

Le Conservatoire numérique : enjeux et perspectives de la numérisation documentaire

The Conservatoire numérique : a digital library in science and technology history

Rodrigo Almeida, Claire Bernardoni, Jean-Paul Cheung, Françoise Chevalier, Mireille Le Van Ho, Pierre Cubaud et Karine Raczynski



Édition électronique

URL : <http://journals.openedition.org/dht/116>

ISSN : 1775-4194

Éditeur :

Centre d'histoire des techniques et de l'environnement du Cnam (CDHTE-Cnam), Société des élèves du CDHTE-Cnam

Édition imprimée

Date de publication : 1 décembre 2009

Pagination : 61-72

ISBN : 978-2-9530779-4-0

ISSN : 0417-8726

Référence électronique

Rodrigo Almeida, Claire Bernardoni, Jean-Paul Cheung, Françoise Chevalier, Mireille Le Van Ho, Pierre Cubaud et Karine Raczynski, « Le Conservatoire numérique : enjeux et perspectives de la numérisation documentaire », *Documents pour l'histoire des techniques* [En ligne], 18 | 2^e semestre 2009, mis en ligne le 24 septembre 2010, consulté le 19 avril 2019. URL : <http://journals.openedition.org/dht/116>

Le Conservatoire numérique : enjeux et perspectives de la numérisation documentaire

Rodrigo Almeida, Claire Bernardoni, Jean-Paul Cheung,
Françoise Chevalier, Mireille Le Van Ho
Bibliothèque du Cnam

Pierre Cubaud
Laboratoire CEDRIC / Cnam

Karine Raczynski
Centre de documentation du Musée des arts et métiers

RÉSUMÉ

Le Conservatoire numérique (<http://cnum.cnam.fr>) est la bibliothèque patrimoniale en ligne du Conservatoire national des arts et métiers, produite par trois de ses entités : la Bibliothèque Centrale du Conservatoire national des arts et métiers, le Centre d'histoire des techniques et de l'environnement (CDHTE) et l'équipe « Interactivité pour lire et jouer » du Centre d'études et de recherche en informatique (CEDRIC). Riche aujourd'hui de plus de 500 000 pages, le Cnum, qui se consacre à la numérisation en histoire des sciences et des techniques, permet la consultation de périodiques savants, de catalogues de constructeurs, d'ouvrages encyclopédiques, de monographies... dont l'ensemble constitue des corpus thématiques autour de l'histoire de la construction, des expositions universelles, des théâtres des machines, de l'électricité, etc. L'installation d'une passerelle OAI permet au Cnum d'une part d'être accessible depuis d'autres bibliothèques numériques, comme Gallica, et d'autre part d'interroger diverses ressources numériques spécialisées. Le Cnum est également un laboratoire de recherche informatique, qui tend à mettre en relation, à travers le développement de la 3D le texte et l'ouvrage, puis l'ouvrage et l'objet.

Résumés et mots clés en anglais sont regroupés en fin de volume, accompagnés des mots clés français

Le Conservatoire numérique des arts et métiers¹ (Cnum) est une bibliothèque numérique consacrée à l'histoire des sciences et des techniques, constituée à partir du fonds ancien de la bibliothèque du Conservatoire national des arts et métiers, riche

de 8 000 ouvrages antérieurs à 1800 et 20 000 ouvrages et périodiques du XIX^e siècle. Le Cnum, qui fêtera ses 10 ans d'existence en 2010, est un projet commun de la Bibliothèque, du Centre d'histoire des techniques et de l'environnement (CDHTE) et de l'équipe « Interactivité pour lire et jouer » du Centre d'études et de recherche en informatique (CEDRIC) du Conservatoire national des arts et métiers (Cnam). Une collection d'ouvrages et de périodiques francophones, significatifs du fonds patrimonial de la bibliothèque du Cnam, est mise en accès libre et gratuit sur Internet. Le Conservatoire numérique s'adresse aux chercheurs et enseignants en histoire des sciences et des techniques, en épistémologie, en didactique, en offrant à la fois des textes et une documentation iconographique spécifiques. Le Cnum est aussi un

¹ <http://cnum.cnam.fr/> Voir *Le Cnam, nouveau pôle associé pour la numérisation partagée dans le domaine des sciences et techniques*, [//www.bnf.fr/pages/infopro/cooperation/cnam.htm](http://www.bnf.fr/pages/infopro/cooperation/cnam.htm) ; Geneviève Deblock, « La Bibliothèque numérique du Conservatoire national des arts et métiers », *L'Archéologie industrielle en France*, juin 2000, n° 36, pp. 28-33 ; Geneviève Deblock, Brigitte Rozet, Pierre Cubaud, « Le Conservatoire numérique des arts et métiers. Une création partenariale », *Bulletin des bibliothèques de France*, tome 46, n° 4, 2001, pp. 43-48. <http://bbf.enssib.fr/consulter/bbf-2001-04-0043-005>.

Le Conservatoire numérique

outil de vulgarisation scientifique qui remplit l'une des missions fondamentales du Conservatoire : la diffusion du savoir et la reconnaissance du patrimoine scientifique et technique francophone. Le Cnum est reconnu Programme Pluriformation pour les années 2007-2010 par le Ministère de la recherche.

Dans le cadre de cette présentation, nous voudrions répondre à plusieurs questions que soulève cette réalisation. Quels sont aujourd'hui la raison d'être et l'avenir de bibliothèques numériques patrimoniales très spécialisées, comme le Conservatoire numérique en regard des entreprises de numérisation massive en mode texte que sont Gallica ou Google books ? Dans quels domaines, sur quels supports précis une politique éditoriale différente peut-elle se définir aujourd'hui ? Quelle est l'appropriation réelle des bibliothèques numériques par les chercheurs et quelles améliorations dans les modes de consultation et de « déambulation » faut-il encore élaborer pour en faire de véritables outils de recherche² ?

La genèse du Conservatoire numérique

Le site du Cnum a été ouvert en février 2000 avec une première campagne de numérisation sur l'électricité et le magnétisme au XVIII^e siècle, à l'occasion d'un colloque sur le bicentenaire de la pile de Volta. Le principe d'associer chercheurs et bibliothécaires à l'établissement des corpus de numérisation est un mode de fonctionnement auquel il ne sera jamais dérogé. En 2007 et 2008, Le CDHTE, la bibliothèque et l'Institut national d'histoire de l'art ont collaboré à la constitution d'un corpus de 600 ouvrages sur la construction qui seront décrits dans la *Bibliographie du livre d'architecture français*, pilotée par l'Institut national d'histoire de l'art (INHA). A cette occasion, la bibliothèque a numérisé plus d'une centaine d'ouvrages de construction et de fortification, dont la rarissime édition de 1623 de la *Manière de bastir pour toutes sortes de personnes* de Pierre le Muet, présentée en juin 2008, en mode texte et images avec un glossaire technique, à l'occasion du Premier congrès francophone d'histoire de la construction. Pour autant, la numérisation en mode texte d'ouvrages antérieurs à 1800, pour lesquels les résultats de l'océrisation sont encore de très médiocre qualité, reste un problème entier, à moins de corriger ou ressaisir à la main l'intégralité du texte. La numérisation en mode texte de

