La organización de colecciones digitales en la Biblioteca Central de la FCEN –UBA y el Modelo OAIS

Ana Sanllorenti
Martín Williman

1er. Plenario Ciclo 2008 del Foro Permanente de Responsables Informáticos de la APN - ONTI - Subsecretaría de Tecnologías de Gestión - Secretaría de la Gestión Pública
Mayo 2008

Instituciones de la memoria

UNESCO

- Bibliotecas, archivos y museos
- Preservación de materiales que son testigo de la memoria intelectual de nuestra civilización

- Producción, recepción, organización, acceso, difusión de colecciones de objetos digitales

Organización de colecciones de objetos digitales

- Caso de la Biblioteca Central "Luis F. Leloir", de la FCEN – UBA
- Presentación según el Modelo OAIS

Nuestra Biblioteca

- Superficie: 4.000 mts^2
- Horario: Atención de lunes a viernes, de 9hs. a 21hs.
- Usuarios activos: 3.500
- Libros: 70.000 vol.
- Publicaciones periódicas: 7.100 títulos
- Servicios: Préstamos de libros, acceso al documento y préstamos interbibliotecarios, referencia, capacitación de usuarios, etc.
- 400 transacciones diarias

- **Política de preservación**:
- Programa continuo desde 2001
- Construcción de la Sala de Preservación

Área Biblioteca Digital

- **Objetos Nacidos Digitales:**
  - Tesis doctorales: 300 (crecimiento anual de 150)
  - Papers: 1,400 (crecimiento anual de 700)

- **Objetos Digitalizados:**
  - Programas de estudio: 1.000
  - Fotos: 400

Proyecto Biblioteca Digital

- Archivo de imágenes / Programa de Historia de la Facultad
- Repositorio de papers
- Tesis doctorales
- Publicaciones de la Facultad
- Videos

Proyecto Preservación Digital en Bibliotecas Argentinas

- Desarrollo de un estudio exploratorio y un proyecto piloto que contemple una estrategia de trabajo colaborativo entre instituciones con problemática similar

Marco institucional:

- ✔ Dirección Técnica: Biblioteca Central de la Facultad de Ciencias Exactas y Naturales - UBA
- ✔ Subsidio del Programa TICs de la SECTIP
- ✔ 15 Instituciones participantes
- ✔ Lic. Fernando Boro, consultor

Resultados

- **Capacitación** de Bibliotecarios y personal de las bibliotecas e instituciones participantes, en la temática de preservación digital
- **Diagnóstico** sobre el estado de las colecciones digitales y los recursos de las bibliotecas participantes
- Elaboración de **procedimientos y pautas** para realizar **mejores prácticas** de digitalización y de preservación digital
- **Experiencia piloto** de biblioteca digital orientada a la preservación con material de Historia de la FCEyN

- **Grupo de discusión sobre Preservación Digital**

http://espanol.groups.yahoo.com/group/PRESERVACION-DIGITAL-BIBLIOTECA-LELOIR-FCEYN-UBA

OAIS
Open Archival Information System

(ISO Reference Model)

ISO 14721:2003

Modelo de Referencia ISO para Sistemas Abiertos de Archivos de Información

OAIS Open Archival Information System

- Modelo de referencia
- Diferentes tipos de archivos
- Sus procesos de gestión son conocidos (abiertos) por lo que promueve la confiabilidad para la comunidad productora y la que accede

