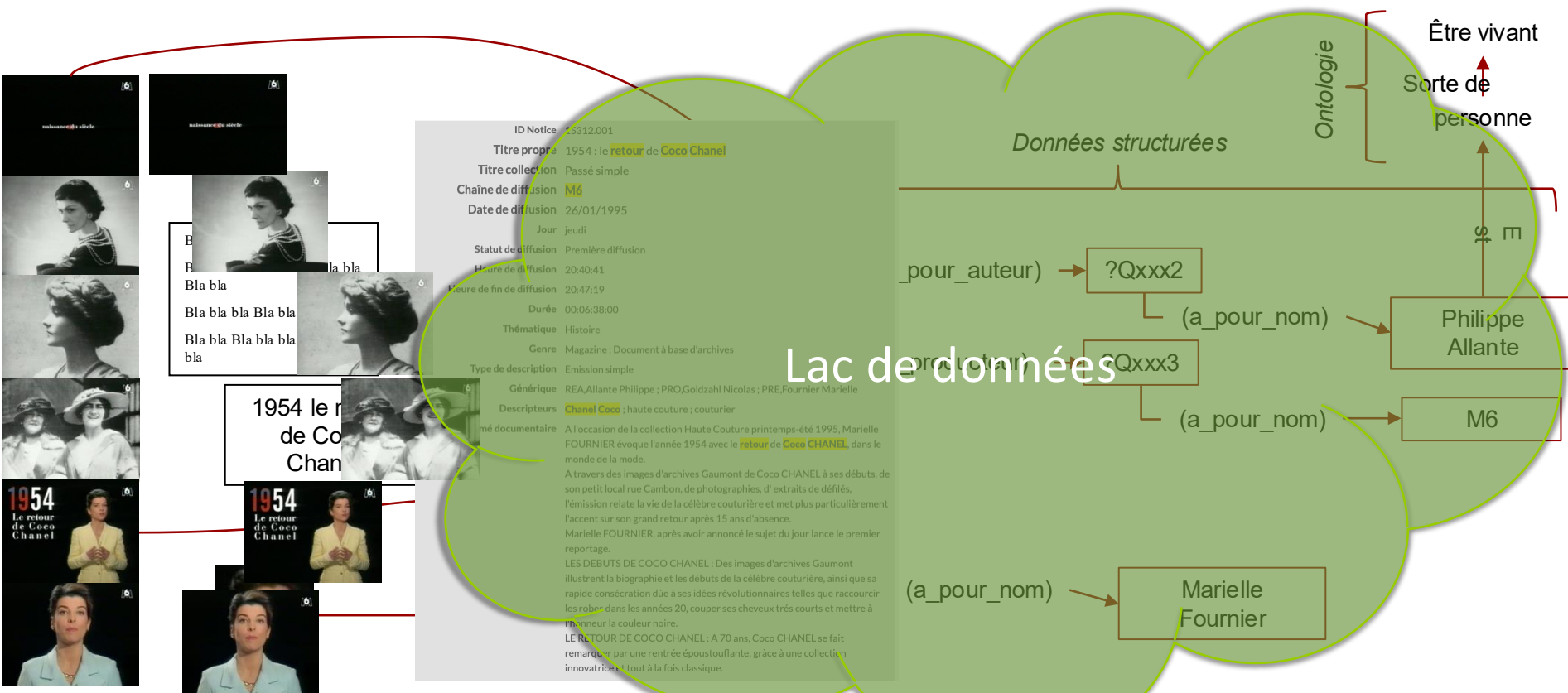
NEWSLETTER
Je m'inscris ou modifie mes abonnements
Mon email ici OK
>> Voir la lettre d'infos de la semaine

Navigateur Plan de travail Editeur Texte Editeur HTML Cartographie

Done

none console

Un autre ch'tiot exemple : la notice documentaire



**IA statistique
(extraction /
reconnaissance)**

**IA Symbolique
(inférence / raisonnement)**

Jacques Derrida

- Parle plutôt de « grammatologie » : le sens ne se dévoile qu'à partir d'une écriture préalable qui l'informe et le déforme par ses propres structures :
 - On n'accède jamais directement à ce qu'on veut dire, mais à travers un dit qui déforme ce qu'on veut dire tout en étant le seul moyen de le dire.
 - C'est une critique de la parole « vivante » comme accès direct au sens : c'est toujours déjà une écriture impliquant espacement (écart entre le dit et le vouloir-dire) et déplacement (déformation).
 - L'expression est toujours le dépôt, la trace d'une origine dont l'expression est la manifestation, qu'elle déforme mais transmet, qu'elle transmet mais déforme.
- « Nous voudrions plutôt suggérer que la prétendue dérivation de l'écriture, si réelle et si massive qu'elle soit, n'a été possible qu'à une condition : que le langage « originel », « naturel », etc., n'ait jamais existé, qu'il n'ait jamais été intact, intouché par l'écriture, qu'il ait toujours été lui-même une écriture. » *De la Grammatologie*, Minuit, 1967. p. 82.

Sylvain Auroux

- L'acception est ici plus technique : l'écriture est l'instrument qui permet d'objectiver la langue et d'en tirer des outils permettant sa description, normalisation et transformation :
 - Grammaire, dictionnaire, rhétorique.
- La grammatisation est donc ce processus d'outillage de langue permettant de la transformer, normer et transmettre.
- « G[rammatisation] désigne le processus par lequel une langue se trouve “outillée”, notamment à l'aide de grammaires et de dictionnaires »,
 - « Grammatisation », *Archives et documents de la Société d'histoire et d'épistémologie des sciences du langage*, 1995, p. 5-6.

Bernard Stiegler

- Il reprend et généralise l'acception de grammatisation par Auroux, ce qui lui permet de la mettre en perspective avec celle de Derrida.
- La grammatisation ne concerne plus seulement les langues ; elle concerne tous les flux temporels humains susceptibles d'être décomposés en unités discrètes : parole, gestes, images, comportements, affects, savoir-faire, données numériques.
- « La grammatisation désigne la transformation d'un continu temporel en un discret spatial »
 - Ars industrialis, vocabulaire.

Pour nous...

- L'écriture est un système d'inscription et de manipulation d'un code de représentation
 - Écriture de la parole : la lettre et la communication
 - Écriture de l'échange : la monnaie et la transaction
 - Écriture du calcul : le nombre et la quantification.
 - Cf Clarisse Herrenschmidt, les trois écritures, Gallimard, 2007.
- La grammatisation est le processus selon lequel une écriture transforme ce dont elle est écriture et fait retour sur elle :
 - L'écriture grammatisé la parole, et on ne parle plus de la même manière quand on sait écrire ;
 - Le numérique grammatisé l'audiovisuel, et on ne conçoit, crée, regarde plus ces contenus de la même manière ;
 - La captation du geste permet d'en faire un enregistrement, sur la base duquel on peut le décomposer, le repenser, le recréer.
 - Etc.

Prologue 2 : mémoire !



Exergue

- « Là est le secret, le principe, la force, l'économie, l'hygiène du génie; cette singulière force de jeunesse: manquer de mémoire, n'avoir pas de mémoire, ne pas traîner derrière soi cette lourde masse, ce train de chemin de fer, ce train de marchandises... »
- PEGUY, Charles, Véronique: dialogue de l'histoire et de l'âme charnelle, Paris, Gallimard, 1972, p.71.

Une expérience quotidienne

- Considérons nos fichiers d'il y a 20 ans...
 - Nous avons relu et retravaillé régulièrement certains d'eux :
 - Le fait de les relire nous a permis de les mettre à jour :
 - Mise à jour technique : migration logicielle
 - Mise à jour intellectuelle : adaptation du contenu
 - Certains n'ont pas été réutilisés :
 - Si jamais on veut les relire :
 - On n'a plus de moyens logiciels pour les lire ;
 - On ne comprend plus le contenu;
- Conclusion
 - C'est en lisant qu'on préserve ;
 - L'utilisation est une condition de l'archivage, non sa conséquence.
 - La consultation comme préalable à la conservation.

Pourtant, un constat global contraire

- Inflation mémorielle
 - Une extension des obligations du stockage et de la conservation:
 - Dépôt légal : exemple en France
 - extension du périmètre Radio et télévision aux chaînes privées, câbles, satellites, réseaux
 - extension du dépôt légal au Web
 - Industrie : conserver les documents de conception pendant 70 ans pour être certifié.
 - Massification des contenus
 - Plus de 1000 000 h/an à l'Ina par exemple
 - Fixation de l'éphémère
 - Conversations, messages, etc., deviennent enregistrables, donc pérennisables.
 - N.B.: l'enregistrement devient un préalable à la communication (IP) alors que c'était l'inverse avant.
 - Cf. internet : l'enregistrement précède la communication.
- Conclusion
 - On conserve et stocke de plus en plus pour l'avenir, mais sans consulter au présent.
 - La conservation comme préalable à la consultation

Deux modèles en tension

➤ Modèle statique

➤ On a des souvenirs :

➤ Traces objectives (documents, vestiges, indices): permet l'histoire ;

➤ Traces subjectives (mémoire, souvenir) : permet la mémoire.

