

# **Présentation des archives numériques**

L'Institut Pasteur fut fondé en 1887, avec un statut de fondation privée qu'il a conservé de nos jours. Cependant il faut attendre 1987 et la célébration de son centenaire pour que l'Institut se dote d'un Service des Archives. Avant cette date les archives ont connu des fortunes diverses. Elles ont été soit conservées au Musée, soit détruites, soit dispersées dans les familles des scientifiques.

A son arrivée comme conservateur détaché des Archives Nationale Mme Denise Ogilvie s'est vue confronter à "une mémoire en miette" et dut construire un Service des Archives dans un milieu assez hostile. Douze ans plus tard le Service des Archives de l'Institut Pasteur fonctionne avec 4 personnes et se penche avec une certaine angoisse sur l'archivage des données informatiques.

## **L'informatique et la biologie**

L'informatique est présent partout dans un institut de recherche comme l'Institut Pasteur. Tous les services et les laboratoires disposent des facilités offertes par la bureautique classique. Mais il faut y ajouter des spécificités pour les laboratoires de biologie.

- Il existe à l'Institut Pasteur de nombreuses bases de données. Les plus importantes sont celles des laboratoires Centre Nationaux de Références (CNR), exerçant une mission de surveillance épidémiologique (mission de service public), pour un certain nombre de microorganismes (listériose, rage, grippe..). Chacun de ces centres dispose d'une base de données de plusieurs dizaines voire centaines de milliers de fiches décrivant des souches bactériennes ou virales.

- Le travail de laboratoire utilise de plus en plus de machines ou d'instruments de mesures qui fournissent il y a peu des résultats sous forme informatiques et papier. De plus en plus ils ne fournissent plus que des données numériques. Pour certains la forme papier n'est de toute façon pas scientifiquement exploitable, je pense ici aux séquenceurs d'ADN qui peuvent fournir des rouleaux de plusieurs dizaines de mètres de long de séquence de gènes décodées.

- Le travail scientifique en biologies utilise également énormément la photographie et la microphotographie. L'Institut Pasteur par le biais de son Service Photographique à choisi une stratégie du tout numérique pour la photographie. Mais aussi pour la microphotographie avec la station centrale de microscopie électronique.

## **Les problèmes rencontrés**

Les problèmes sont communs avec la plupart des personnes présentes. Comment sauvegarder ces informations alors qu'il n'existe aucun standards ni méthodes, ni modes de conservations totalement satisfaisant.

Pour ce qui est de la sauvegarde classique, c'est à dire à court terme, le service de l'Informatique Scientifique, représenté ici par O. Perret, possède le matériel et les moyens d'effectuer des sauvegardes régulières des ordinateurs du réseau. Il suffit d'en faire la demande, en effet pour l'instant ce service est basé sur le volontariat. Cette sauvegarde s'effectue sur des cartouches DLT 7000 (340 possibles) capacité en volume 34 terra octets, entre 1/3 et 1/2 utilisé avec 53 postes clients.

On peut voir comme problèmes immédiats :

- L'obsolescence de certains matériels et de certains formats : (disquette 5 pouces 1/2 , logiciel spécifique de certains instruments comme les séquenceurs ou les appareils de microphotographies)

Les risques d'obsolescence des "nouveaux" supports (CD-rom, DVD-rom..). En particulier pour toutes les photographies numérisées.

- Les problèmes de certifications des nouveaux supports la signature électronique et la certification obligatoire du Service des Brevets et inventions. Celui ci, pour l'instant, utilise des cachets de cire pour authentifier les CD-rom

- La sauvegarde de bases de données de laboratoires qui ferme ou qui sont menacés de fermetures. Pour ce cas précis nous avons déjà pris contact avec Mme Leblanc (du centre des Archives Contemporaines de Fontainebleau, CONSTANCE) pour avoir des choses à proposer aux laboratoires (pour l'instant le cas ne s'est pas encore présenté)

- Répondre aux exigences des différentes normes et référentiel de l'assurance qualité qui exige la conservation de certains fichiers informatiques en général sur deux types de supports et à deux endroits différents.

- Éclatement des lieux de production et de stockage des fichiers, les Projets Transversaux de Recherches impliquent plusieurs laboratoires et déjà pour leurs papiers oblige à la conservation de doubles et à un circuit de document compliqué.

### **L'action du Service des Archives**

- Le Service des Archives de l'Institut Pasteur cherche à sensibiliser les laboratoires qui n'auraient pas encore pris conscience du problème. (Néanmoins ils sont de plus en plus conscients des dangers et demandeurs de solutions) Un guide d'archivage interne met en garde les laboratoires en ce qui concerne les problèmes de pérennisation des données numériques. (sans hélas leur fournir de réponses adaptées)

- Le Service des Archives essaie de faire de la veille technologique, et fait partie du groupe PIN

- Le Service des Archives participe aussi avec le Service de l'Informatique Scientifique et l'INRIA à un essai de construction d'un prototype de "cahier de laboratoire augmenté" (Réalité augmentée. Permettant par exemple l'import direct des résultats fournis par un appareil de mesure dans le cahier de laboratoire. En particulier Wendy Mackay souhaite maîtriser l'archivage du document électronique "double" du papier.