

Lecture active et annotations

Yannick Prié
LIRIS – Université Lyon 1
M2R / 2006-2007



Plan

- Lecture active et annotations
- Applications sur le web
- Applications dans l'audiovisuel

2

Plan

- Lecture active et annotations
- Applications sur le web
- Applications dans l'audiovisuel



[Bellenger, 1978]

Plusieurs types de lectures

- lecture jouissance
 - goûter le texte en tant que tel
- lecture appropriation
 - comprendre en profondeur la teneur du texte
- lecture exploration
 - visant la recherche d'une information supposée ou sue présente dans le texte
- lecture assimilation
 - visant une connaissance du texte en tant que tel, par exemple pour faire un compte-rendu
- lecture d'inspiration
 - en vue de rebond vers d'autres lectures ou des réalisations externes au texte

lecture professionnelle

Flechon et Enard (1994)



3

4

Lecture active

- Lire en écrivant - marquant
 - annotations comme marques dans un espace de connaissance documentaire, largement textuel
 - activité tendue vers la réutilisation de la lecture
- Etudier la lecture active
 - décrire les grandes catégories d'annotations
 - penser et théoriser la pratique d'annotation comme utilisation du document en vue de produire d'autres documents
 - réfléchir de façon plus générale sur la lecture électronique, ce l'écriture hypertexte, la création et le partage de connaissances, etc.
- Analyses
 - Dimensions de l'annotation [Marshall 98]
 - Forme, lieux et objectifs [Virbel & co.]



[Marshall 98]

Dimensions de l'annotation forme de l'annotation

- Annotation formelle vs informelle
 - des méta-données avec un schéma pour l'interopérabilité ... aux notes de marges
 - possibilité de passage de l'informel vers le formel (ex : smiley)
- Annotation explicite vs tacite
 - lié à l'intelligibilité de l'annotation à son auto-suffisance
 - du « No! » au soulignement,
 - rarement au-delà de la stricte nécessité pour la tâche en cours
 - pourquoi expliciter si ce n'est pas nécessaire ?
 - une annotation peut être interprétable
 - par son auteur (un certain temps)
 - par autrui éventuellement



5

6

Dimensions de l'annotation fonctions de l'annotation

[Marshall 98]



- Annotation comme écriture vs annotation comme lecture
 - « communier » avec le document, entre
 - à un bout, le lecteur de M. de Certeau, qui ne fait que se promener... et oublier
 - à l'autre, l'« idéal postmoderne de l'hypertexte polyvocal » : le texte comme un médium participatif dans lequel l'annotateur est écrivain
 - l'annotation est un pont entre l'écriture et la lecture.
 - fonctions [Desmontil / Jacquin 2002]
 - information, illustration, extension du document
 - forum
 - opérationnalisation
 - aide au processus rédactionnel
 - support de lecture

7

Dimensions de l'annotation comm. entre auteurs et lecteurs

[Marshall 98]



- Annotation *hyperextensive* vs *extensive* vs *intensive* (niveau de lecture du document)
 - annotations orientées liens (le texte dans son entour hypertexte, lecture superficielle)
 - notes de lecture rapide (typographie : mise en gras ou en italique par exemple, surlignage)
 - annotations orientées texte uniquement (pour une lecture approfondie)
- Annotations permanentes vs éphémère
 - note de bas de page vs note de rédaction

8

Dimensions de l'annotation comm. entre auteurs et lecteurs

[Marshall 98]



- Annotations publiées vs privées
 - privé pur (evt. diffusé)
 - note pour soi (evt. passées anonymement)
 - public – publié
 - one to many
 - note de bas de page
 - édition savante
 - many to many
 - forum
- Annotation globale vs institutionnelle vs groupe de travail vs personnelle
 - tout est possible

9

« Ecole toulousaine » Forme et lieux de l'annotation

[Virbel & co.]



- Formes
 - mise en valeur
 - surlignage, changement de mise en forme
 - liens
 - à partir du texte, vers le texte, à double sens
 - ajouts
 - postits, objets quelconques
 - pour modification
 - correction, restructuration
 - annotations type forums
 - annotations d'annotations
- Lieux
 - dans la page (sur, autour du texte)
 - entre les pages (signet, intercalaire)
 - hors-document (cahier de note)

10

« Ecole toulousaine » Objectifs de l'annotation

[Virbel & co.]