➤ Si on garde ces souvenirs intacts, on se souvient de manière fidèle.

➤ L'enjeu est la préservation à l'identique: **intégrité et authenticité**.

on se souvient parce qu'on a des souvenirs

➤ Modèle dynamique

➤ les souvenirs ne sont jamais stables:

➤ Corruption physique naturelle

➤ Si on ne fait rien, ils deviennent inaccessibles

➤ En se souvenant, on les réactive : mise à jour technique, actualisation conceptuelle; l'enjeu est l'action de se souvenir.

➤ L'enjeu est de maintenir **vivant** le contenu pour en faire un souvenir

➤ **On a des souvenirs parce qu'on se souvient**

Une tension créatrice

- La faculté de mémoire est donc partagée entre deux contraintes :
 - Conserver et stocker pour permettre l'exercice de la mémoire
 - C'est la conservation du patrimoine et de ses objets
 - Utiliser et entretenir la culture et la pratique de ces objets
 - C'est la transmission de la tradition et la réinvention permanente de la mémoire et de ce dont il y a mémoire.
- Le numérique:
 - Accroît cette tension
 - Permet de la repenser et de mieux comprendre les conditions de possibilité de la mémoire et des outils, méthodes et politiques à mettre en œuvre.

Mais de quels souvenirs parle-t-on ?

- Les contenus conservés sont au croisement de 4 préoccupations distinctes mais pas incompatibles:
 - L'archive comme **valeur probatoire / source de preuve** :
 - Les archives administratives, les archives nationales.
 - L'archive comme **valeur culturelle / collection d'œuvres** :
 - Les bibliothèques, les dépôts légaux, le génie propre d'une nation ou d'une culture
 - L'archive comme **valeur informationnelle / ressource d'information** :
 - La documentation pour la recherche d'information
 - L'archive comme **valeur mémorielle / référence identitaire** :
 - Le patrimoine : tout objet susceptible de renvoyer à un passé qui serait le mien, qui me constitue, me définit, et dans lequel je me reconnais ou identifie.

Différents rapports au passé

Preuve du passé **Archives**

Sumer



Témoignage de l'esprit **Bibliothèques**



Edit de Montpellier: 1537

Fréteval: 1194



Edit de Villers-Cotterêts: 1539



Source d'information **Documentation**

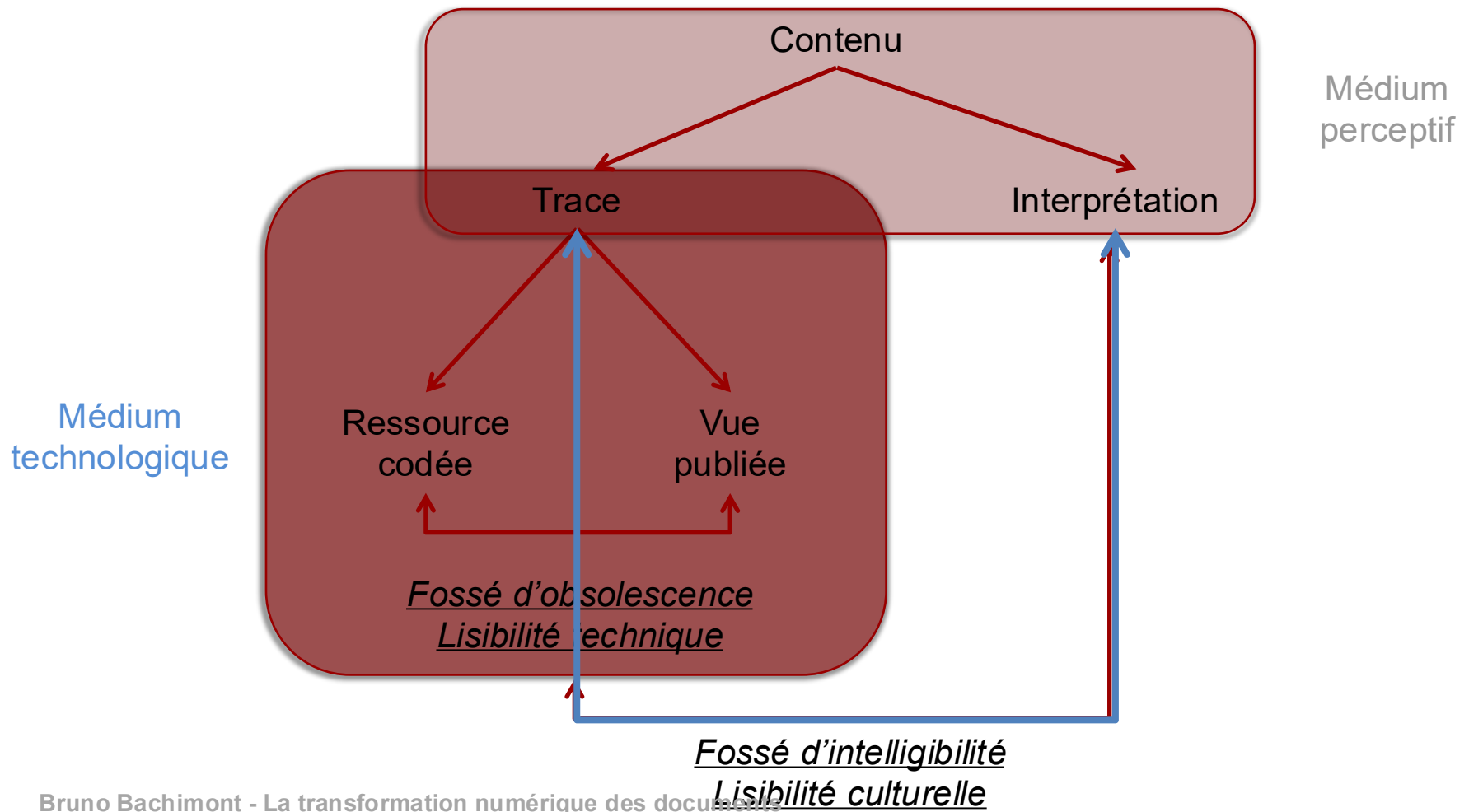
Suzanne Briet



Support de mémoire **Patrimoine**



Support et contenu



Enjeux des arts de l'enregistrement

- Massification des contenus :
 - Opportunité :
 - émergence des « humanités numériques », une compréhension nouvelle grâce à l'analyse globale des contenus numérisés et massifiés
 - Production et traitement par l'IA (ML, DL, IAg).
 - Menace : infobésité, montée de l'insignifiance, perte de la relation au contenu et au réel.
 - Tension : savoir structurer, trier, explorer, analyser.
- Complexités
 - Complexité physique : corruption des supports toujours plus fragiles.
 - Même si le numérique est autonome vis-à-vis du support, il n'en est pas indépendant.
 - Complexité technologique:
 - Obsolescence accélérée des outils, formats et normes;
 - Sophistication des outils de lecture et d'accès;
 - Complexité culturelle:
 - Une culture de l'information à re-configurer.
- Tension
 - Toute technique d'enregistrement entraîne par nature ces enjeux, le numérique étant la première technique universelle car portant sur l'information

Prologue 1 + Prologue 2

- Que fait la grammatisation à la mémoire ?
 - Derrida : la mémoire n'existe que par sa transmission mais qui la déforme ; elle n'est pas une parole qui se transmet, mais une écriture qui la déplace et la réinvente ;
 - Auroux : la transmission de la mémoire repose sur des outils d'objectivation qui la transforme de ce fait ;
 - Stiegler : les techniques grammatisent tous les procédés et contenus et les transforment de cette manière. La technique s'uniformise à travers la numérisation globale que l'on connaît.

Objectifs et plan du cours

Archivistique audiovisuelle et numérique



Mais bon, c'est bien gentil tout ça, mais pourquoi sommes-nous là ?

- Essentiellement :
 - Pour avoir des crédits (3 en fait)...
- Mais aussi :
 - Pour penser cette grammatisation comme processus en cours de notre rapport aux contenus et à la transmission (mémorielle, patrimoniale, documentaire) qu'ils permettent :
 - Modification de notre rapport à la mémoire ;
 - Compréhension de ce que change le numérique ;
 - Dégager le statut des outils actuels et les enjeux sur nos pratiques.

Objectifs

➤ Savoir:

- Maîtriser un cadre conceptuel pour problématiser les archives audiovisuelles et numériques.
- Avoir des repères normatifs et méthodologiques.

➤ Savoir-faire:

- Savoir dégager l'enjeu et la difficulté archivistique d'un domaine donné de contenus:
 - Problématiser, pour cribler les outils et méthodes des autres cours.

➤ Savoir-être:

- Se comporter en bricoleur théorique et en praticien rigoureux :
 - Tout est à inventer : il faut penser expérimental !
 - Les enjeux sont importants : bricoler avec rigueur !
 - Cf. Michel de Certeau : bricolage et braconnage !