OAIS Open Archival Information System

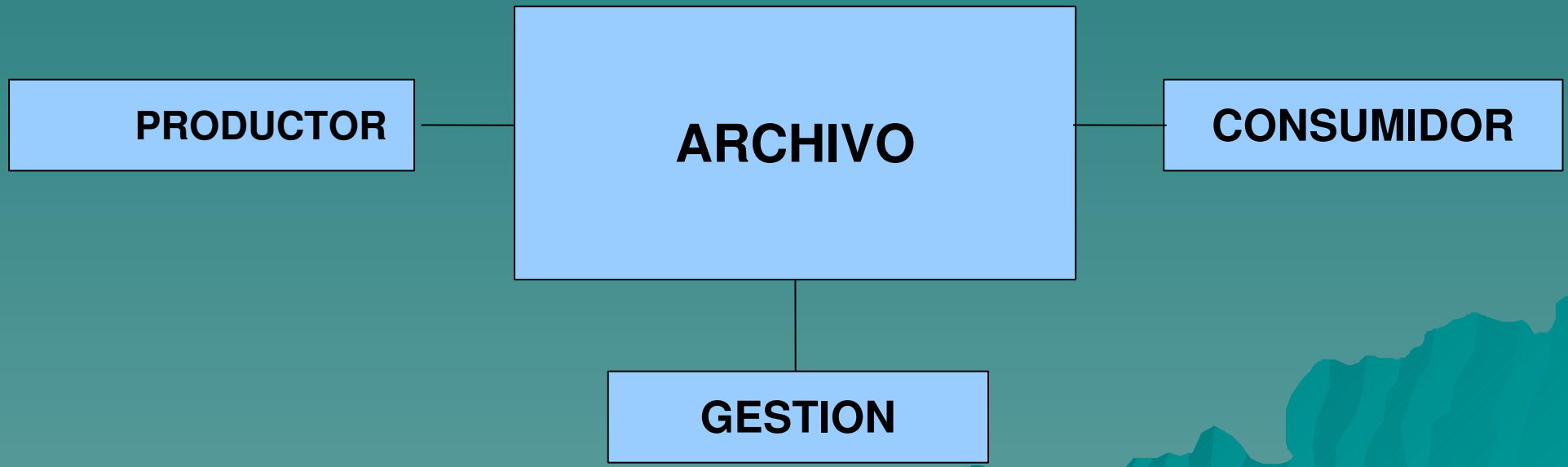
- Ofrece un marco internacional para la definición de entidades, procesos y funciones de los archivos de datos
- Estandariza las relaciones y los componentes de un sistema de archivos
- Establece un vocabulario común
- Facilita la comparación de sistemas de archivos
- Facilita la comprension y aplicación de conceptos para la preservación de información digital a largo plazo

OAIS Open Archival Information System

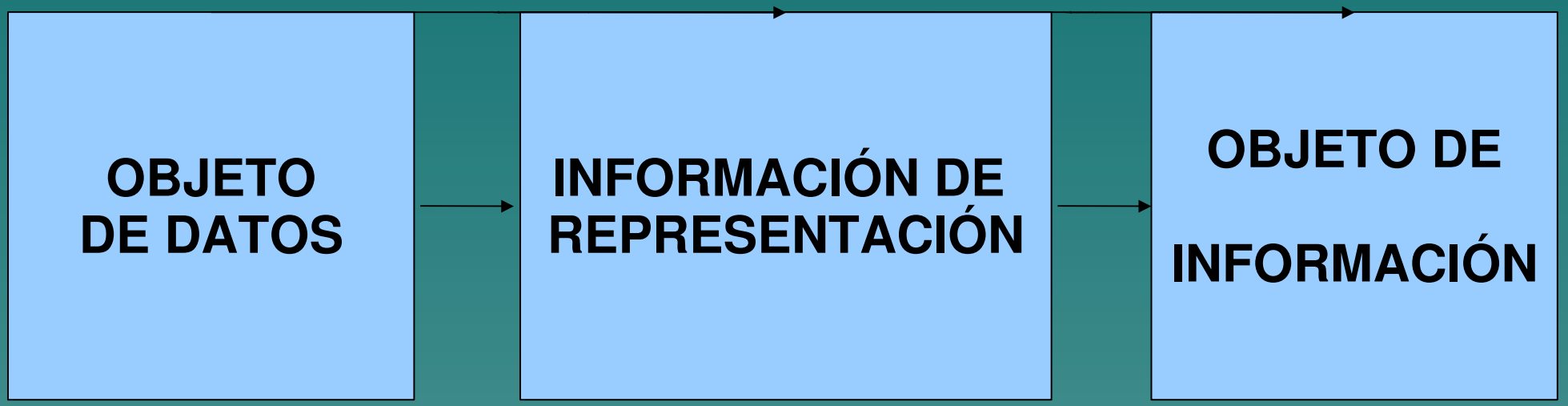
- NO se restringe a los objetos digitales, es aplicable a datos y objetos analógicos
- NO especifica la implementación de un sistema informático

OAIS Open Archival Information System
Entidades del Modelo

- **Productor:** personas o sistemas que envían la información que hay que preservar
- **Gestión:** personas o sistemas que definen y ejecutan las políticas y el alcance de las mismas
- **Consumidor:** personas o sistemas cliente, que interactúan con los servicios OAIS para encontrar y recuperar información preservada