- Classifier
 - annotation = trace de restructuration du document par le lecteur
 - hiérarchiser
 - jugement qualitatif suivant l'importance, la représentativité
 - architecturer
 - affectation de catégories logico-linguistiques à différents objets
 - document entier, division, liste, paragraphe, unité de texte
 - contextualiser une partie du document par une autre
- Compléter
 - reformuler : modifier le contenu du texte
 - développer, reformuler, résumer
 - commenter
 - documenter : attacher un objet quelconque
- Corréler deux passages)
- Programmer de future actions
 - traduire, relire

11

Quelques remarques



- Liens entre objectifs et moyens
 - ambiguïté de la relation, multivocité
 - certaines relations sont fortement improbables
 - difficile de reformuler en soulignant
- L'annotation est rarement un but en soi
- Statut de la lecture scientifique professionnelle
 - parangon qui en vaut un autre
 - largement fondée sur une pratique annotative
 - les chercheurs pensent à partir de leurs pratiques

12

Passage au numérique...



- Mimer ce l'activité de lecture active le papier
 - annotation en même temps qu'on lit pour aider la lecture
 - architecturer
 - résumer
 - préparer un traitement futur
- Automatiser des tâches à partir des annotations comme matériel numérique exploitable pour des outils sophistiqués
 - résumé : représentation concise
 - relecture assistée, mémoire
 - recherche d'information
 - classification / comparaison de documents
 - conversations
 - personnalisation
- Multiples formes possibles pour les annotations
 - mise en forme, posts, threads de discussion (conversation), liens, tours guidés, tags, notes, marques d'approbation...

13

Christian Jacob La lecture assistée par ordinateur



- « L'idée essentielle me semble être que l'on déconstruit le texte comme réalité physique associée au livre, au profit d'un modèle topographique dynamique, d'un ensemble de lieux qui peuvent être reliés soit par un parcours linéaire bidirectionnel (forward/backward) soit sous une forme plus cartographique avec des figures arborescentes voire labyrinthiques, de manière à suivre toutes les connexions qui se forment au cours d'un travail d'interprétation ou au cours d'une problématique.
- Ce double processus de balisage et d'établissement de liens correspond à une étape essentielle du travail intellectuel et permet une mobilisation immédiate, réitérable et visuelle de l'ensemble des intuitions que l'on peut avoir au cours de la lecture linéaire de l'oeuvre. »

14

Questions liées au numérique

[Veron 97



- Stockage
 - qui, où, comment
- Filtrage, sécurité et droits de lecture/écriture/modification des annotations
- Utilisabilité
 - liberté d'utilisation
- Uniformité
 - tout *in situ*, de la même forme : visualisation / annotation
 - extensibilité ?
- Robustesse et cohérence
- Acceptabilité par l'utilisateur
- Réseaux :
 - échelle et performance
 - protocoles, formats, plateformes

15

Annotation, navigation, édition électronique : vers une géographie de la connaissance

[Stiegler 2000



- Lecture savante = lecture professionnelle
 - Ex : lecture approfondie d'une œuvre, lecture d'une œuvre dans le contexte d'une autre
 - Retour à une situation d'avant le bon à tirer ?
- Deux grandes classes d'interventions sur le texte
 - Hiérarchisation
 - Qualification
- Annotation permanente, inscription opportuniste, mais mémoire finie
 - L'effet miette de pain du Petit Poucet
 - La machine est là pour aider, organiser l'appropriation des inscriptions, l'accès à la lecture passée

16

Annotation, navigation, édition électronique : vers une géographie de la connaissance

[Stiegler 2000



- Approche d'une lecture active et des descriptions formalisées des structures (DFS) en trois niveaux.
 - Niveau du fonds
 - Ex. structure de la bibliothèque
 - Niveau de la grille de lecture
 - Découpage / annotation / grille
 - Niveau de l'utilisation
 - synthèse, navigation, etc.
 - « niveau des historiques de la lecture effective, constituée en objet hypertextuel qui est la grille de lecture dynamique. »
- D'où 4 grandes catégories d'opérations, qui sont les fonctionnalités de base de l'annotation électronique.
 - Hiérarchisation
 - Qualification (annotation / commentaire)
 - Navigation et recherche
 - Représentation (assistance à la navigation et à la recherche).
- Besoins de standard normatifs suivant les activités, mais qui permettent l'émergence de singularités.

