

La mise en œuvre du dépôt légal électronique en France

Catherine Lupovici

Directrice du département de la bibliothèque numérique

Direction des services et des réseaux

Bibliothèque nationale de France

Le dépôt légal reçu à la BnF

- ◆ 1537 Imprimés
- ◆ 1648 Estampes dont cartes et plans
- ◆ 1793 Partitions musicales
- ◆ 1925 Photographies et phonogrammes
- ◆ 1975 Vidéogrammes et documents multisupports
- ◆ 1992 Edition électronique sur support
 - Radio et télévision INA (Institut National de l'Audiovisuel)
 - Films cinématographiques CNC (Centre National de la Cinématographie)
- ◆ 2004 Projet de loi sur le droit d'auteur : dépôt légal des publications en ligne (Web)

Chapitre 1er : Objectifs et champ d'application du dépôt légal

◆ Article L131-1

- La collecte et la conservation des documents
- La constitution et la diffusion de bibliographies nationales
- La consultation des documents, sous réserve des secrets protégés par la loi, dans les conditions conformes à la législation sur la propriété intellectuelle et compatibles avec leur conservation

◆ Article L131-2

- Les documents imprimés, graphiques, photographiques, sonores, audiovisuels, multimédias, dès lors qu'ils sont mis à la disposition d'un public
- Les progiciels, les bases de données, les systèmes experts et les autres produits de l'intelligence artificielle, dès lors qu'ils sont mis à la disposition du public par la diffusion d'un support matériel

Projet loi droit d'auteur et droits voisins dans la société de l'information

- ◆ Tout ce qui fait l'objet d'une communication publique en ligne
- ◆ Obligation pour les éditeurs et producteurs de sites
- ◆ Modalités : collecte automatique et/ou dépôt en accord avec les éditeurs et producteurs
- ◆ Exception au droit d'auteur : reproduction pour les nécessités de collecte, de conservation et de consultation sur place
- ◆ Consultation sur place pour des chercheurs accrédités sur des postes individuels dédiés

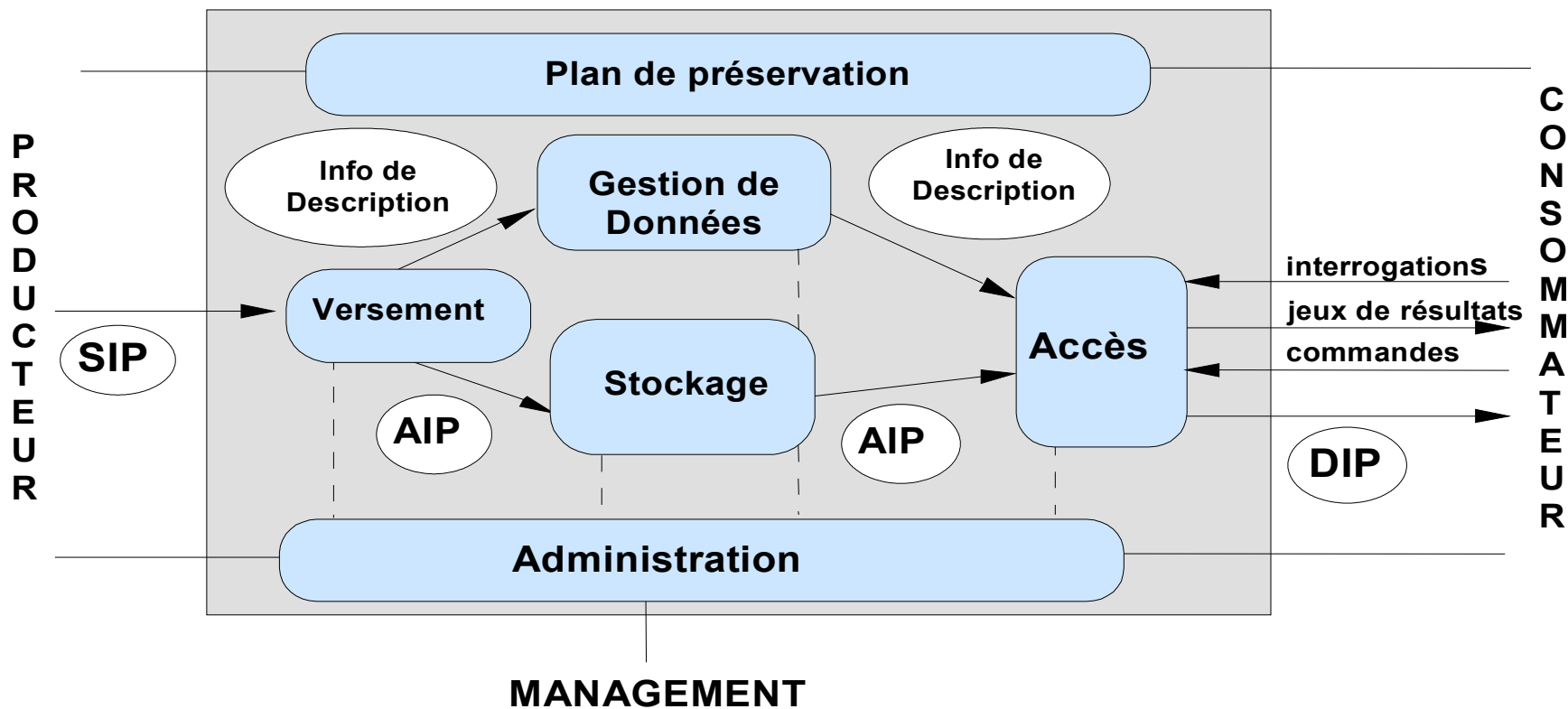
Le dépôt légal : une mission patrimoniale