OAIS Open Archival Information System
La información en el Modelo

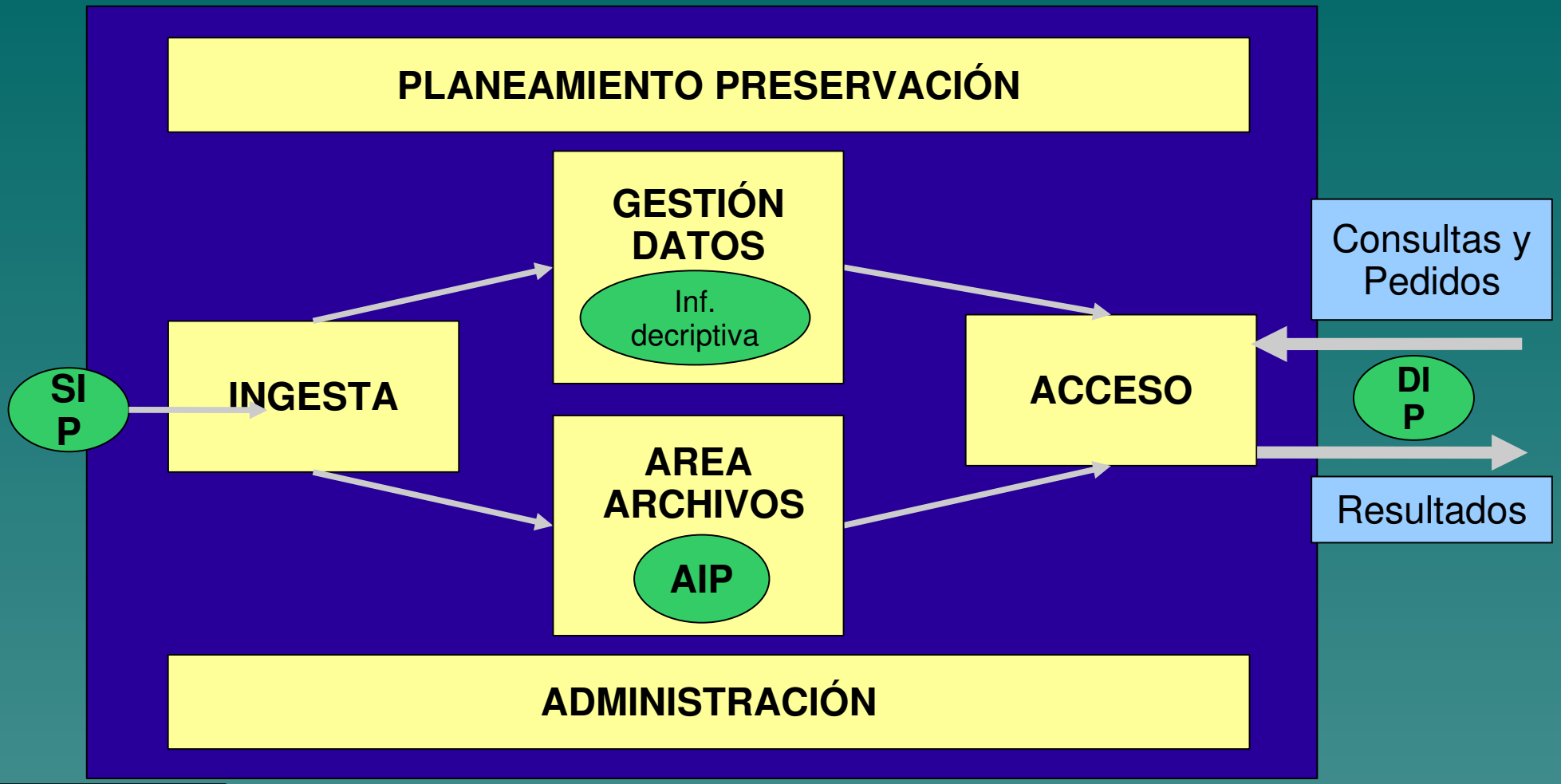


Ej.: Documento en pdf

Metadatos del documento

Documento + metadatos

OAIS Open Archival Information System



SI P — Paquete de información de ingesta (Submission Information Package)