Structure du cours

Document et archive

- Le document
 - L'information
 - Les dimensions documentaires
- L'archive
 - L'archivistique générale
 - Les problèmes philosophiques

Support et nouveaux enjeux

- Le support:
 - Théorie du support
 - Le numérique
- Grammatisations numériques
 - Données, traces, enregistrement
 - IA symbolique et neuronale
 - IA générative
 - Archivage du Web

Plan du cours

Séances	Horaires	Sujet
13 juillet 2026 matin	9h-12h	Introduction / Document
13 juillet 2026 après-midi	13h-16h	Document (suite)
14 juillet 2026 matin	9h-12h	Archives : principes et disciplines voisines
14 juillet 2026 après-midi	13h-16h	Philosophies de l'archive
15 juillet 2026 matin	9h-12h	Méthodologie
16 juillet 2026 matin	9h-12h	Théorie du support et support numérique
16 juillet 2026 après-midi	13h-16h	Les tendances du numériques
17 juillet 2026 matin	9h-12h	Données et grammatisation de l'information
17 juillet 2026 après-midi	13h-16h	Introduction à l'IA : IA symbolique
20 juillet 2026 matin	9h-12h	Introduction à l'IA : IA neuronale
21 juillet 2026 matin	9h-12h	Introduction à l'IA : IA générative
23 juillet 2026 matin	9h-12h	Archivage du Web : généralités et le cas de l'Ina
24 juillet 2026 matin	9h-12h	Présentations orales

Évaluation

- Une étude sur l'archivage d'un type de contenus particulier:
 - Travail collectif par binôme, sur une grille d'analyse fournie, mais avec corpus / terrain / objet à définir, trouver et analyser.
 - Soutenance + remise du support de présentation :
24 juillet 2026
- Sujet de réflexion individuelle
 - Travail rédigé de 3000 mots environ.
Remise le 31 août 2026
- Par voie électronique uniquement :
 - bruno.bachimont@utc.fr

Principes de l'étude

- Posture :
 - Pré-étude d'un secteur pour déterminer les enjeux et les opportunités d'envisager la transformation numérique du secteur et ses modalités ;
 - Motiver un avis auprès d'un décideur dans l'attitude d'une AMO (assistance à maîtrise d'ouvrage).
- Approche :
 - Choisir un domaine défini en termes de « contenus », culturels ou professionnels.
 - Articuler un cadre général (e.g. BD) et un cas concret pour fixer les idées (Webtoons).
 - Mener la recherche par les outils du Web. Rencontrer les porteurs du cas concret, mais pas nécessité en soi.
- Résultat
 - Présentation orale et remise du support de présentation (ppt).

Sujets d'étude

Des types de contenus

- Archivage du web ;
- Archivage des arts médiatiques ;
 - La musique contemporaine : IRCAM, INA/GRM;
 - Les dispositifs multimédia et vidéo : fondation Langlois
- Archivage des contenus scientifiques ;
- Archivage des programmes informatiques ;
- Archivage des contenus télévisuels;
- Archivage des contenus cinématographiques;
- Archivage des contenus radiophonique;
- Autres...

Des types d'institution

- Radio Canada
- Cinémathèque du Québec
- BAC
- Les archives de la RAI
- ...

Des types de support

- Préservation holographique
- Préservation non numérique
- ADN, plaque de verre, etc.

Exemples de sujets traités par le passé

- L'orchestre de l'université de Montréal
- Club vidéo d'un CEGEP
- La musique des jeux vidéo
- Association Montréalaise « du monde à bicyclette »
- Archivage des réseaux sociaux
- Festival « nuits d'Afrique »
- ...
- Archives audiovisuelles de la congrégation Sainte Croix
- Wakoponi Mobile
- Vidéos projections architecturales (Artiste : Grandmaison)
- Conservatoire d'art dramatique de Montréal
- Un support matériel : Sony Rolly
- ...

Sujet de réflexion

- En quoi le fait de recourir à des médiums technologiques (audiovisuel, numérique) entraîne-t-il un nouveau rapport à la mémoire et à la connaissance ? Que deviennent les documents, leur rôle, et sous quelle forme ?

Documents

Notions à propos du document



Contenu

- **Forme sémiotique d'expression** associée à un **support matériel de manifestation** qui lui prête sa matérialité.
- La forme d'expression correspond à ce qui est perceptible et interprétable comme véhiculant un sens.
- Une forme d'expression peut être
 - sonore ou visuelle (type de modalité perceptive mobilisée),
 - verbale, imagée ou plus largement esthétique. (type de langage exprimé)

Forme et expression : l'héritage de Hjemslev

Contenu

Forme



Substance

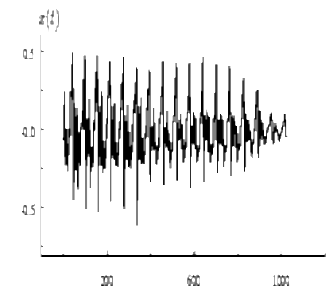
*tout ce qui
est pensable*

Expression

Forme

chien

Substance



Du contenu à l'inscription

➤ Inscription : Définition

- Tout contenu fixé sur un support matériel. La notion de fixation indique que le contenu obtient une stabilité et pérennité grâce au lien qui l'unit au support matériel. La pérennité et stabilité du support se communique à la forme d'expression.

➤ Inscription et contenu : Remarque

- Tout contenu n'est pas nécessairement une inscription :
 - un contenu verbal n'est pas fixé sur un support et disparaît avec l'évanescence du flux sonore.
 - L'inscription correspondrait ici à son enregistrement sur une piste magnétique, où un signal pérenne sur le support magnétique consigne la forme verbale orale.

De l'inscription au document

- Document:
 - Le document est une inscription qui s'insère dans un système de production et de consultation.
 - Le contenu associé prend son sens en fonction de la nature et du contexte documentaire.
- Il s'inscrit dans un contexte éditorial :
 - Des genres d'écriture (des comptes-rendus, des thèses, des articles scientifiques, etc.)
 - Des genres de lecture
 - Dans la lignée du genre d'écriture (étude pour un manuel, loisir pour un roman, etc.)
 - Dans une perspective méta : humanités (numériques), corpus d'apprentissage (IA), etc.

Deux grands types de documents

➤ Enregistrement (record)

- Tout document créé par une personne physique ou morale dans le cadre de ses activités comme un instrument ou un produit de cette activité.
L'enregistrement est donc un effet de bord, un épiphénomène de l'activité.

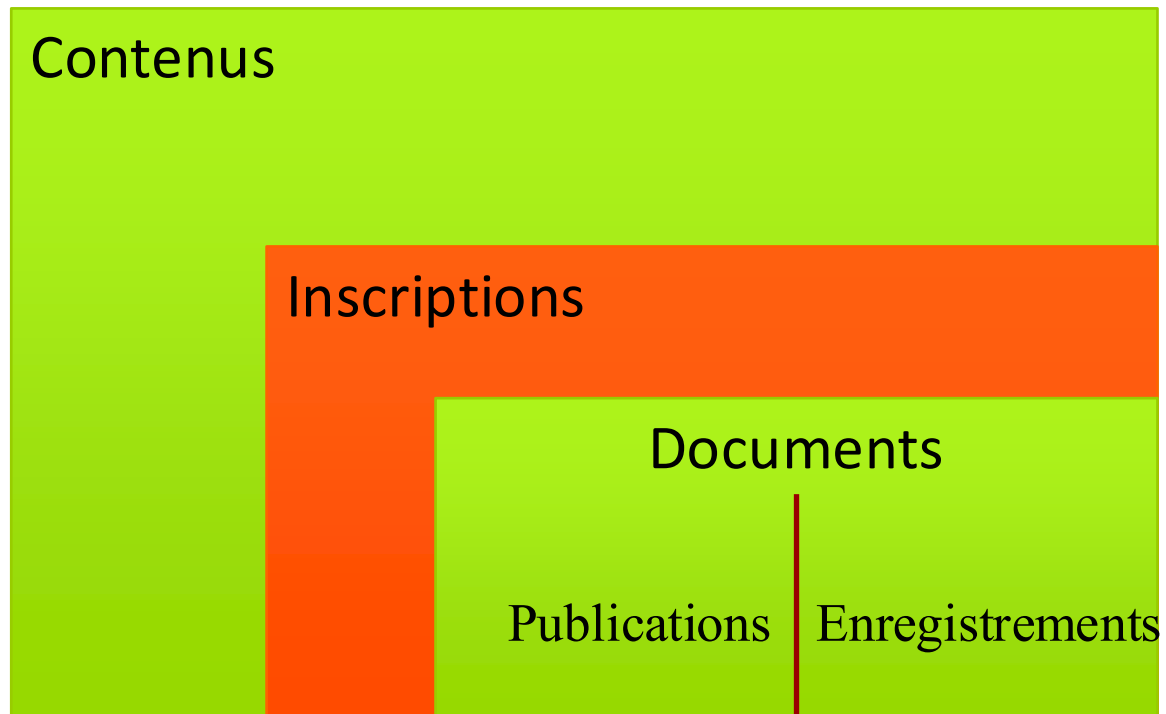
- L'enregistrement est un moyen au service d'une activité;

➤ Publication

- Tout document conçu pour être diffusé et distribué dans une communauté ou vers des publics au sens large.