17

Annotation et lecture active



- L'annotation
 - peut être considérée à de nombreux niveaux
 - est rarement une fin en soi
 - ressort de l'écriture
 - s'intègre dans une activité productrice de sens, et souvent de documents
- Actuellement
 - des systèmes qui imitent la pratique papier, l'enrichissent
 - standby comme le livre électronique ?
 - constitution d'une lecture véritablement active
 - sur le web, des pas vers un hypertexte plus riche
 - en parallèle à une socialisation grandissante : blogs, web 2.0, document de groupes (vs. auteurs), documents pour l'action
 - ailleurs : exemple audiovisuel

18

Plan

- Lecture active et annotations
- **Applications sur le web**
- Applications dans l'audiovisuel



19

Le web, un système d'information global

- Espace de connaissances
 - composé de ressources plus ou moins structurées
 - avec des liens simples et universels
 - ... et ça fonctionne !
- Idée d'amélioration
 - associer aux ressources des informations utiles afin de favoriser leur exploitation
 - annotations / méta-données
- Questions
 - quelles ressources ?
 - quelles informations
 - utiles pour qui ?
 - utiles pour quoi faire ?



20

Première vision

- Web = espace documentaire ultime (pour l'instant)
 - naturellement distribué et bruité
 - juste un peu hypertexte (comparé à l'[[hypertexte historique](#)])
 - essentiellement dédié à la lecture, peu à l'écriture
 - auquel manquent des choses par rapport à l'hypertexte ultime
 - lecteur passif → lecteur actif
 - one to many → many to many
 - interopérabilité des marques laissées
 - repérage communautaire
 - etc.
- Nécessités d'annotations
 - lecture,
 - enrichissement, liens,
 - discussion – commentaires,
 - partage
 - etc.



21

Deuxième vision

- Des informations
 - partout, dans tous les sens, incohérentes
 - auxquelles on peut difficilement accéder sans médiateurs automatiques
 - moteurs de recherche
- Nécessité
 - de liens riches
 - de méta-données structurées, sûres, et signifiantes
 - pour les agents logiciels
- Solution : structurer sémantiquement tout ça
 - annotations sémantiques
 - assertions logiques
 - pour un web sémantique



22

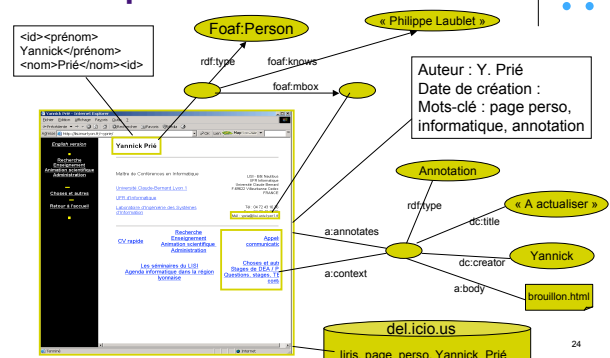
Annotation / méta-donnée

- Définition
 - « quelque chose » que l'on ajoute à une ressource Web
- Conséquences
 - cette information doit d'une manière ou d'une autre être structurée (utilisable) et descriptive de la ressource ou de son utilisation
 - la ressource en question doit exister et pouvoir être exploitée sur le Web indépendamment des informations qui lui sont associées



23

Exemples d'annotations



24

Annotations informelles



- Annoter
 - Construire des cartes de navigation
 - Localiser et étiqueter les informations intéressantes dans un document
 - Commenter sa lecture
 - Évaluer le contenu d'un document
 - Décrire des ressources
 - ...
- Pour
 - Visualiser plus tard les annotations
 - les siennes, celles d'autres
 - Construire automatiquement des index, des tables des matières, ...
 - Rechercher des informations : requêtes, filtrage
 - Résumer automatiquement des documents
 - Aider la navigation
 - Adapter des documents
 - ...