AI P — Paquete de Información de Archivo (Archival Information Package)

DI P — Paquete de Información de Difusión (Dissemination Information Package)

OAIS - Aplicación del modelo a la Biblioteca Digital de la FCEN

- **Productores**
  - **Sector Académico**
    - Investigadores y Docentes: Tesis, papers, etc
    - Alumnos: Tesinas
    - Programa de Historia de la FCEN: Imágenes, documentos con valor histórico
    - Departamentos Académicos: Programas de estudio
  - **Gestión de la Facultad**
    - Consejo Directivo, Decanato y Secretarías: Actas, Memorias.
    - Departamentos administrativos: Registros contables

OAIS - Aplicación del modelo a la Biblioteca Digital de la FCEN

- **Gestión:**
  Biblioteca Central Luis F. Leloir de la FCEN

- **Consumidor:**
  - Investigadores y Docentes
  - Alumnos
  - Usuarios externos: Particulares o privados (Empresas, Laboratorios, otras universidades, etc.)

OAIS - Aplicación del modelo a la Biblioteca Digital de la FCEN

- **La información**
  - **Objetos de Datos:**
    - Tesis Doctorales en diferentes formatos
    - Papers en formatos en diferentes formatos
    - Otros documentos (imagenes tif, videos, etc.)
  - **Información de Representación:**
    - Metadatos ingresados en el área de Procesos Técnicos de la Biblioteca
    - Metadatos ingresados en el área Biblioteca Digital

OAIS - Aplicación del modelo a la Biblioteca Digital de la FCEN

- **Objetos de información**

  Colecciones digitales de tesis, fotos, Programas de estudio y Videos con su respectiva descripción

OAIS - Aplicación del modelo a la Biblioteca Digital de la FCEN

- **Ingesta**

  Para planificar esta etapa deben considerarse las posteriores acciones de la Gestión: carga de información de representación (metadatos), conservación de la información a largo plazo y recuperación y visualización de los objetos de información.

  - Digitalización, según pautas preestablecidas
  - Recepción de objetos nacidos digitales
  - Chequeos de integridad iniciales: completitud y función
  - Validación de formatos
  - Estandarización / migración de formatos
  - Nombramiento de archivos

OAIS - Aplicación del modelo a la Biblioteca Digital de la FCEN

- **Ingesta**
  - Digitalización: aplicación de parámetros de escaneo:
    - Digitalización de fotografías blanco y negro a color.
    - Tomando la resolución que garantice 3000 pix en el lado largo (una resolución aproximada de 600 dpi)
    - Estrategia de Masters tiff y copias de uso en jpg
  - Nombramiento de archivos
    - Esquema para el nombramiento de archivos. El nombre del archivo muestra información sobre su contenido

    Ej. 001-Diaz de Guijarro-0156

OAIS - Aplicación del modelo a la Biblioteca Digital de la FCEN

- **Ingesta**
  - Pautas para la entrega de tesis por parte del productor:
    - Resolución CD 2053 del 2004: Establece el depósito obligatorio de la versión digital de las tesis doctorales
    - Actualmente se trabaja en el establecimiento de requerimientos de formatos para la entrega, en una modificatoria de la Resolución.
    - Estos requerimientos están supeditado a las posibilidades del autor y los objetivos de acceso.

OAIS - Aplicación del modelo a la Biblioteca Digital de la FCEN

- **Gestión**
  - Acciones de Gestión de datos
    - Agregado de información de representación (metadatos)
    - Marcado de los objetos (Lenguajes de marcado)
  - Acciones de Archivo
    - Organización de archivos en carpetas
    - Guarda
    - Acciones de Preservación: Migraciones, copias redundantes

OAIS - Aplicación del modelo a la Biblioteca Digital de la FCEN

- **Acciones de Gestión de Datos**
  - Metadatos (Información de representación)
    - Utilización del esquema Dublin Core
    - Metadados específicos por colección
    - Metadatos de preservación digital