- La publication est une fin en soi.

En gros...



Le contenu: notion hybride

- Le contenu a un support:
 - Objet matériel, structuré techniquement.
 - La structuration technique se fait par des *formats*:
 - Format physique (comment on réalise des 0 et des 1)
 - Des trous/bosses, tensions fortes/faibles, etc.
 - formats de codage (ce que codent les 0 et les 1) :
 - Des chiffres ou des lettres (e.g. ASCII...)
 - formats logiques (ce que représentent les 0 et les 1) :
 - Un schéma de base de données, une syntaxe formelle
- Le contenu exprime un sens :
 - *Forme* sémiotique permettant de la part d'un interprète de mener un parcours interprétatif.
 - Cette interprétation dégage des *signes*, ce qui fait sens, signifie.

Format et forme

Formatage

A priori
Construction technique des
unités

Données



Information

Forme

A posteriori
Unités définies au terme
des parcours
interprétatifs

Signes



Signification

Format et forme

Formatage

**Système technique:
Manipulation que l'on peut
mener**

Données



Information

Forme

**Système de valeurs:
Sens à partir d'un
système d'opposition
(Saussure)**

Signes



Signification

Manipuler, interpréter

Unités techniques de manipulation (UTM)

- Ce que la technique saisit pour transformer le contenu matériel.
- Définies a priori.
- Plus petite unité manipulable :
 - Pixel, caractère, image....

Unités sémiotiques d'interprétation (USI)

- Ce qui fait sens dans une interprétation.
- Définies a posteriori, comme résultant de l'interprétation et non comme la conditionnant.
- Plus petite unité porteuse de sens :
 - Signe, morphème...

L'enjeu est d'articuler UTM et USI : surmonter l'arbitraire de l'interprétation.

Manipuler et interpréter

- Problème - Cohérence entre les UTM's et les USIs :
- ce que l'on sait manipuler n'a pas toujours du sens ;
- ce qui a du sens n'est pas adressable par la technique.

Ceci est un texte qui ne dit pas grand chose
car il n'est là que pour servir d'exemple.

Cependant, il doit se prêter aux besoins de
l'illustration du propos et s'adapter à ses
exigences.

Par exemple, à la queue leu leu est une
expression lexicalisée où les blancs sont non
signifiants.

Notions supplémentaires

- Contenu culturel (e.g. vidéo, texte)
 - Contenu ***formaté techniquement*** pour être manipulable
 - Les unités composant le format technique n'ont pas forcément de sens, et leur combinaison donne un sens imprédictible a priori.
 - Les unités significantes dégagées par l'interprétation ne sont pas manipulables au niveau du format technique.
- Contenu formalisé / information (e.g. base de données)
 - Inscription ***formatée logiquement*** pour être manipulable en fonction du sens, et pour définir un sens en fonction de la manipulation.
 - Les unités logiques sont des unités techniques manipulables et des unités sémiotiques interprétables, la manipulation ayant et donnant du sens.
- Données (e.g. âge = 14) :
 - La plus petite partie significative d'une information.
 - Assertion minimale formatée et formalisée
 - Minimale: ne contient pas de donnée
 - Formatée : exprimée dans un code une syntaxe (on l'analyse / *parse*)
 - Formalisée : la syntaxe possède une sémantique associée.

Contenu, format, forme

- Format d'inscription :
 - c'est le format physique dans lequel l'information est consignée sur le support : l'encre du papier, le code binaire, etc.
- Format de manipulation :
 - c'est le format sous lequel l'information consignée est manipulable : le code binaire, la feuille de papier, l'alphabet ascii, le pixel, etc.
- Forme d'interprétation :
 - c'est le format sous lequel l'information s'interprète : les unités signifiantes. Elles sont dégagées par le parcours interprétatif, et amorcées par la segmentation perceptive du format d'inscription (le blanc entre les mots).

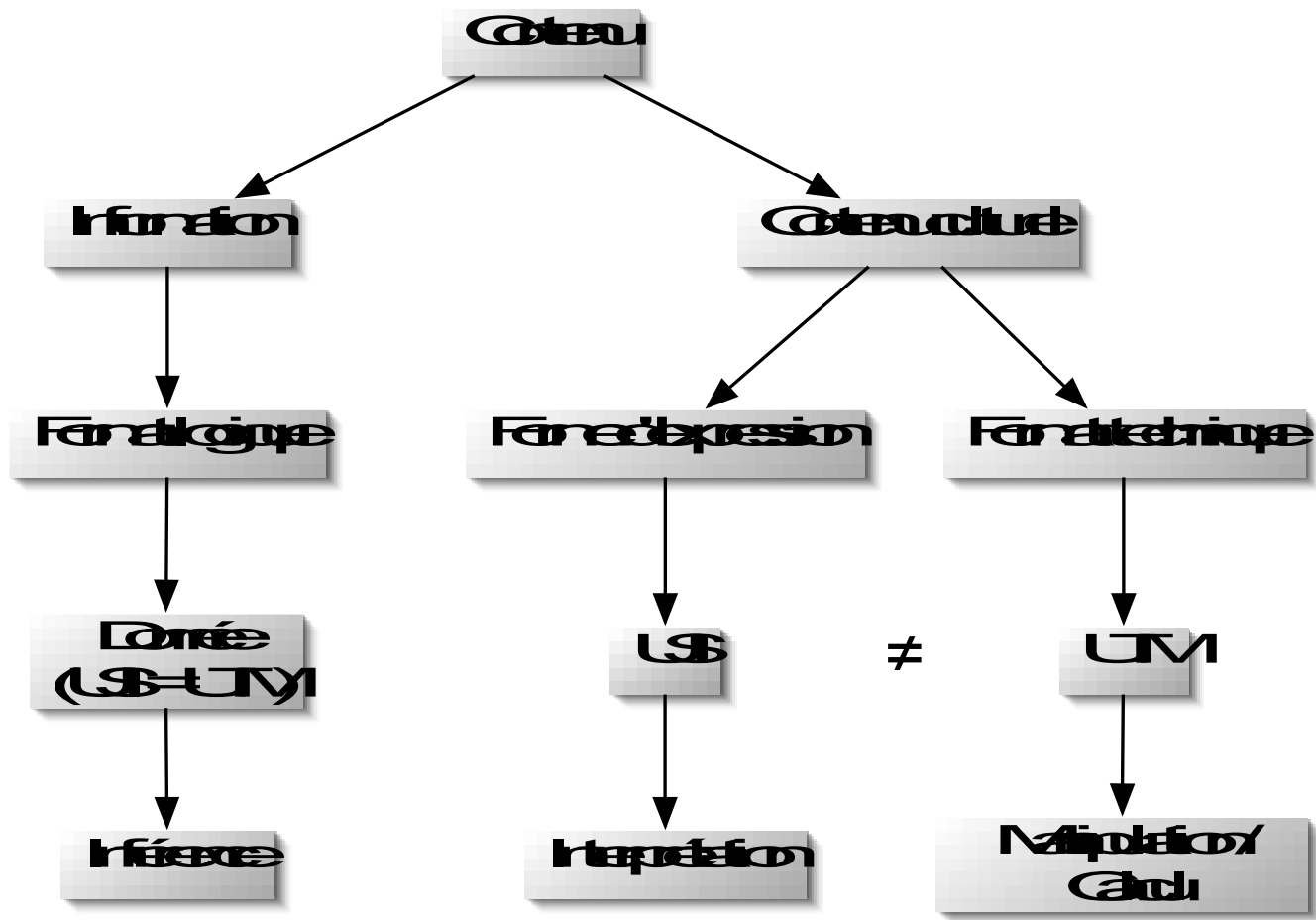
Contenus numériques, culturels, formalisés

- On pose que :
 - un contenu est numérique
 - quand les formats d'inscription et de manipulation sont confondus.
 - un contenu est culturel
 - quand la forme d'interprétation et le format de manipulation sont irréductibles l'un à l'autre et en un rapport arbitraire : manipuler le contenu à des conséquences arbitraires et non prévisibles sur l'interprétation.
 - Un contenu est formalisé
 - quand le format de manipulation devient une syntaxe permettant d'assumer et de contrôler une interprétation, en général définie par rapport à un modèle du monde.

Tension entre signe et information

- Toute inscription formatée a un double statut:
 - Comme information :
 - Le formatage technique permet de manipuler et d'associer un contenu informationnel ;
 - Comme signe:
 - En tant que contenu, l'inscription adopte une forme sémiotique appelant l'interprétation via un système de valeurs, comme signe.
- Ces deux statuts ne sont pas d'emblée cohérents
 - Ce qui peut être dit d'un point de vue informationnel peut ne pas avoir de corrélat au niveau sémiotique
 - Ce qui a un sens au niveau sémiotique peut ne pas se définir au niveau informationnel.

Contenu, forme, format.