25

Annotation sur le WS



- Web sémantique
 - Des langages pour décrire des ressources de façon formelle de telle sorte que des machines puissent (également) utiliser ces ressources
 - annotations = instances de concepts OWL
- Pour rendre de nouveaux / meilleurs services aux utilisateurs du web
 - Recherche d'information « sémantique »
 - Rechercher des annotations-concepts avec des inférences
 - Utiliser les ontologies pour transformer les requêtes / réponses
 - Répondre à des questions
 - Composition
 - de documents, de fragments de documents, de services en vue de construire de nouveaux documents (cours à la carte, catalogues, réponses à une question, etc.) ou de nouveaux services (hôtel + spectacle)
 - Adaptation
 - Ex. construire un cours automatiquement
- Remarque
 - Utilisateur pas forcément conscient de l'utilisation de la ressource et des annotations liées

26

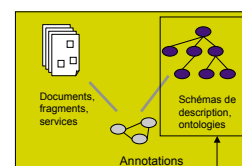
Annotations / méta-données



- Pas de différences fondamentales
 - annotations plus situées dans les documents
 - métadonnées plus éloignées
- Critères pour comparer les annotations sur le web
 - types de ressources concernés, plus ou moins fragmentées,
 - automatisation plus ou moins marquée de la mise en place,
 - structuration plus ou moins forte des « schémas »,
 - tâches soutenues (ou sous-tendues),
 - utilisation qui en est faite par les agents logiciels, plus ou moins « intelligente »

27

Discussion



- Web vs web sémantique
 - différence subtile, graduelle
 - différentes acceptions de « sémantique »
- Annotation informelle vs formelle
 - idem : visions caricaturales, mais pratiques
 - formalisme à divers degrés, sémantique à divers degrés
 - web (un peu) sémantique vs web sémantique
- Deux classes d'applications
 - annotations dans un web (un peu) sémantique
 - annotations dans (pour) le web sémantique

28

Fonctionnalités Création des annotations



- Ancre de l'annotation
 - document entier / fragment d'un document
- Valeur de l'annotation (contenu, forme...)
 - Annotation informelle
 - texte, la machine n'a pas accès au contenu « ultime »
 - Annotation formelle
 - obéissance à un schéma plus ou moins précis : ontologie ou autre
 - la machine est capable de faire des choses

29

Fonctionnalités Représentation, stockage, accès



- Représentation
 - texte structuré : XML, RDF/OWL
 - tables de BD plus ou moins dédiées
- Stockage
 - dans / hors document
 - éventuellement doublé si « moissonnage » des informations par des robots crawlers
 - local / distant
 - possibilité de proxy
 - centralisé / distribué
- Accès
 - contrôle d'accès : privé, public, groupe
 - local / distant

30

Fonctionnalités Utilisation des annotations



- Visualisation des annotations
 - Localisation
 - à côté du document
 - dans le document (nécessité reconstruction)
 - Nécessité de suivi de la navigation
 - interception / proxy
- Exploitation des annotations
 - côté client ou serveur
 - génération de sites, moteurs de recherche, etc.
 - utilisation des fonctionnalités dues à la représentation de connaissance sémantique... ou non

31

Quelques exemples Web (un peu) sémantique



- Annotation collaborative de pages
 - <http://www.sitesays.com/> (un parmi beaucoup)
- Del.icio.us
 - Gestion de bookmarks avec tags
- Flickr
 - Image, tags, commentaires de parties d'images

32

Quelques exemples Web sémantique



- Outils expérimentaux
 - Annozilla
 - Protocole W3C Annotea
 - Outil de création / annotation de pages
 - [[Ontomat](#) / [CREAM](#)]
 - SMORE (+modification ontologie)
 - COHSE : hypertexte sémantique
 - Ontologies simples + outils multiples
 - FOAF : Friend of A Friend
 - DOAP : Description Of A project

33

Discussion Développement des systèmes



- Systèmes généraux informels (tags, textes)
 - adoption à l'échelle du web
 - gros acteurs
- Systèmes plus formels
 - pas d'adoption à grande échelle
 - challenge technique (passage à l'échelle) / social (collaboration ?)
 - Pour une application qui perce
 - le bon langage
 - expressivité / possibilités d'inférences
 - une bonne représentation des connaissances
 - exprimée dans ce langage sur un certain domaine
 - une application d'annotation
 - annotation automatique / manuelle / semi-automatique
 - un système complet
 - sur le web : distribution des annotations et des calculs
 - rendant un vrai service
- Pas évident !