² Sur ces questionnements, voir Isabelle Westeel, Martine Aubry édcs., *La numérisation des textes et des images : techniques et réalisations*, actes des journées d'études organisées à la Maison de la recherche, Université Charles-de-Gaulle, les 16 et 17 janvier 2003, Villeneuve d'Ascq, Éd. du Conseil scientifique de l'Université Charles de Gaulle-Lille 3, 2003.

corpus anciens pose aussi les questions de la restitution des graphies anciennes, de l'établissement de glossaires, toutes choses que ni Gallica ni Google books n'offrent à ce jour, à savoir un travail scientifique sur le texte lui-même. De telles entreprises ignorent également tous les documents de grand format au-delà du 4^e : les volumes de planches de grand format, dépliantes le plus souvent, n'ont pas vocation à figurer dans des projets de numérisation de masse. La qualité de restitution de l'illustration technique, avec une acquisition en haute définition, est au cœur des préoccupations du Cnum depuis ses débuts et doit permettre des recherches et une analyse précise de ce corpus iconographique, de plus en plus prisé, à partir de la bibliothèque numérique elle-même, étude aujourd'hui impossible à mener à partir de la numérisation proposée par Google-books. L'étude de l'illustration technique ne peut se faire qu'à partir de corpus scientifiquement élaborés, sur des thématiques précises et des niveaux d'acquisition élevés, en dehors de trains de numérisation de masse privilégiant les petits formats et les structures éditoriales simples. Dans cette perspective, il convient de signaler dans la collection du Cnum, riche de 1036 volumes, dont 709 téléchargeables en PDF et plus de 500 000 pages en ligne à ce jour :

- les premiers recueils imprimés de dessins de machines, appelés « théâtres des machines » richement illustrés de gravures sur bois tout au long du XVI^e siècle puis de gravures sur cuivre au début du XVII^e siècle : machines de guerre (fig. 1), machines hydrauliques, moulins, fontaines. Toutes ces représentations se caractérisent par une abondance de détails et un goût très baroque pour l'extraordinaire et le complexe comme *Le Diverse et artificieuse machine* d'Agosto Ramelli (1588) (fig. 2) ou *Les Raisons des forces mouvantes* de Salomon de Caus (1615) ;
- des monographies scientifiques et techniques comme la *Description des arts et métiers* (fig. 3) ou *Les grandes usines* de Julien Turgan (fig. 4) ;
- les grandes revues scientifiques du XIX^e siècle, essentielles pour la recherche comme *les Annales du Cnam* (1861-1933), *La Lumière électrique* (1879-1894 ; 1908-1916), *les Mémoires de la Société des ingénieurs civils* (1848-1965) à partir de la collection de l'Écomusée du Creusot, ou *Le Génie industriel*. Vient d'être mis en ligne le plus grand périodique sur les innovations industrielles du début du XIX^e siècle jusque dans les années 1940, le *Bulletin de la Société d'encouragement pour l'industrie nationale* (1801-1943), avec accès au sommaire de tous les fascicules du bulletin. Une revue de vulgarisation scientifique comme *La Nature* (1873-1905) (fig. 5), très richement illustrée de gravures spectaculaires est le best-seller

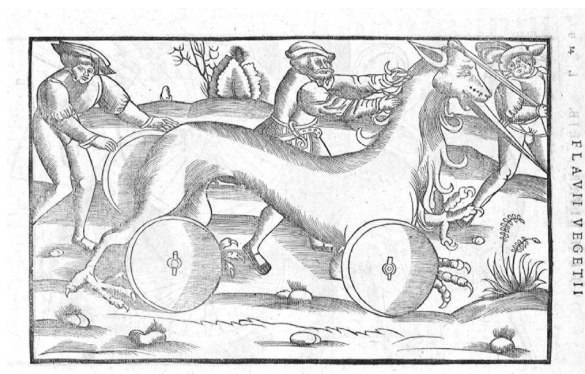


fig. 1 - De Re militari libri quatuor, Végèce, Parisiis, ex officina C. Wecheli, 1535.

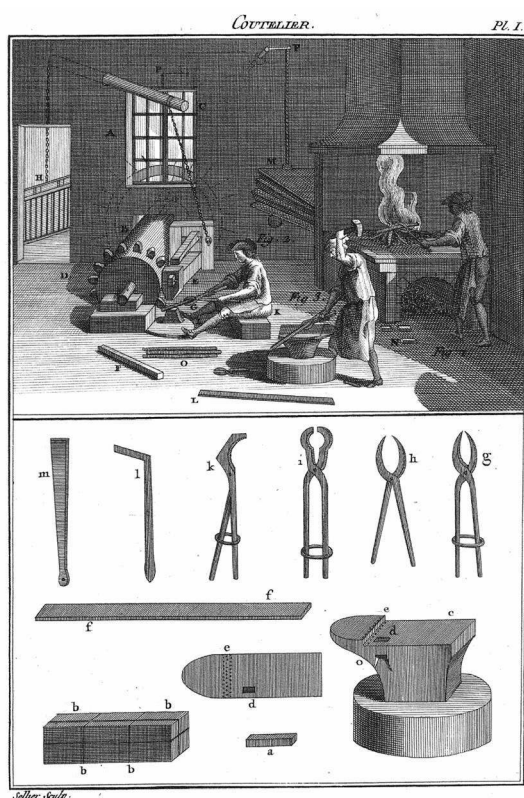


fig. 3 - Descriptions des arts et métiers, faites ou approuvées par Messieurs de l'Académie royale des sciences de Paris/ Académie des sciences (France), éd. Bertrand, Jean-Elie, éd. - A Neuchâtel : de l'imprimerie de la Société typographique, 1771-1783 ; à Paris, chez Calixte Volland, an VII [1798].

du Cnum et a fait l'objet d'une étude des usages de consultation en 2008³.

En 2005, le Cnam et la Bibliothèque nationale de France ont signé une convention de coopération documentaire dans le cadre du réseau des pôles associés. Les deux établissements se sont concertés pour établir un corpus de numérisation à partir de leurs

3 //cnum.cnam.fr/RUB/usages_cnum.pdf

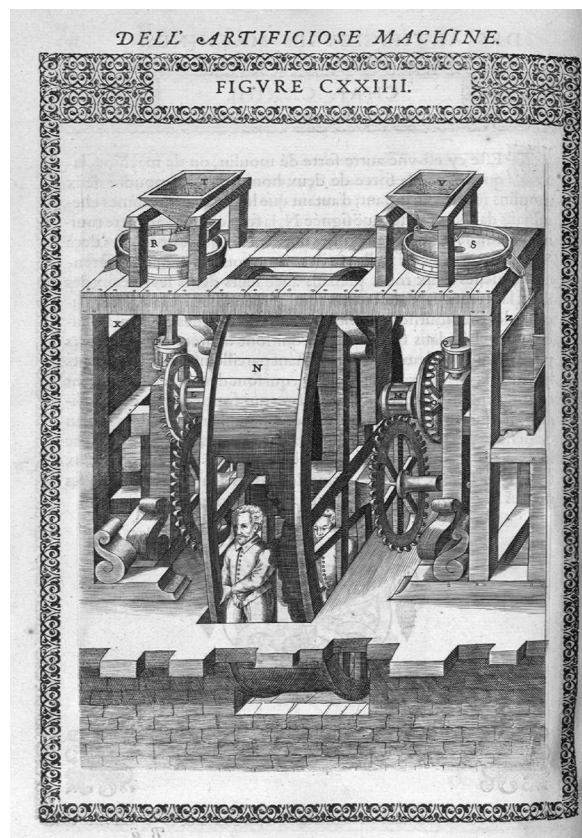


fig. 2 - Le diverse et artificiose machine/Agostino Ramelli, Parisi, in casa dell'autore, 1588.