- ◆ Constitution de la mémoire : exhaustivité de la production ou représentation significative d'un type de publication et de diffusion (le DL n'est qu'un élément de preuve)
- ◆ Pas de concurrence avec la diffusion courante
- ◆ Collecter, conserver, donner accès, sans limite de durée
- ◆ Pour les dépôts électroniques, assurer l'accès aux contenus à long terme : gérer l'obsolescence technologique

L'archive numérique pérenne

- ◆ Modèle conceptuel de référence OAIS (Open Archival Information System)
 - Comité consultatif pour les systèmes de données spatiales
 - NASA, ESA, CNES, NARA
 - ISO 14721:2002
- ◆ Adopté par tous les programmes de préservation et accès à long terme pour les archives, bibliothèques, institutions de recherche
- ◆ Travaux en cours pour une certification
 - explicitation et normalisation des procédures
 - tiers de confiance et conservation partagée

Modèle fonctionnel de référence pour une archive pérenne



Mise en œuvre du dépôt légal de l'Internet à la BnF

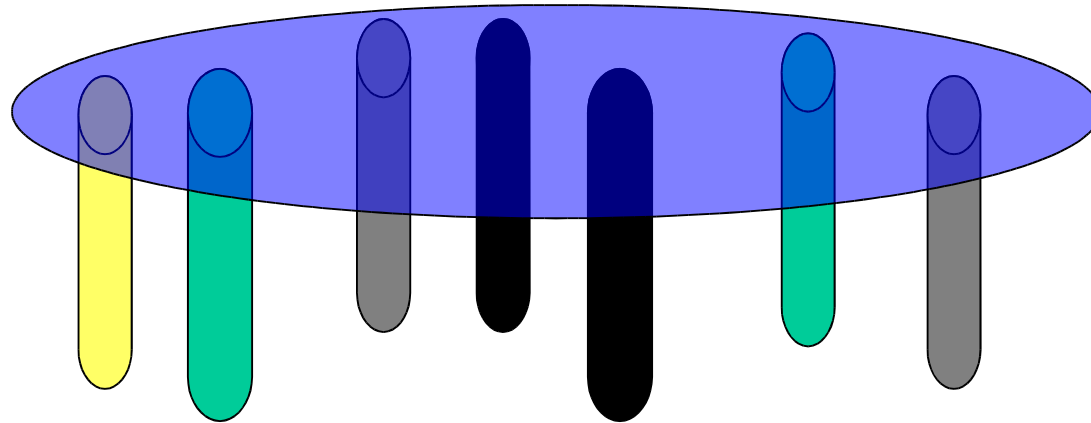
- ◆ Travail préparatoire dès 1998
- ◆ Collecte automatique par robot
 - Thématique élections 2002 et 2004 : 4000 sites, 1 To, 50 millions de fichiers
 - Global sur tout un domaine « national » : travaux de l'IIPC (International Internet Preservation Consortium). Normes, méthodes, outils, interopérabilité des entrepôts
- ◆ Dépôt de portions de sites inaccessibles aux robots
 - Notification du robot, évaluation de l'intérêt par la bibliothèque, contact avec le producteur et définition de modalités de dépôt
 - Outil pour le reformattage en XML des bases de données servant de passerelles documentaires vers les contenus du site
 - Organisation par secteur disciplinaire, en particulier secteur juridique : Journal officiel électronique

Conclusion

- ◆ La mémorisation de l'Internet en est à ses débuts
- ◆ Importance du suivi et de la contribution à l'évolution des technologies de traitement automatique de grandes masses d'informations liées à l'évolution du web
 - pour le recueil des informations à conserver
 - pour l'organisation et la conservation à long terme des informations recueillies
 - pour offrir l'accès à des entrepôts interopérables utilisables par les chercheurs de demain

Le web invisible (deep web)

- ◆ Outre les documents auxquels on accède par une recherche, restent inaccessibles les pages
 - récentes
 - volontairement non liées (à l'intérieur d'un site)
 - dont l'accès a été volontairement protégé par un mot de passe



Modèle fonctionnel intégré défini par la BnF