En résumé

- Contenu :
 - Expression interprétable associée à un support matériel
- Inscription :
 - Toute information associée ou fixée à un support dans une forme stable. Elle est formatée techniquement en unités techniques de manipulation (UTM). Elle est interprétée en unités sémiotiques d'interprétation (USI).
- Document :
 - Inscription dans un contexte de production et de réception (genres d'écriture et de lecture)
 - Enregistrement :
 - Document produit par effet de bord de l'activité
 - Publication :
 - Document constituant la finalité de l'activité.

- UTM
 - Unité technique de manipulation
 - Plus petite entité manipulable
- USI
 - Unité sémiotique d'interprétation
 - Plus petite entité porteuse de sens
- Contenus numériques :
 - Format d'inscription = format de manipulation
- Contenus culturels :
 - UTI ≠ USI
 - Le format de manipulation et forme d'interprétation irréductibles
- Contenus formalisés
 - UTI = USI
 - Information, dont la plus petite partie est la donnée.
 - Forme d'interprétation rabattue sur le format de manipulation

Documents

Cerner la notion de document dans son contexte



Étymologie

- Latin documentum, dérive du verbe docere, enseigner;
- Dérivation savante à partir d'un mot latin suffixé en -entum : ce sont essentiellement des objets (outils, produits d'une action)

Document	Docere = enseigner
Instrument	Instruere = construire
Monument	Monere = avertir
Tourment	Torquere = torturer
Ferment	Fervere = bouillir

Étymologie *documentum*

- Au départ : *documentum* signifie la leçon, l'enseignement :
« faites voir que vous profitez des bons documents qu'on vous donne ».
- Toute chose qui apporte un enseignement, qu'elle soit écrite ou non, permanente ou fugitive.
 - 18e : sens judiciaire ; On passe de l'enseignement au renseignement, de la source de savoir, quelle qu'en soit la nature, à la pièce écrite. Le sens de document se restreint au domaine de l'écrit et du matériel durable, propre à être une pièce à conviction.
 - 19e : Est document toute source d'information authentique permettant d'étayer une affirmation, une théorie, une étude, notamment les sources écrites. « Chose qui enseigne ou renseigne ; titre, preuve. Un document précieux. Les documents font défaut pour établir ce point d'histoire ». Littré.

Étymologie : *documentare*

- Au départ :
 - documenter, c'est renseigner quelqu'un, l'avertir.
- Ensuite :
 - documenter, c'est étayer un ouvrage à l'aide de documents.

Document et documentation

➤ Définition :

- « Document : toute base de connaissances, fixée matériellement, susceptible d'être utilisée pour consultation, étude ou preuve. Exemples : manuscrits, imprimés, représentation graphique ou figurées, objets de collections, etc. »

➤ Pionniers :

- Paul Otlet (1868-1944) : les objets eux-mêmes peuvent être des documents : fouilles archéologiques notamment.
- Suzanne Briet (1894-1989): « un document est une preuve à l'appui d'un fait (..) tout indice concret ou symbolique, conservé ou enregistré, aux fins de représenter, de reconstituer ou de prouver un phénomène ou physique ou intellectuel » Qu'est-ce que la documentation ? 1951.

Jusqu'où étendre la notion de document ?

- « L'antilope qui court dans les plaines d'Afrique ne peut être considérée comme un document, (...) mais si elle est capturée (...) et devient un objet d'étude, on la considère alors comme un document. Elle devient une preuve physique (...) Les articles scolaires écrits sur l'antilope sont des documents secondaires, car l'antilope elle-même est le document premier. » S. Briet, ibidem

Documents, ou pas ?

Objet	Document
Étoile dans le ciel	Non
Photo d'étoile	Oui
Pierre dans la rivière	Non
Pierre dans un musée	Oui
Animal en liberté	Non
Animal dans un zoo	Oui

Document et intentionnalité

- Un document est la trace d'une expression. Objet matériel, il faut le considérer non pour ce qu'il est matériellement, mais pour ce qu'il n'est pas mais signifie:
 - un document est un objet institué par un point de vue intentionnel comme expression, trace matérielle d'un contenu.
 - Le document n'est pas un instrument ou un outil, mais un renseignement : il apporte une médiation par la signification et non par son rôle causal physique (livre qui bloque la porte versus livre qui m'apprend ce qu'est une porte).
- Le point de vue intentionnel peut être :
 - A priori : le document est créé pour être la trace d'un contenu. C'est le cas de tous les documents publiés, des notes personnelles, etc.
 - A posteriori : un objet matériel est considéré comme une trace d'une activité que l'on veut comprendre à partir de l'interprétation de l'objet. Ce dernier devient un document, une trace ou témoignage dont l'interprétation peut enseigner sur l'activité dont il est issu.
 - Exemple: les objets archéologiques.

Caractéristiques d'un document

- Être délimité dans le temps et l'espace :
 - le document a un début et une fin (livre, vidéo) ;
 - un flux infini n'est pas un document : flux d'une caméra de surveillance, diffusion en continu ;
 - un réseau infini n'est pas un document : le Web ;
- Être permanent :
 - Assurer la conservation d'un contenu ; témoignage d'une expression et support d'une transmission.
- Être perceptible :
 - Pouvoir percevoir le document comme tel, le reconnaître comme porteur de sens et d'enseignement sur une des entités qui lui sont externes.
- Être intelligible :
 - Pouvoir interpréter la forme signifiante perçue comme telle (signifiante, donc signifiant quelque chose). Le document donne lieu à une interprétation pour recouvrer le contenu à partir de la trace qu'il constitue.

Tentative de caractérisation (Buckland)

- un document est matériel : les objets physiques et les signes physiques seulement.
- il est intentionnel : il doit se rapporter à autre chose (*intendo* en latin, *aboutness*) ;
 - Ce rapport est un rapport de preuve, ou de documentation : apporter une information sur quelque chose.
- les objets doivent être traités : ils doivent être transformés en document ;
- il a une position phénoménologique : l'objet doit être perçu pour être un document.

« C'est la qualité d'être placé dans une relation organisée et significative avec d'autres preuves qui donne à l'objet le statut de document. »

Du document au complexe documentaire

- Le document est un objet matériel singulier permettant :
 - D'assurer une permanence dans le temps
 - De permettre un accès direct au contenu, ce dernier étant perceptible et intelligible selon une forme sémiotique.
- Si un objet ne présente pas ces caractéristiques, il ne peut être un document:
 - Le document audiovisuel, numérique est-il un document ?