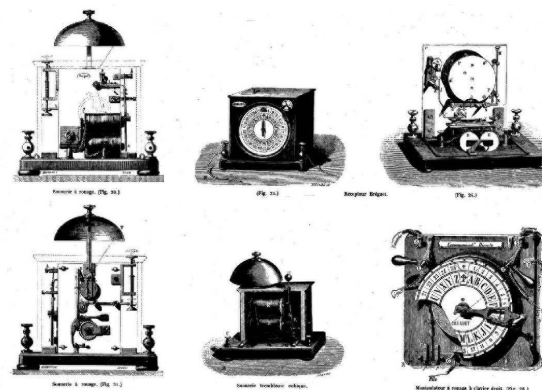


fig. 4 - Les grandes usines de France/Turgan, Paris : Calmann-Levy, 1878.

collections respectives sur le thème des Expositions universelles du XIX^e siècle : rapports officiels des jurys internationaux, rapports publiés par la France sur les nouvelles inventions, congrès, revues, documents préparatoires aux expositions, catalogues... (fig. 6) Les collections de la bibliothèque du Cnam recensent plus de 1 000 monographies et titres de périodiques sur ce sujet qui constitue l'une des principales sources historiques sur l'innovation industrielle et technique de la période contemporaine. Les compléments

que le poids supporté, par chaque arbalétrier, sera de 500 000 kilogrammes, mais à la condition expresse que la pression se répartisse également sur chacun d'eux; car, s'il en était autrement, un arbalétrier arriverait à porter beaucoup plus que les autres, ce qui pourrait devenir dangereux.

Pour que cette égale répartition soit obtenue, il est essentiel que tous les arbalétriers soient exactement montés et qu'ils portent de la même manière sur leurs fondations; il faut, pour nous servir d'une comparaison familière, que la Tour soit comme une

table dont les quatre pieds seraient bien calés. A cet effet, elle est construite de manière à permettre un réglage parfait de ses arbalétriers. Ceux-ci portent, à leur extrémité inférieure, une pièce d'acier en forme de chapeau qui pénètre dans l'intérieur du sabot de fondation et s'appuie, par ses bords, sur le contour de ce sabot (fig. 1). C'est entre les bords du chapeau et le sabot que l'on vient interposer les cales en fer destinées à régler la position exacte de l'arbalétrier.

Pour faire cette opération, on introduit dans l'in-

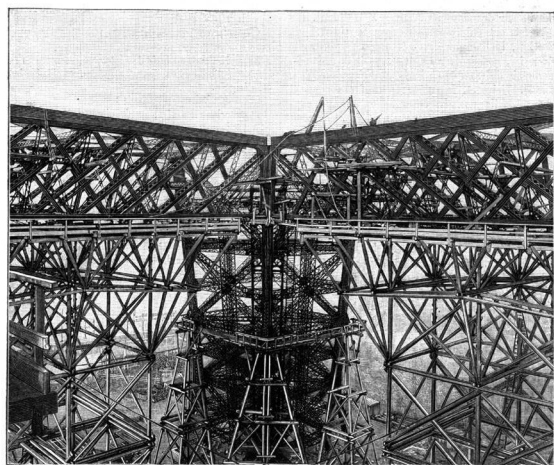


Fig. 5. — Etat des travaux de la Tour Eiffel en janvier 1888. Aspect d'un montant de la Tour à sa partie supérieure. (D'après une photographie de M. Durandelle.)

tableau du sabot, par l'ouverture carrée ménagée dans l'une de ses parois, un vérin ou presse hydraulique d'une grande puissance dont le piston, en s'élevant, vient pousser le fond du chapeau en acier et, par cela même, soulever l'arbalétrier, ce qui permet d'augmenter ou de diminuer le nombre des cales de réglage.

Le vérin se compose d'un piston de 0^m,450 de diamètre qui se meut dans un cylindre de 95 millimètres d'épaisseur correspondant à un diamètre extérieur de 0^m,620 (fig. 2). Piston et cylindre sont en acier forgé. L'eau pénètre dans le fond du cy-

lindre par un tuyau de 6 millimètres de diamètre. Cette eau est comprimée par une pompe foulante que des hommes actionnent au moyen d'un levier. Le poids normal qu'un vérin est capable de soulever est de 800 000 kilogrammes. Chaque vérin a été essayé avant de sortir des ateliers des constructeurs : MM. Vallot, Badois et C^o, à une pression de 600 atmosphères, qui correspond à un poids de 900 000 kilogrammes environ.

Nous terminerons cette notice en reproduisant une photographie donnant l'aspect actuel d'un des montants de la Tour vu de la partie supérieure du

fig. 5 - La Nature. Revue des sciences et de leurs applications aux arts et à l'industrie. Suivi de : Bulletin météorologique de La Nature, Boîte aux lettres, Nouvelles scientifiques/Tissandier, Gaston, réd. en chef ; Parville, Henri de, dir., Paris, Masson, 1888.

apportés par les collections de la BnF ajoutés aux facilités de navigation liées à la mise en place d'un entrepôt OAI font de cette collection numérique une source exceptionnelle qui sera étudiée dans le cadre du colloque : Les expositions universelles en France au XIX^e siècle. Techniques, publics, patrimoine, organisé par le CDHTE en 2010⁴.

En 2008, le Conservatoire numérique et le Musée des arts et métiers ont publié une partie de la collection des cent quatre-vingts catalogues de constructeurs⁵ du fonds Alain Brioux (fig. 7), libraire, expert et spécialiste d'instruments scientifiques, dont la bibliothèque personnelle a été acquise en

4 Anne-Marie Blanchenay, Une coopération autour des expositions universelles : http://chroniques.bnf.fr/numero_courant/cooperation/expos_universelles.htm

5 On trouve fréquemment le terme « catalogues de fabricants » dans les catalogues de vente des libraires ou chez les collectionneurs. Nous avons retenu l'expression plus appropriée de « catalogues de constructeurs ».



Phot. L. Baschet

TOUR DU MONDE
(Champ de Mars)

fig. 6 - Exposition universelle internationale de 1900 à Paris. Rapport général administratif et technique/ France. Ministère du commerce, de l'industrie, des postes et des télégraphes Picard, Alfred, dir., Paris, Imprimerie nationale, 1902-1903.