34

Discussion Modification, vie privée



- Disparition d'une ressource
 - Effacer l'annotation orpheline ?
 - Et ses réponses ?
 - Repositionner ?
- Confidentialité des données
 - Transmission protégée ?
 - Stockage protégé ?

35

Discussion Partage



- Tout le monde n'est pas intéressé par l'annotation de tout le monde
 - à part l'auteur de la page
 - cf. papier : trop d'annotation tue l'annotation
- Liberté de n'importe quoi
 - cf. noms de domaines
 - ne voir que les annotations de confiance?
 - review ?
- Droits d'auteurs d'annotation ?
 - autorisations d'accès

36

Discussion

Annotation sémantique



- Sur la sémantique du lien d'annotation
 - annoter une ressource ou décrire le monde ?
- Sur la notion d'auto-indexation
 - annotations a posteriori sur les ressources
 - vs conception de page / intégration de méta-données
 - est-on expert de la description de ses documents ?
- L'ontologie dirige ?
 - annoter en utilisant une ontologie
 - annoter en construisant/étendant
 - annoter d'abord, formaliser ensuite
 - SMORE vs OntoMat
- Maîtrise-t-on l'ontologie et ses conséquences pour pouvoir annoter ?
 - Qui est expert ?
 - Qui voit les méta-données ?

37

Conclusion

Annotations sur le web



- La situation change
 - Lecture / écriture
 - communautés, wiki, partage de tag, Web 2.0
 - Amazon, Google, Yahoo!
 - Web sémantique
 - intégration de quelques outils
 - pas vraiment d'application
 - Annotation par l'usage
 - traces
- A suivre

38

Plan



- Lecture active et annotations
- Applications sur le web
- Applications dans l'audiovisuel

39

Documents audiovisuels



- Familial / loisirs
 - Télévision cinéma : films, documentaires, émissions, reportages...
 - Vidéos personnelles (plus ou moins montées)
- Pro
 - Transmission de connaissances
 - professionnel, général, manuels techniques
 - Recherche
 - sémiotique, anthropologie, linguistique...
 - Temps réel
 - vidéo-conférences, surveillance, médecine...
 - ...
- Evolution du médium depuis les frères Lumière

40

AV : documents temporels



- Temporalités natives
 - temps du flux
 - lecture contrôlée par l'outil
 - temps du document
 - 1h pour lire un document d'une heure
- Lecteurs standards
- Pour aller au delà
 - obligation de construire des structures à base d'annotations

41

Systèmes d'Information AV



- Fonctionnalités
 - Indexation
 - Recherche
 - Visualisation résultats
 - Réutilisation
- Exemples
 - TF1
 - YouTube
 - Vidéos personnelles sur son disque
 - ...

42

SIAV et métadonnées



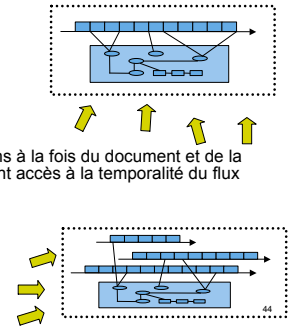
- Annotations
 - informations attachées à un fragment spatio-temporel du flux
 - pour
 - décrire, rechercher
 - présenter
 - réutiliser
- Modèles pour des annotations AV
 - fragment + informations
 - Exemples
 - MPEG7
 - Annodex

43

Le cadre des hypervidéos



- Documents AV annotés
 - document AV
 - plus structure d'annotation
- Vue →
 - « façon de présenter » un document audiovisuel annoté
- Hypervideo
 - Vue qui utilise des informations à la fois du document et de la structure d'annotation, donnant accès à la temporalité du flux
 - Hyper → accès direct
 - Video → flux temporel
- Généralisation pour plusieurs documents AV



44

Lecture active de documents AV



- Activité impliquant
 - Annotation,
 - un ou plusieurs documents temporels
 - Structuration
 - les annotations
 - les contraintes sur les annotations (schéma)
 - Visualisation
 - documents et annotations sous la forme d'hypervideo
 - Echange
 - pour pouvoir partager ses annotations, analyses