La déconstruction du document

Permanent

Permanent

Permanent

Non
perceptible



Non
permanent

Perceptible

Complexe documentaire



Perceptible

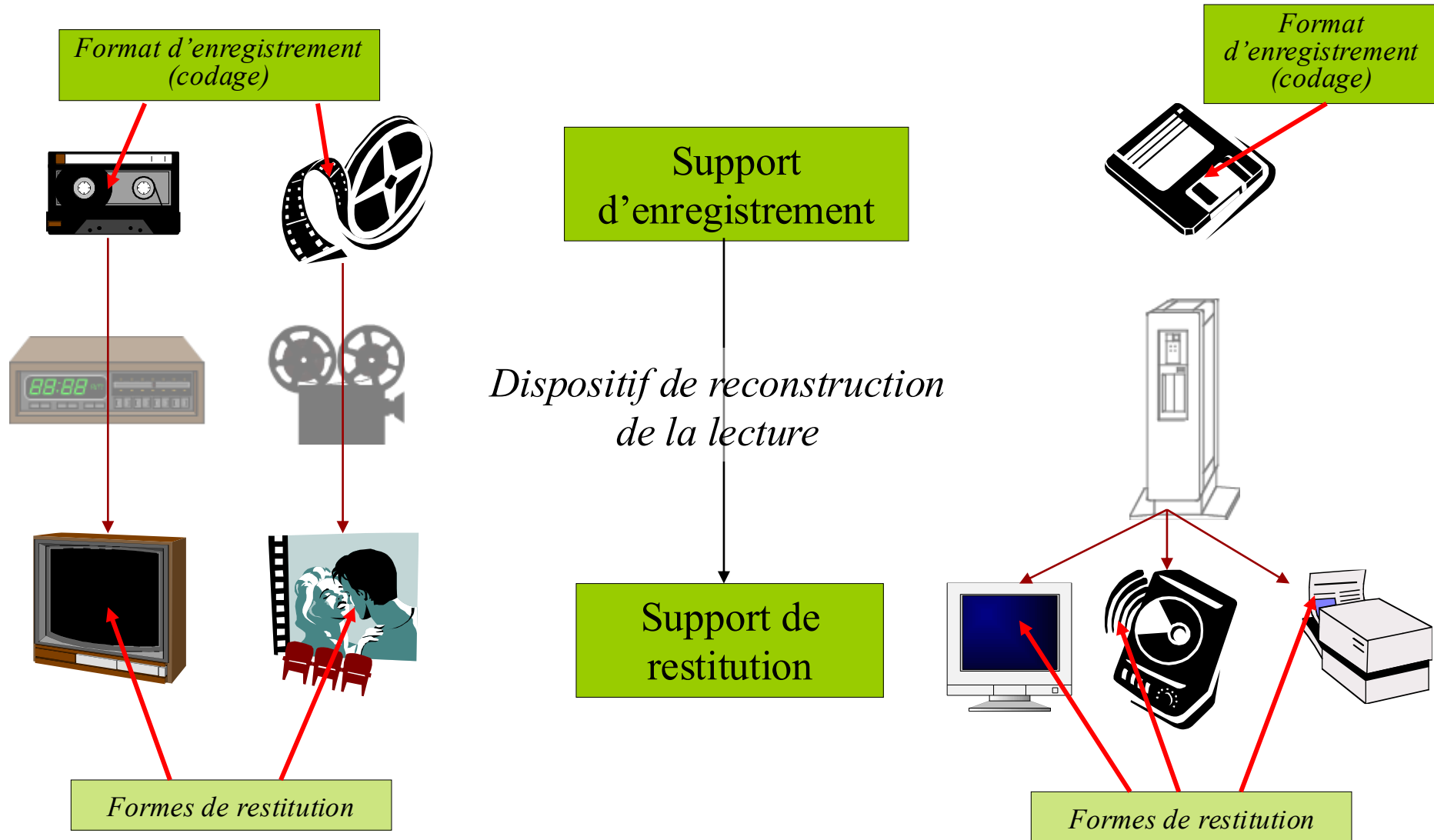
Perceptible

Dimensions d'un document

Archivistique audiovisuelle et numérique



Les dimensions documentaires⁶⁶



Dimension documentaire : enregistrement

- Support d'enregistrement :
 - support sur lequel un document est enregistré ;
 - C'est la dimension de la conservation et de la préservation du contenu. Ce support est nécessairement spatial, car seul l'espace propose une dimension de la permanence, alors que le temps est la dimension de la succession.
- Format d'enregistrement :
 - format sous laquelle un document est enregistré ; c'est l'enregistrement interne, manipulable ou calculable ;
 - format binaire ; codage ascii ; signal magnétique; écriture alphabétique ; etc. ;

Dimension documentaire : restitution

- Support de restitution :
 - support sur lequel est projeté l'enregistrement interne en une forme appropriable par un lecteur humain ; exemple : écran, papier, haut-parleur ;
- Mode de restitution :
 - modalité perceptive sollicitée par la forme de restitution ;
 - exemple : vue, ouïe, toucher ;

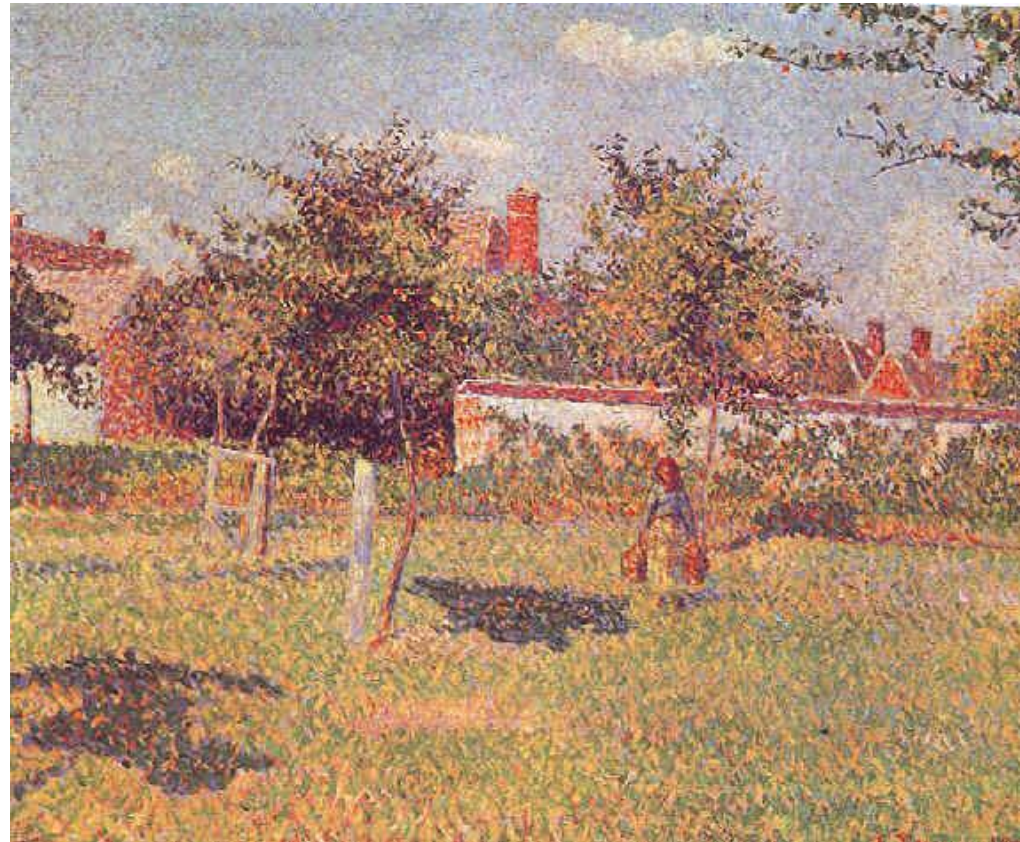
Formes de restitution

- Format physique de restitution :
 - signal physique reconstruit à partir de la forme d'enregistrement, de manière à constituer un objet matériel possédant un sens directement appréhendable par un être humain. C'est un signal physique véhiculant une forme sémiotique.
 - Pixels d'un écran de télévision, points d'une page imprimée...
- Forme sémiotique de restitution :
 - forme sémiotique sous laquelle est configuré l'enregistrement interne permettant à un lecteur humain d'appréhender le contenu de l'enregistrement ;
 - Écritures, graphiques, images, sons, etc. ;