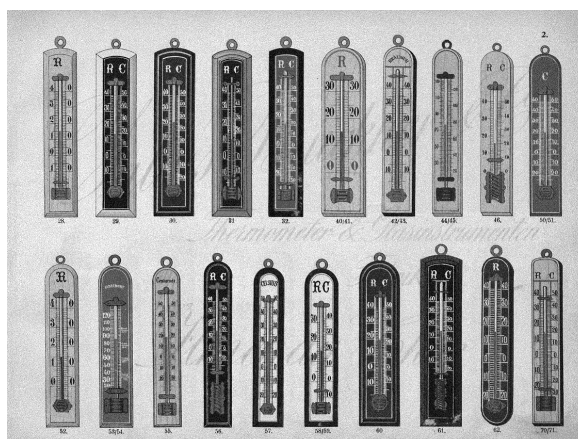


fig. 7. Gravures du prix-courant A.2 de thermomètres et instruments en verre / Brückner, Julius & Co, Ilmenau (Thüringen), Kunstanstalt W. Leopold, [1900].

2004 par le musée⁶; l'intégralité sera consultable fin 2009 sur le Cnum. Ce fonds, incontournable pour les historiens des techniques, est un outil précieux pour les chargés de collection du musée puisqu'il documente les objets concernant les domaines de la communication, de l'énergie, de l'instrumentation scientifique et de la mécanique. Hormis un catalogue datant de 1739 (Van Musschenbroek)⁷, l'ensemble couvre la période 1815-1970. Y figurent

6 Voir la bibliographie reprenant l'ensemble des titres sur le site du Musée des arts et métiers (<http://www.arts-et-metiers.net>), rubrique « Ressources documentaires, bibliographies ».

7 Van Musschenbroek : machines de physique, mathématique, chirurgie.

des constructeurs français (*Adnet, Alvergniat, Brewer, Charrière, Chauvin, Deleuil, Lerebours, Nachet...*)⁸, anglais (*Amadio, Arnold & sons, Beck, Becker...*)⁹ et allemands (*Brückner, Geissler Nachfolger, Schleicher, Zeiss...*)¹⁰. Ces documents se présentent sous forme de catalogues détaillés, notices techniques ou publicitaires. Riches en contenu, abondamment illustrés, ils sont le plus souvent accompagnés de tarifs, parfois de courriers.

La numérisation d'un grand nombre de catalogues sur le Cnum, sous l'intitulé « Catalogues de constructeurs », est l'occasion d'inaugurer une bibliothèque numérique spécialisée pour ce type de documents. La spécificité du Cnum permettant une recherche par mots des tables des matières et des illustrations est un avantage par rapport à un simple mode image.

Le centre de documentation du musée continue d'acquérir régulièrement des catalogues de constructeurs avec le soutien du département scientifique du musée. Pour 2009, il est prévu de mettre en ligne vingt-cinq catalogues, acquis en 1998, provenant de la bibliothèque du photographe Marcel Bovis (1904-1997), et portant sur les appareils photographiques et leurs accessoires ; ce rythme annuel devrait désormais être régulier. Notons que des catalogues numérisés par d'autres institutions et accessibles sur Internet, sont recensés par *The Scientific Instrument Commission*¹¹.

Fin 2010, le Cnum aura doublé son volume avec 650 000 pages en ligne. En 2008, 10 millions de pages ont été consultées sur le Conservatoire numérique et 50 000 fichiers téléchargés.

Les projets de numérisation pour 2010 et les années suivantes concernent deux thématiques pour lesquelles les fonds patrimoniaux de la bibliothèque sont particulièrement riches :

- les technologies de la communication (photographie, phonogramme, microfilm) pour lesquelles un corpus de 200 ouvrages sur les techniques et les procédés de

la photographie à ses débuts a déjà été constitué en collaboration avec le Musée des arts et métiers.

- L'innovation technique dans les moyens de transport de 1880 à 1914.

Les modes de consultation et les usages du Cnum

L'interrogation en mode plein texte s'effectue sur les notices bibliographiques, les tables des matières et des planches qui sont les véritables points d'entrée des documents.

Articulation principale de l'ouvrage, la table des matières offre, en un seul regard, la structure et le contenu du document. Souvent associée et complétée par une table des planches, la consultation d'une table des matières offre un premier panorama de l'ouvrage.

La typologie des tables des matières est variable :

- Les *tables chronologiques* suivent l'ordre des pages et peuvent présenter des constructions hiérarchiques plus ou moins complexes : section, sous-section, chapitre, paragraphe. Autant de segmentations qui organisent la pensée et les propos.
- Les *tables index* proposent des entrées par ordre alphabétique, comme un dictionnaire, qui soulignent ainsi la variété et le grand nombre des termes scientifiques. Si elles offrent un riche vocabulaire, elles ne rendent pas compte de la structure de l'ouvrage.
- Les *tables des matières* des revues suivent l'ordre de parution des numéros toutefois, au fil des années, elles peuvent modifier leur présentation, changer la structure des rubriques.

Ces multiples cas de figure sont analysés et un soin particulier est apporté à leur restitution fidèle dans le Cnum.

Symbolisée par l'icône ci-contre (loupe), la recherche en mode texte est possible uniquement dans les notices bibliographiques, les tables des matières, des planches et des illustrations. Les boutons radio permettent de multiplier et combiner les critères de recherche : possibilité de choisir le type de document (monographie ou périodique), les types de table (matière ou illustration) pour rechercher un ou plusieurs mots, une expression ou une partie seulement d'un mot.

La recherche par début ou fin de mot peut être utilisée quand on recherche des informations sur une thématique précise. Si on s'intéresse aux techniques du son, la recherche début de mot « audio » donne les résultats suivants : audio, audiomètre, audiométrique. La recherche « ortho » propose : orthose, orthoptère, orthochromatique, orthographe, orthopédie, orthopédique, orthographie, orthoscopique, orthodiagraphe...

8 Adnet : instruments de laboratoire ; Alvergniat : verrerie de laboratoire, microbiologie ; Brewer frères : verrerie et instruments de laboratoire ; Charrière : instruments de médecine et de chirurgie ; Chauvin : appareils de contrôle électriques ; Deleuil : instruments de chimie, optique, mathématiques ; Lerebours : instruments d'optique ; Nachet : instruments d'optique, micrographie.

9 Amadio : microscopes et instruments d'optique ; Arnold & sons : instruments de chirurgie ; Beck : microscopes ; Becker : instruments de physique.

10 Brückner : appareils scientifiques, verrerie de laboratoire ; Geissler Nachfolger : instruments de chimie, physique, bactériologie ; Schleicher : filtres pour instruments de laboratoire ; Zeiss : microscopes.

11 Scientific instrument Commission - Online scientific instrument trade catalogues. //www.sic.iuhps.org/refertxt/catalogs.htm

Le Conservatoire numérique

Il est possible de rechercher le suffixe « scope » en fin de mot, recherche qui donne 111 occurrences dans *La Nature*, 16 dans la *Lumière électrique* et 282 au total.

Interroger les tables des illustrations permet de trouver des informations non mentionnées dans les tables des matières. Les termes « acanthoglosse » et « kiwi » sont uniquement présents dans la table des illustrations qui sont d'une grande richesse. La recherche « mots à la suite » permet de retrouver des expressions ou mots composés : Photographie céleste, Tour Eiffel. La recherche « mot à proximité » est plus aléatoire : le mot « vitesse » est dans la même phrase que « lumière » mais ne concernera pas forcément la vitesse de la lumière. Depuis cette interface, il est aussi possible de faire une recherche dans les titres de GALICA. La convention avec la Bibliothèque nationale de France a permis de développer une passerelle de navigation d'après le protocole OAI-*PHM*¹² entre Gallica et le Cnum et depuis 2009 entre le Cnum et Gallica. D'autres bibliothèques numériques peuvent également enrichir automatiquement leurs recherches en ciblant le Cnum par ce protocole.