45

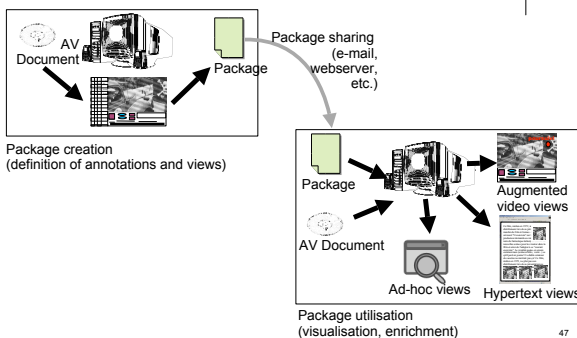
Framework / outil : Advene



- Annotate Digital Video, Exchange on the Net
- <http://liris.cnrs.fr/advene> (LIRIS, 2002-...)
- Echange de recueils indépendants du flux
 - Schémas
 - Types d'annotations et de relations
 - Annotations et relations
 - Vues
 - Statiques
 - Dynamiques
 - Ad-hoc
 - Requêtes
 - Ressources supplémentaires

46

Advene : principes

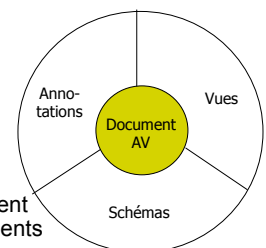


47

Lecture active dans Advene



- Créer / utiliser
 - schémas
 - annotations et relations
 - vues
- Plusieurs possibilités de lecture active
 - plusieurs catégories d'utilisateurs qui réutilisent plus ou moins des éléments d'hypervidéos



48

Lecture active : navigation

Lecture active : annotation

50

Lecture active : structuration

51

Lecture active : création d'hypervidéos

Outils pour la lecture active

- Outils pour bonne gestion des différents éléments
- Médium temporel
 - Collisions de temporalité
 - flux / lecture active
 - Nécessité de spatialisation
 - ...
- Projets en cours pour Advens
 - Cinélab
 - IncorporActions

53

Démonstration

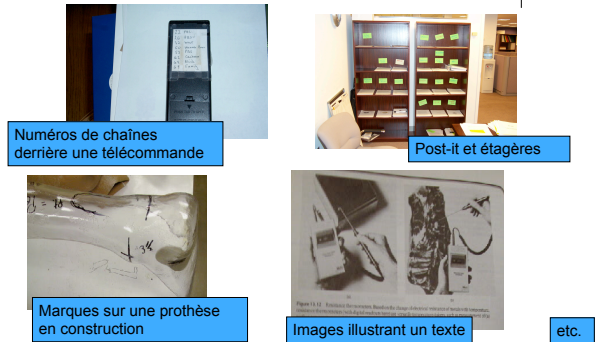
54

Conclusion générale

- Cours introductif à l'annotation documentaire pour la lecture active et/ou dans le cadre du web sémantique
 - documents complémentaires (notamment cours des années précédentes) disponibles
- Les traces de lecture de documents (annotations) sont partout et forment le support d'une « communauté des documents » qui croise et résulte des activités humaines
- Ce n'est pas fini
 - annoter le monde
 - documentariser le monde

55

Des annotations partout !



Annexes

- Bush et l'hypertexte
- CREAM / Ontomat

57

Aux origines de l'hypertexte

- V. Bush, As we may think, The Atlantic Monthly, 1945.
- Principe général
 - Photos, facsimilés, ... : milliards de documents.
 - Traitements : calculs par des machines très rapides, simple sélection (pensée non créative).
 - Les progrès de l'informatique et du calcul ne pourront pas grand-chose pour la Science : l'intuition et la créativité ne sont pas dans la machine.

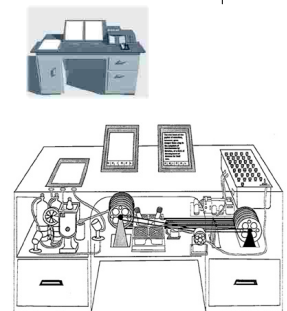
Bush dans le texte

- ...Our ineptitude in getting at the record is largely caused by the artificiality of systems of indexing. When data of any sort are placed in storage, they are filed alphabetically or numerically, and information is found (when it is) by tracing it down from subclass to subclass. It can be in only one place, unless duplicates are used; one has to have rules as to which path will locate it, and the rules are cumbersome. **Having found one item, moreover, one has to emerge from the system and re-enter on a new path...**
- The human mind does not work that way. It operates by association. With one item in its grasp, it snaps instantly to the next that is suggested by the association of thoughts [...] the speed of action, the intricacy of trails, the detail of mental pictures, is awe-inspiring beyond all else in nature.
- Man cannot hope fully to duplicate this mental process artificially, but he certainly ought to be able to learn from it. [...] The first idea, however, to be drawn from the analogy concerns selection. **Selection by association, rather than by indexing, may yet be mechanized.**