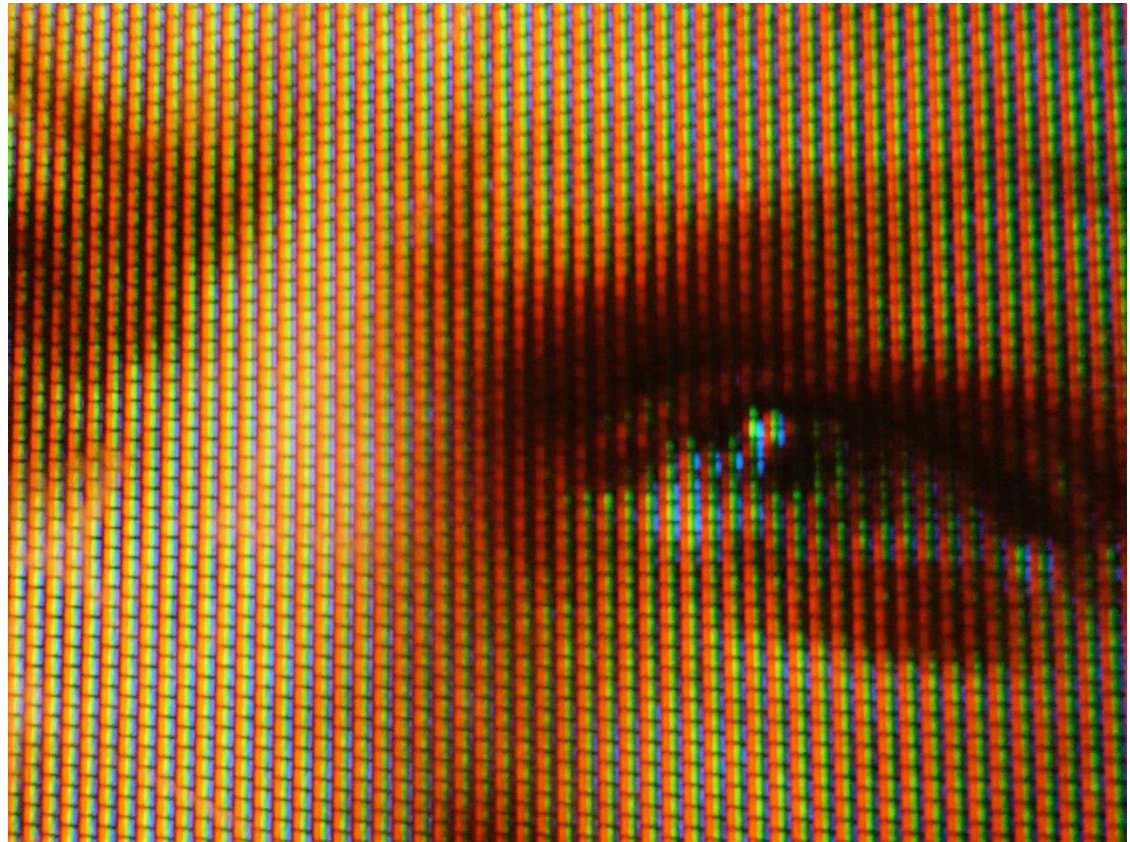
Vieux constat : le pointillisme



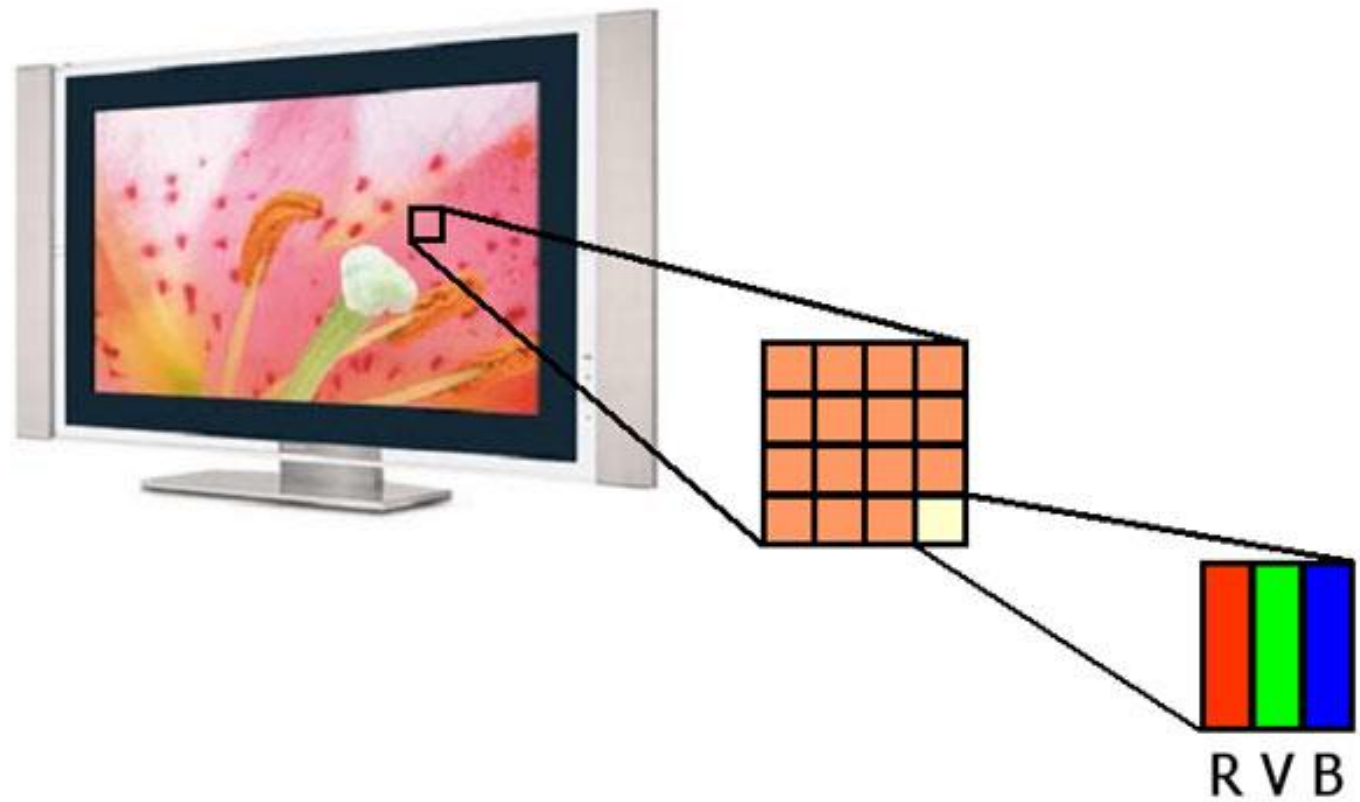
De même...



Plus moderne...



C'est-à-dire...



Précisions

- Multimodal :
 - un document est multimodal quand il mobilise plusieurs modes de restitution ;
- Multimédia :
 - un document est multimédia quand il mobilise plusieurs formes sémiotiques de restitution.

Lecture et décodage

- Distinguer entre les formes d'un document :
 - le format d'enregistrement est liée au support d'enregistrement. Il est donc spatial. C'est lui qui véhicule l'information pour accéder au contenu via le support. Il constitue donc un codage physique du contenu sur le support.
 - La consultation du format d'enregistrement relève d'un décodage.
 - la forme sémiotique de restitution est la forme sous laquelle la conscience accède au contenu. C'est donc la forme sous laquelle le contenu fait sens, fait signe à la conscience.
 - La compréhension de la forme sémiotique relève d'un apprentissage culturel.

Dispositifs

- Le document mobilise différents dispositifs:
 - le ***dispositif d'enregistrement***, qui permet de coder le contenu dans la forme d'enregistrement, qui informe le support d'enregistrement.
 - le ***dispositif de restitution***, qui permet de construire la forme physique de restitution à partir de la forme d'enregistrement.
 - le ***dispositif de lecture***, qui permet d'accéder au contenu à partir de la forme sémiotique de restitution, véhiculée par la forme physique de restitution.
 - le ***dispositif d'écriture***, qui permet d'exprimer le contenu dans une forme sémiotique de restitution.

Remarques sur les dispositifs

- Le dispositif d'enregistrement et de restitution peuvent être mécaniques ou automatiques. Les dispositifs de lecture et d'écriture ne peut l'être, puisqu'il s'agit d'un accès au sens.
- Il n'existe pas d'accès direct au contenu ; par conséquent :
 - la lecture correspond au passage de la forme sémiotique du document à une forme mentale, orale, etc.. Il s'agit d'une réinscription, et la compréhension ou accès au sens correspond au processus de la réécriture.
 - l'écriture correspond au passage d'une forme sémiotique interne (mentale, corporelle, etc.) ou externe (la forme d'un autre document), à la forme sémiotique du document.



Quelques cas particuliers, mais exemplaires

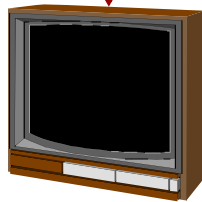
Les documents papiers / Les documents temporels / Les documents numériques

Livre :
Consultation



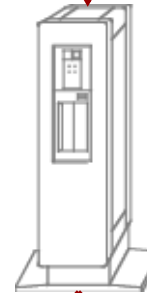
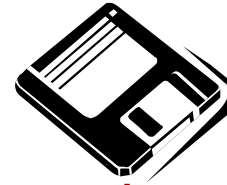
Ce qui est consulté est identique à ce qui est conservé : archiver c'est garder intact.

Audiovisuel :
Reconstruction

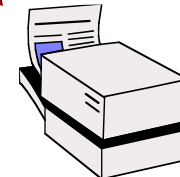
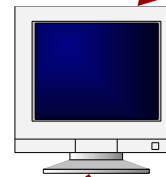


Ce qui est consulté est reconstruit de manière dédiée à partir de la ressource: archiver c'est reconstruire à l'identique.

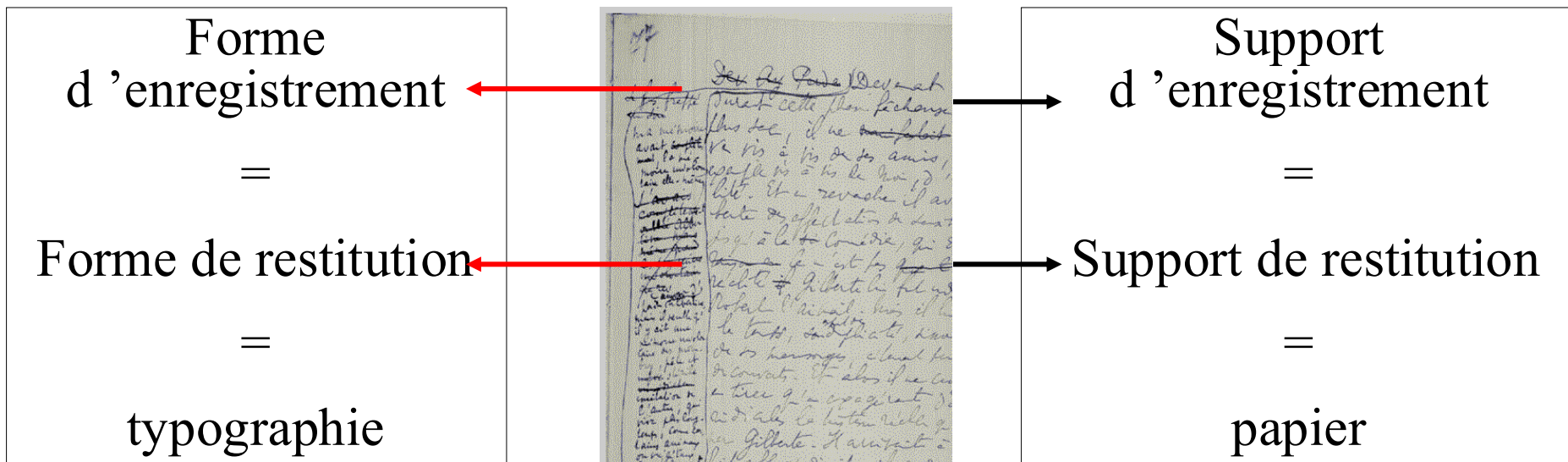
Numérique :
Ré-invention



Ce qui est consulté est reconstruit de manière plurielle à partir de la ressource: archiver c'est confronter des variantes à une ressource anonyme



Un cas particulier: le papier



Dispositif de lecture : le lecteur

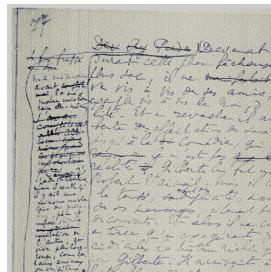
Conséquences pour les documents papiers

- ce qui conservé est ce qui est relu.
- symétrie entre l'enregistrement et la lecture :
 - l'enregistrement est une écriture, i.e. expression en une forme sémiotique d'un contenu, c'est-à-dire une autre inscription puisque le contenu est inaccessible en tant que tel ; la lecture est l'inscription en une forme orale ou mentale du contenu, c'est-à-dire de la forme sémiotique du document.
- Il n'y a pas de dispositif d'enregistrement ou de restitution, mais seulement de lecture et d'écriture.