L'OAI est un protocole d'interopérabilité qui permet l'échange des métadonnées (les notices) entre des bibliothèques numériques. Ces métadonnées contiennent, entre autres, le titre, l'auteur, la date d'édition et les mots-matières de chaque volume numérisé. Elles sont associées à une adresse URL pointant directement sur le site d'origine où le document numérique est consultable. Ces données sont exportées dans un fichier XML au format Dublin Core lors d'une requête sur le serveur OAI.

En octobre 2005, le Cnum a mis en place un serveur OAI, permettant le « moissonnage » régulier des métadonnées de ses collections qui peuvent être intégrées dans les bases des sites « moissonneurs » partenaires dont Gallica. Grâce à ce mécanisme, un utilisateur de Gallica qui fait une recherche en histoire des techniques, voit alors apparaître des résultats provenant, à la fois du catalogue Gallica et des catalogues d'autres bibliothèques « moissonnées » dont le Cnum. Sur la page d'affichage des résultats de GALICA, des liens hypertextes du CNUM redirigent

alors l'utilisateur vers ses propres tables des matières. Par ailleurs, dans le protocole OAI, chaque notice peut être associée à une, voire plusieurs « catégorie » (« SET » dans le jargon OAI). Cette catégorie désigne l'ensemble – qu'il soit basé sur un critère sémantique, éditorial ou arbitraire – auquel chaque volume est lié. En moissonnant le Cnum, il est possible de cibler une ou plusieurs catégories et ainsi, seul un sous-ensemble de la base est exporté. Actuellement, le Cnum dispose de quatre catégories : « Expositions universelles », « Catalogues commerciaux », « Revues » et « Monographies ». En indiquant une de ces catégories au moment du moissonnage, nos partenaires peuvent ainsi suivre les mises à jour de notre fonds en se focalisant sur le domaine ou le type de document qui les intéresse. Par la suite, d'autres catégories plus fines pourront être créées (par exemple la catégorie « Fortifications » ou « Catalogues de microscopes »), en fonction des centres d'intérêts des partenaires.

Depuis janvier 2009, à son tour, le Cnum moissonne et intègre les métadonnées de GALICA. Ainsi, lorsqu'un utilisateur du Cnum ne trouve pas suffisamment de références pertinentes, il a la possibilité d'élargir sa requête en sélectionnant le catalogue Gallica sur la même page d'interrogation. Après avoir lancé sa recherche, il accède directement à la page du document hébergé sur le site partenaire. Actuellement, nous travaillons à l'identification et au repérage précis des catégories du catalogue Gallica qui correspondent aux thématiques du Cnum. Par la suite, ce filtrage permettra d'affiner les notices importées pour que les résultats de recherche correspondent davantage aux exigences de nos lecteurs, dont les pratiques d'interrogation ont fait l'objet d'une première étude en 2008.

Cette étude portait sur les usages des internautes, à partir d'un million de logs¹³. Elle s'appuyait sur deux critères : les logs de connexion disponibles sur le serveur hébergeant le Cnum, sur une période de 23 jours. Dans un deuxième temps, ont été retenus et étudiés les logs se référant au périodique *La Nature, revue des sciences et de leurs applications aux arts et à l'industrie*, publication de vulgarisation scientifique et technique du XIX^e siècle, l'une des plus importantes du fonds du Cnum, et surtout la plus consultée.

¹² Open Archives Initiative's Protocol for Metadata Harvesting. Voir Geneviève Deblock, Pierre Cubaud, « La coopération autour des programmes de numérisation : les projets fondés sur l'OAI, avec le réseau des bibliothèques associées : présentation du projet réalisé avec le Cnam, autour des expositions universelles et le protocole OAI et la bibliothèque numérique du CNAM », 10^e journées des pôles associés. BNF 6 décembre 2006, Paris : //www.bnf.fr/PAGES/infopro/journeespro/pdf/poles_pdf/poles2006_pdf/Deblock.pdf

¹³ Les logs sont les journaux de bord de connexion des internautes, lors de la consultation de bibliothèques numériques. Voir Claire Bernardoni. *Étude des parcours individuels de consultation à partir des logs du Cnum, 2008.* http://cnum.cnam.fr/RUB/usages_cnum.pdf

L'étude s'est d'abord intéressée au point d'entrée des utilisateurs dans le Conservatoire numérique. Un fichier *log* enregistre l'historique des connexions au serveur Web où le site est hébergé. Cet historique indique, pour chaque connexion, la date et l'heure de la connexion ou de l'action, l'adresse IP de l'utilisateur, la page consultée, le *referer* de la requête.

À partir du *referer*, un des éléments composant un *log*, on a pu déterminer la provenance des internautes, qui s'effectue soit à partir d'un moteur de recherche, soit via un accès direct. La date et l'heure déterminent le temps de connexion d'un internaute sur la bibliothèque numérique, ainsi que le temps passé, par exemple, lors de la consultation d'un document. Le fait qu'un *log* identifie la page consultée ou demandée permet, lors de son analyse, de savoir quels sont les documents ou pages les plus visualisés. Dans leur structuration, les logs comportent la présence d'équations de recherches, qui ont révélé les termes ou thèmes retenus et choisis par les internautes.

Une étude sur le type de recherches des internautes privilégiant un accès direct au Cnum a été réalisée. Suite à un inventaire des occurrences de recherches, treize thèmes apparaissent et trois sont privilégiés : les transports, les sciences de la vie et la communication. Cet inventaire a révélé une précision du vocabulaire de recherches, indiquant que les internautes maîtrisent parfaitement des termes techniques et le fonds documentaire. Ils ont connaissance du foisonnement des sujets abordés dans *La Nature* et de la richesse et la variété des illustrations.

Pour ceux qui accèdent directement au Conservatoire numérique (en inscrivant l'URL dans le navigateur, ou en enregistrant l'adresse dans leurs favoris ou en la désignant en marque-page), il est possible de caractériser deux types de lecteurs-internautes : des chercheurs spécialisés en histoire des techniques qui utilisent un vocabulaire technique savant et un public moins averti, mais tout aussi curieux. L'analyse du temps de consultation des internautes a pu être réalisée, car les logs comportent des éléments indiquant les dates et heures de chaque action faite sur la bibliothèque numérique. En regroupant les logs indiquant la même adresse IP on peut définir le temps de connexion. À travers cette analyse, on remarque que la majorité parcourt rapidement le Cnum : quelques minutes, voire quelques secondes. Mais en parallèle, il existe un public, moins nombreux, qui interroge de façon extrêmement pointue et qui consulte plus longuement car il a une grande connaissance du domaine d'interrogation et a su

s'approprier le Cnum et ses outils, tant dans le mode de recherche que dans la navigation. Depuis un an, on constate aussi une augmentation significative de la durée des sessions de consultation, indiquant que le Cnum devient un outil de travail pour la recherche.