59

Memex

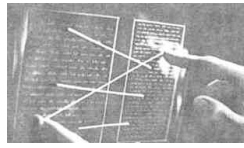
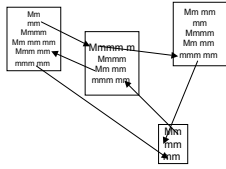
- Un bureau à microfilms : livres, communications, ...
- Lecture en utilisant l'indexation et la sélection : livre, page.
- Possibilité de voir plusieurs pages en même temps, de les rappeler, etc.
- Possibilité d'ajouter des pages (scanner)



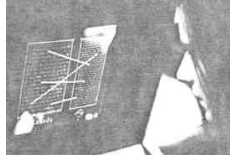
60

Memex

- Annotation dans les marges
- Construction de trails (chaines de liens)



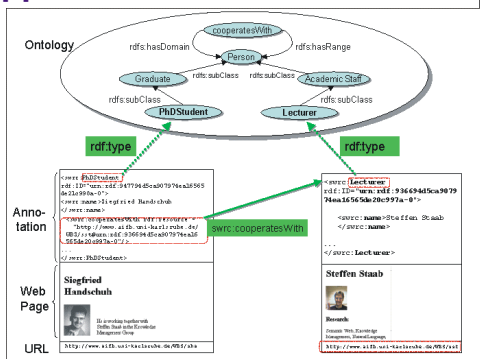
Theodor H. Nelson, "As We Will Think,"
Proceedings of Online 72 Conference, England, 1972.



Bush dans le texte

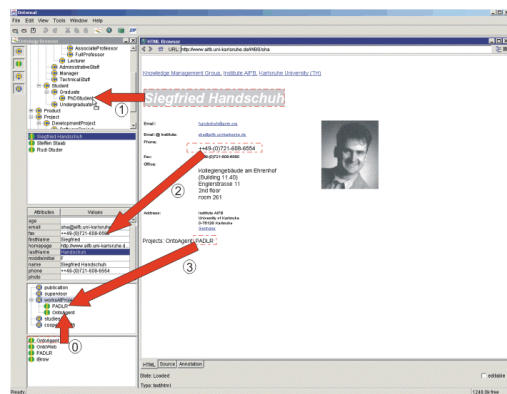
- All this is conventional [...] It affords an immediate step, however, to **associative indexing**, the basic idea of which is a provision whereby any item may be caused at will to select immediately and automatically another. This is the essential feature of the memex. **The process of tying two items together is the important thing.**
- When the user is building a **trail**, he names it, inserts the name in his code book, and taps it out on his keyboard [...] the items are permanently joined. [...]
- Thereafter, at any time, when one of these items is in view, the other can be instantly recalled merely by tapping a button below the corresponding code space. Moreover, when numerous items have been thus joined together to form a trail, they can be reviewed in turn, rapidly or slowly, by deflecting a lever like that used for turning the pages of a book. It is exactly as though the physical items had been gathered together to form a new book. It is more than this, for any item can be joined into numerous trails.

Approche CREAM



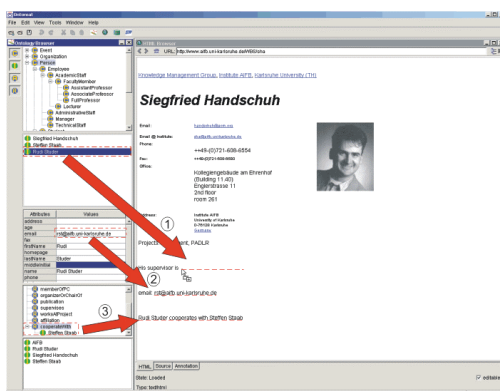
63

Ontomat



64

Ontomat



65

66