Documents temporels

Formes *statiques/spatiales*

- Les structures interprétables sont présentées simultanément sur le support de lecture ;
- L'ordre et le rythme de lecture sont déterminés par le lecteur.



Formes *dynamiques/temporelles*

- Les structures interprétables sont présentées successivement sur le support de lecture ;
- L'ordre et le rythme de lecture sont imposés par le dispositif de lecture.



Caveat emptor !

- Une difficulté :
 - La lecture est un processus temporel imposé par la forme de restitution ;
 - L'enregistrement est spatial et statique ;

La situation est grave...

- Problème:
 - le format d'enregistrement est nécessairement spatial.
 - Si la forme sémiotique est temporelle, elle ne peut être véhiculée par la forme d'enregistrement.
 - Il ne peut donc y avoir coïncidence entre la forme physique de restitution et la forme d'enregistrement.

Mais pas désespérée : la solution du codage

- La forme d'enregistrement assure la conservation du contenu sous une forme spatiale, à la condition que cette forme permette de reconstruire le contenu :
 - la forme d'enregistrement est un codage de la forme physique de restitution ;
 - un dispositif mécanique de restitution permet de reconstruire la forme physique de restitution à partir de la forme d'enregistrement.
 - le document archivé n'est plus le document, mais l'enregistrement de son codage.

La nécessité du codage

- L'enregistrement des documents temporels n'est possible que si l'on sait jouer le temps, si l'on sait reconstruire un déroulement temporel à partir d'une forme spatiale :
 - on a une ingénierie documentaire puisqu'il faut un mécanisme qui reconstruise la forme temporelle à partir du code ;
 - la notion de mécanisme est celle de la prescription d'un déroulement temporel en fonction d'un agencement spatial ;
 - le mécanisme est donc par essence un programme, si bien que le numérique n'est que l'aboutissement inéluctable de la mécanisation de la reconstruction documentaire.

Mécanique et espace

La boîte à musique

www.universe-of-luxury.com



Le principe du codage sur le cylindre



Document numérique

- Le document numérique n'est pas un document :
 - l'enregistrement numérique constitue une ressource à partir de laquelle un document peut être reconstruit sur un support de restitution ;
 - le document est une entité dynamique n'existant que lors de la projection ;
- Évolution avec le numérique :
 - Les documents habituels confondent formes et supports de restitution et d'enregistrement ;
 - le numérique distingue restitution et enregistrement et rompt l'unité documentaire ;

Conséquence pour les documents numériques

- Le document n'est plus ce qui est conservé, mais reconstruit à partir d'un potentiel calculatoire :
 - projection canonique qui restaure la forme de restitution pour laquelle le document a été créé ;
 - E.g. : reproduire une vidéo à partir de son enregistrement ;
 - projections non canoniques permettant de construire d'autres documents ;
 - E.g. : faire une reproduction sonore à partir de l'enregistrement d'une vidéo.

Quel statut des documents reconstruits ? Quelle authenticité ?

Document temporel et numérique

➤ Ressource :

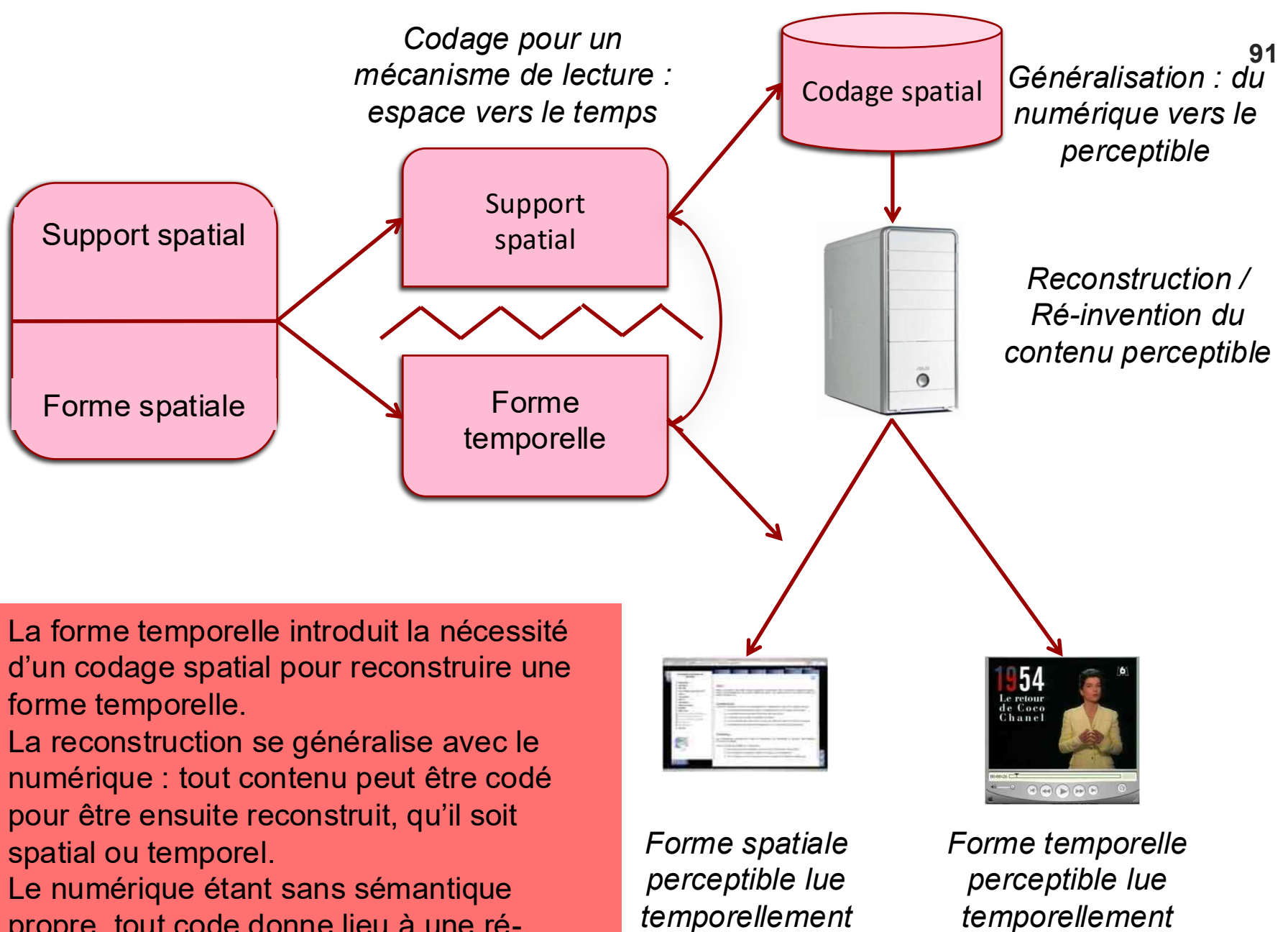
- ce n'est pas le document, ni le contenu, mais un objet à partir duquel reconstruire une forme lisible.

➤ Programme :

- entité calculant et transformant la ressource.

➤ Formes publiées :

- elles sont des reconstructions à partir de la ressource. Elles sont multiples et donnent autant de points de vue possibles sur la ressource.



1. La forme temporelle introduit la nécessité d'un codage spatial pour reconstruire une forme temporelle.
2. La reconstruction se généralise avec le numérique : tout contenu peut être codé pour être ensuite reconstruit, qu'il soit spatial ou temporel.
3. Le numérique étant sans sémantique propre, tout code donne lieu à une ré-invention en fonction du codage adopté

Où est le document ?

- Le document n'est pas la ressource, car elle n'est pas perceptible ;
- Le document n'est pas la forme publiée, car elle n'est pas permanente (cf. film).
- Le numérique fait rupture car il remet en cause, au sein même de l'intégrité documentaire, son identité : il est par définition toujours reconstruit et réinventé.
- *Que signifie conserver un contenu dont la monstration est par nature une ré-invention ?*
- *Nouvelle problématique: comment fonder une authenticité sur un absence d'intégrité ?*