Au-delà du Web : perspectives en termes technologiques

Le site Cnum et ses collections numériques sont aussi un lieu d'observation et de recherche en matière d'interaction humaine médiatisée. Ces travaux sont menés au sein du groupe de recherche « Interactivité pour lire et jouer » du Centre d'études et de recherche en informatique du Cnam (CEDRIC). Ils sont de deux ordres : d'abord dans l'enrichissement des typologies de documents potentiellement numérisables et, conjointement, dans l'étude d'interfaces de consultations plus performantes que ce que peut offrir le Web d'aujourd'hui. Nous ne décrivons ici que ceux pouvant intéresser plus directement la communauté de la recherche en histoire des techniques.

Le livre comme volume

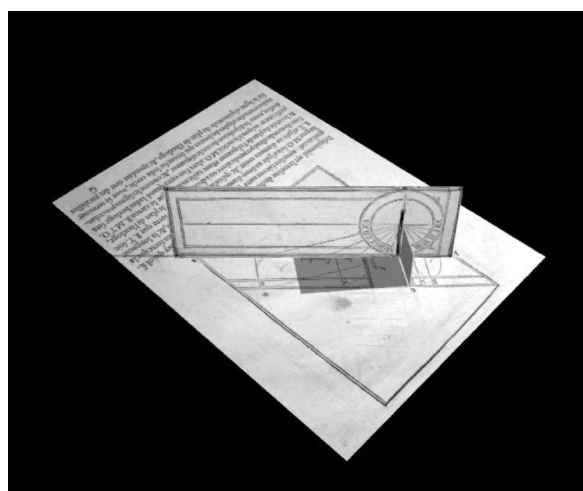
La numérisation de livres est aujourd'hui pratiquée à une échelle quasi industrielle avec des machines de grande capacité, après massicotage ou placement à plat sous une caméra. Dès lors que l'objet livre devient lui-même centre d'intérêt, la mise à plat devient de toute manière insuffisante. On peut citer ainsi les livres d'artistes, avec souvent de nombreux collages superposés, les herbiers, et bien sûr les livres « à systèmes », c'est à dire avec parties mobiles. Les livres à systèmes ont une longue et fascinante histoire, qui remonte aux manuscrits du Moyen Âge. Ils ont d'abord été conçus à des fins didactiques, ou pour réaliser des outils d'aide au calcul. Depuis le XIX^e siècle, les livres à systèmes sont principalement réalisés pour les enfants et utilisent une très grande variété de techniques d'animation : le tableau ci-dessous en récapitule les principaux. À partir des années 1960, les procédés s'industrialisent et une écriture spécifique à ce type de média apparaît¹⁴.

De tels ouvrages procurent une expérience de lecture allant beaucoup plus loin que le simple parcours du texte et des illustrations. Nous prendrons comme exemple un texte de l'ingénieur de la Renaissance Salomon de Caus, consacré à la gnomonique. La page reproduite en figure 1 inclut un modèle en papier de gnomon. Il est composé de deux plats posés initialement l'un sur l'autre. Le plus petit plat, qui représente le stylet du gnomon, doit être levé en premier. Le grand

¹⁴ Peter Haining. *Movable books, an illustrated history*. New English Library, 1979.



fig. 1 - ci-dessus : Salomon de Caus, *La pratique et démonstration des horloges solaires*, Paris, Drouart, 1624 (Bib. du Cnam).
ci-contre : reconstruction 3D avec simulation de l'ombre.



plat se lève ensuite en glissant le stylet dans une échancrure située sur le bord du grand plat. Une fois le modèle construit, le lecteur doit placer le livre au soleil, l'orienter dans la direction Est-Ouest et, grâce à l'ombre du stylet, lire l'heure solaire sur les graduations latérales.

La numérisation de tels textes serait extrêmement utile, car leur manipulation répétée cause des dommages irréversibles. On conçoit toutefois qu'elle soit délicate, dans la mesure où elle impose souvent un démontage des parties mobiles du livre. Une numérisation 3D par le dessus, ne suffit en effet pas pour des montages de papiers perpendiculaires au plan du livre, ni pour résoudre les occlusions de parties cachées. L'exemple évoqué ici est toutefois suffisamment simple pour se prêter à des prises de vues successives des différentes parties du système.

Une fois photographiées les différentes parties du système, il est possible d'en décrire l'organisation sous la forme d'un arbre¹⁵. Chaque partie mobile est nommée et typée (volet, tirette, disque mobile, etc.) et possède une vue avant et arrière. À chaque type sont associés des attributs géométriques spécifiques (position et orientation de l'axe de rotation, direction de rotation, position initiale). Un attribut Parent permet de situer la partie mobile dans la hiérarchie du système, la racine de l'arbre étant la page du livre sur laquelle le système est posé. Un certain nombre d'actions sont associées à la partie mobile, qui sont effectuées par le lecteur et qui peuvent affecter le positionnement d'autres parties mobiles.

Deux options sont envisageables pour la reconstruction : nous pouvons, d'une part, choisir

de reconstruire virtuellement toutes les parties du système (même celles qui sont invisibles, souvent cachées dans l'épaisseur de la page) afin d'obtenir une copie stricte de l'original : tous les mécanismes, leurs articulations et leurs effets seront préservés. Leur intégration se fera grâce à l'utilisation d'un logiciel gérant la physique de l'animation 3D dans l'environnement de consultation. Celui-ci déterminera les positions de chaque partie du système, en temps réel, en fonction des actions de l'utilisateur. Nous pouvons, d'autre part, décider que ces mêmes mécanismes sont d'un intérêt négligeable et que leur numérisation n'apporterait rien de fondamental. Seules les parties terminales (actionneur et actionné) seront donc prises en considération. Cette seconde option permet de s'abstenir de la reconstruction des structures cachées de la chaîne cinématique. D'un point de vue fonctionnel, la perte d'information (liens entre les parties du système) devra être compensée. Ceci pourra être réalisé par la création d'animations, mettant en scène les mouvements simultanés des systèmes liés entre eux. Idéalement, ces séquences permettront de visualiser toutes les configurations relatives du système, en partant de la position initiale jusqu'à la position finale des parties mobiles. À chaque action de l'utilisateur pour modifier la configuration d'un système donné, le fichier de l'animation sera joué, en avant ou en arrière à partir de la position courante, afin de faire correspondre à l'action effectuée (interaction utilisateur) les mouvements des différentes parties mobiles.

Le document et son contexte

Il semble ensuite intéressant d'étudier les modalités de la coexistence, au sein d'un unique environnement de visualisation 3D, de documents textuels et d'objets 3D reconstitués. Une application de tels dispositifs se trouve en histoire des techniques, où l'on voudrait pouvoir associer

¹⁵ Pierre Cubaud, Jérôme Dupire, Alexandre Topol, "Digitization and 3D modeling of movable books. In Proceedings of the 5th ACM/IEEE-CS Joint Conference on Digital Libraries", Denver, CO, USA, June 07 - 11, 2005, JCDL '05. ACM, New York, pp. 244-245.



fig. 2 - Polarimètre Laurent et sa reproduction, dans D. Sidersky. *Polarisation et saccharimétrie*, Paris, Gauthier-Villars, 1900, p.61 (coll. part.).

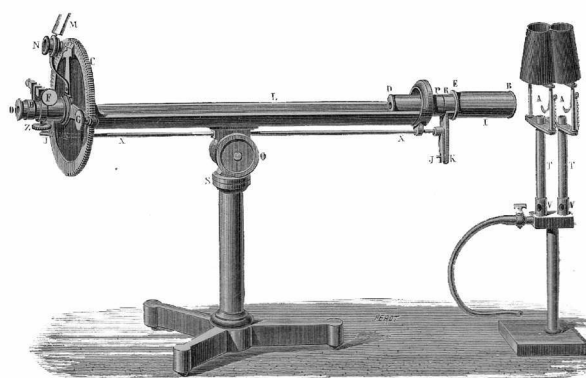
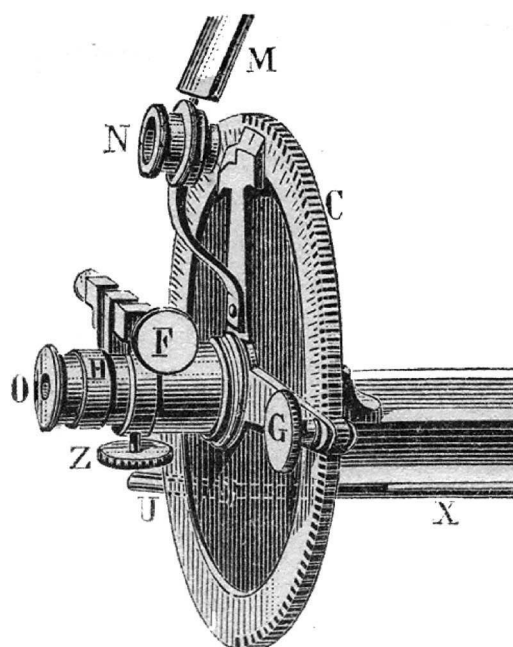


fig. 3 - ci-dessous : Détails du cadran gradué et du porte-oculaire.



des reproductions d'appareils scientifiques, par exemple, et les ouvrages qui les décrivent. Nous nous intéressons en particulier aux situations où ce type d'association contextuelle n'est pas conçue *a priori*, par l'auteur d'un hypermédia, mais *a posteriori* par un lecteur qui a recours à plusieurs sources d'informations qu'il confronte.

Prenons comme exemple une situation que nous jugeons assez typique des usagers du site Conservatoire numérique. Il s'agit ici d'étudier le fonctionnement d'un appareil complexe ancien (ici, un polarimètre) à partir de documents historiques. La polarimétrie est une technique de chimie analytique qui permet le dosage de la teneur d'une solution en substances optiquement actives. On y a recours pour évaluer par exemple la teneur en sucre du jus de betterave. Le polarimètre de Laurent (1878) a été particulièrement employé à cette fin (fig. 2). Cet appareil a été décrit

dans de nombreuses publications de l'époque, dont certaines sont d'ores et déjà numérisées et accessibles sur le Web (au Cnam et à la BnF, par ex.)

La confrontation entre la gravure et l'objet réel est intéressante (fig. 3). D'abord, leur similarité rassure. Un examen de détail révèle ensuite des différences. Le capot du polarimètre est absent de la gravure, sans doute pour exposer les parties plus essentielles qu'il cache. On voit aussi que l'exemplaire conservé est incomplet : il manque le miroir M. Le montage du bouton de réglage F diffère. Une rotation de 120° de celle-ci, possible avec les trois vis de fixation, permettrait de reproduire l'installation décrite sur l'illustration. L'objet a-t-il été mal (re)monté ? S'agit-il de versions différentes de l'appareil, produit pendant plusieurs décennies, ou bien s'agit-il d'un modeste exemple d'adaptation par l'utilisateur, pour une meilleure



fig. 4a - Une table de travail avec les documents originaux.

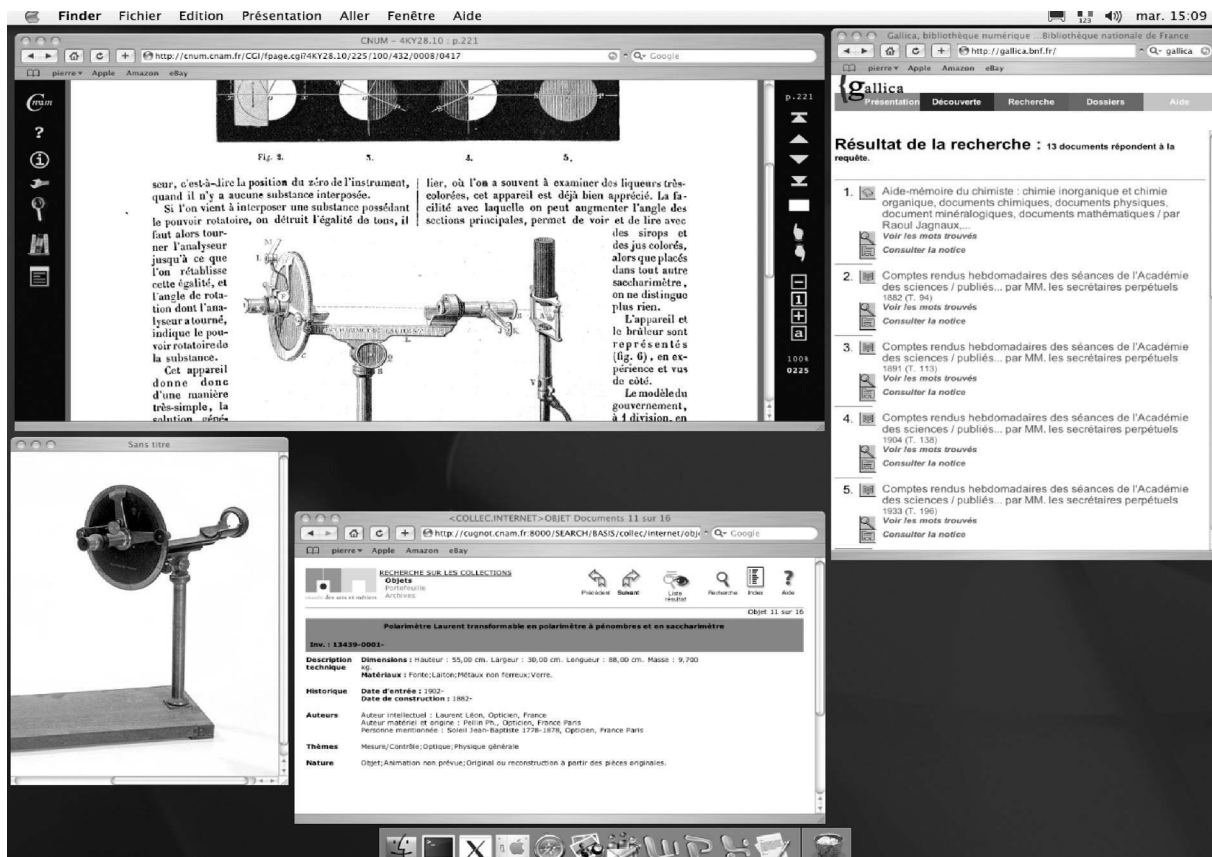


fig 4b - une session de travail avec les documents numérisés sur le Web (sites Gallica/BnF, Conservatoire numérique/Cnam et Musée des arts et métiers)

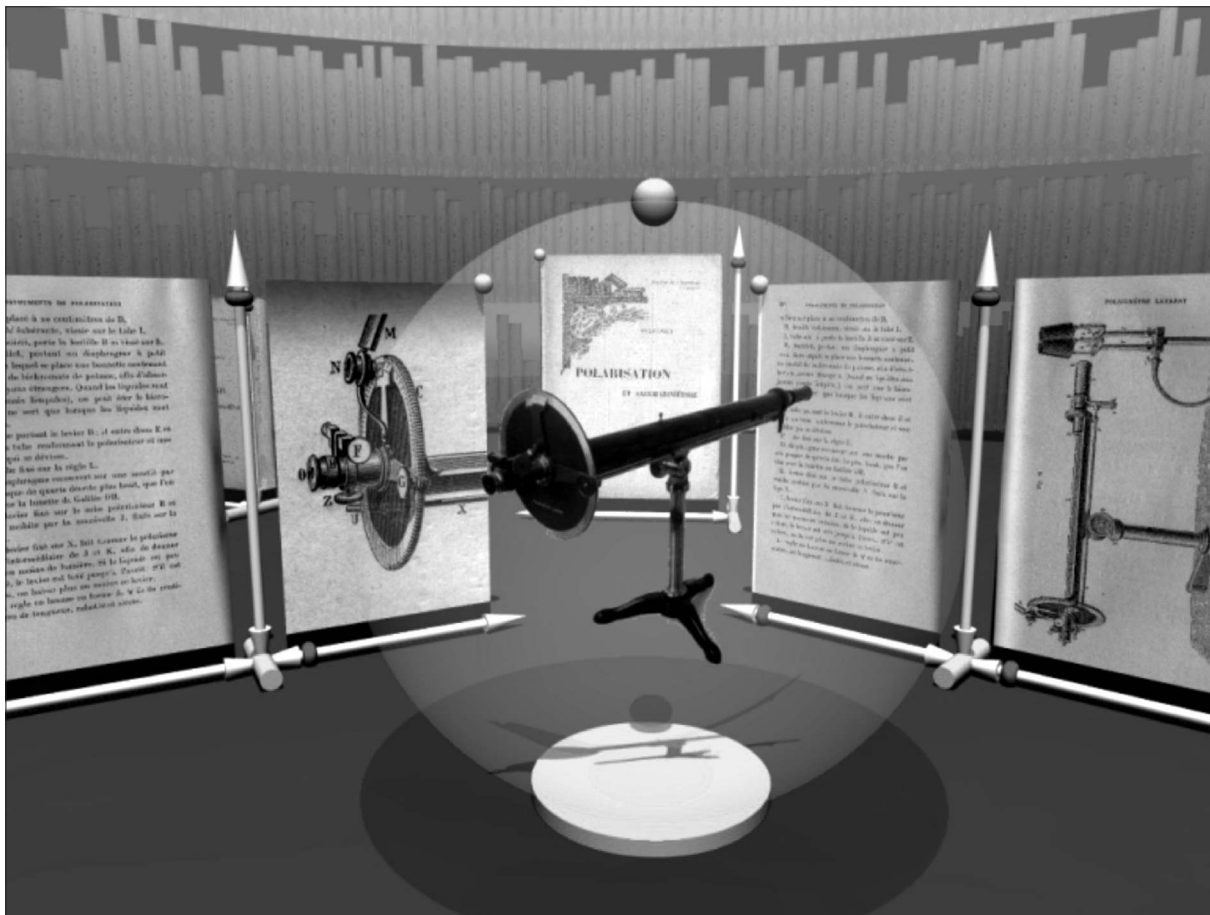


fig. 5 - Simulation d'une session de travail dans le prototype d'environnement 3D.

ergonomie au quotidien, par exemple ? Il n'est pas dans notre propos de répondre à ce type de question, mais plutôt de s'interroger sur les outils informatiques à fournir pour assister ces questionnements lorsqu'ils s'effectuent uniquement à l'aide de documents numérisés.

Les figures 4a et 4b opposent une session de travail effectuée avec les documents et objets originaux, avec celle que peut effectuer aujourd'hui un lecteur du Web. Tout espace de lecture nécessite une organisation, une ergonomie voire une esthétique adaptées pour favoriser un travail efficace. Plus généralement, un espace de travail tel qu'un bureau, réel ou métaphore informatique, est organisé en fonction des aptitudes et préférences de son utilisateur. Ainsi, chacun étale à sa guise, selon ses propres critères, les différents documents sur lesquels il travaille. La mémoire spatiale du lecteur et son sens de l'organisation sont mis à profit et laissés libres sur une surface réduite limitée à la taille de son bureau. La métaphore du bureau apparue avec les interfaces WIMP (Windows Menus Icons Pointer) reprend ce principe et permet également à l'utilisateur d'organiser les différentes fenêtres. Les

documents peuvent y être regroupés, dimensionnés très librement. Cependant, la surface de ces bureaux virtuels est très réduite et ne permet pas d'atteindre une productivité comparable à un bureau classique. Les fenêtres recouvrantes habillant les documents conduisent inévitablement et rapidement à un enchevêtrement de données qui s'oppose clairement à une bonne organisation.

L'environnement reproduit en figure 5 utilise la visualisation 3D temps réel pour faciliter l'organisation des documents¹⁶. Les différences de dimension entre les facsimilés textuels peuvent être compensées par un positionnement en profondeur. Le placement de côté offre une vue en perspective qui, sans offrir une lisibilité totale, permet de conserver une trace des activités de lecture. Les objets 3D sont eux aussi mobiles et manipulables selon un angle de vue arbitraire.

Passer de l'interface textuelle, unidimensionnelle,

¹⁶ Pierre Cubaud, Jérôme Dupire, Alexandre Topol, « A fluid interaction for the document in context », *Proceedings of the 7th ACM/IEEE-CS Joint Conference on Digital Libraries*, Vancouver, BC, Canada, June 18 - 23, 2007, JCDL '07. ACM, New York, pp. 504-504.

Le Conservatoire numérique

au volume, comme les technologies 3D temps réels l'autorisent dès à présent, pourrait permettre d'élargir le flux de communication entre les bibliothèques numériques et leurs usagers. Les interfaces présentées ci-dessus ne sont qu'à l'état de maquette et de prototypes. Il reste encore à comparer, dans un cadre expérimental commun, les procédés proposés par les différentes équipes qui contribuent à ce thème de recherche.

Les techniques de visualisation tridimensionnelle ou les techniques de visualisation, comme la page de vignettes, utilisée lors des contrôles qualité devraient permettre une « déambulation » plus rapide dans les bibliothèques numériques, freinées aujourd'hui dans

leur usage, par la lenteur du chargement des pages et l'absence de « visualisation globale » de l'objet-livre. L'avenir des bibliothèques numériques spécialisées, comme le Conservatoire numérique, passe par la mise en place de ces déambulations rapides, plus encore que par « l'océrisation » massive, si décevante encore pour le livre ancien. Les corpus d'images techniques de grande qualité qu'il met en ligne méritent que les chercheurs en fassent un objet d'études. La plus grande durée des sessions de consultation est sans doute un premier indicateur intéressant d'appropriation des bibliothèques numériques comme outil de recherche, et non plus seulement comme catalogue d'images feuilletées au détour d'une interrogation rapide.