OAIS - Aplicación del modelo a la Biblioteca Digital de la FCEN

- **Gestión**
  - Esquema de metadatos de preservación

Hardware de captura	Información sobre los el hardware de captura Ej: Escaner o cámaras fotográficas
Software de creación	Software con el que fue creado el archivo.
Software de lectura	Software con el que se puede leer el archivo.
Software de edición	Software con el que fue editada la imagen. Se entiende por edición, el uso de compresión, marcas de agua, etc.
Acciones de preservación	Información sobre las acciones especificas orientadas a la preservación que deben ejecutarse sobre el objeto digital.
Procedencia e Historia del objeto digital	Información sobre las condiciones del objeto digital original.
Formato digital (Extracción automática)	Nombre del archivo / extensión de archivo / formato del archivo / peso /

OAIS - Aplicación del modelo a la Biblioteca Digital de la FCEN

- **Gestión**
  - Acciones de Almacenamiento
    - Estrategia de copias múltiples en distintos soportes (ópticos y mágneticos)
  - Acciones de preservación a futuro
    - Copias de refresco y migraciones

OAIS - Aplicación del modelo a la Biblioteca Digital de la FCEN

- **Acceso**
  - Selección e implementación de Software para bibliotecas Digitales
    Greenstone, Dspace, E-print, Gallery, TEDE
  - Derechos de acceso y uso
    Proyecto Portal de tesis UBA, formulario para autores de tesis - Grupo de Trabajo SISBI

Universidad de Buenos Aires

PORTAL INSTITUCIONAL DE TESIS Y DISERTACIONES ELECTRONICAS
Resolución (R) Nº 69/06

AUTORIZACION DEL AUTOR (1)

En calidad de TITULAR de los derechos de autor de la obra que se detalla a continuación, en la presente **INFORMO** a la Universidad de Buenos Aires **mi decisión** de concederle en forma gratuita la autorización para publicar el texto completo de mi trabajo inédito más abajo indicado, **exclusivamente en medio electrónico**, en el sitio web de la Facultad y/o Universidad, por INTERNET, a título de divulgación gratuita de la producción científica generada por la Universidad, a partir de la fecha especificada:

1.-AUTOR (APELLIDO/S y Nombre/s):

_____

Nacionalidad _____

Tipo y Nº Doc._____

Domicilio:_____

_____Código postal – Localidad _____

Teléfono/s_____

email_____

Egresado de _____ Carrera_____

Título_____

Lugar de desarrollo de la Tesis_____

2.- Identificación del tipo de material bibliográfico:
Tesis [ ] APROBADA [ ] Disertación [ ] Monografía [ ] OTROS (indicar) .....
[ ] EN CURSO

3.- Identificación del documento / autor:

Facultad en la que se presenta:_____

Carrera: _____

Area del Trabajo:
Doctorado de la UBA en _____
Maestría de la UBA en_____
Otros_____

Director/Orientador/Docente responsable: _____

Codirector_____

Consejero de Estudios (*) _____
(de la Facultad en que se defenderá, cuando el Director no es Profesor, ó no lo es en esa Casa)

TITULO del TRABAJO: _____

_____

_____

_____

Aprobado por: Expte Nº / - Acta Nº

Fecha de defensa_____/____/_____ -

4.- a) SÍ AUTORIZO : marque dentro del casillero ☐

Para uso exclusivamente personal de: a) lectura; b) bajada del archivo [download] y/o c) envío posdatado.

b) Cuándo Ud. autoriza su publicación en texto completo? :
Puede publicarse por este medio:
[ ] a partir de su aprobación
[ ] 1 año después de su aprobación
[ ] 2 años después de su aprobación

c) En qué formato digital ? :[1]
PDF: [ ] HTML o similar: [ ] OTRO: [ ] Cuál ?:.....

La UBA procederá al registro del trabajo en el carácter de "obra publicada en medio electrónico"

5.- NO AUTORIZO: marque dentro del casillero ☐

NOTA: Las tesis no autorizadas para ser publicadas EN ESTE PORTAL en TEXTO COMPLETO, serán difundidas en el mismo mediante su cita bibliográfica completa, incluyendo Tabla de contenidos, bajo la leyenda DISPONIBLE SOLO PARA CONSULTA EN SALA en su version COMPLETA , en la BIBLIOTECA DEPOSITARIA.

6.- Condiciones de acceso a la versión COMPLETA en la BIBLIOTECA DEPOSITARIA: marque dentro del /los casilleros:

sólo lectura en sala ☐ fotocopia parcial *hasta..... págs.* ☐ *fotocopia completa para uso personal* ☐

_____ _____
firma del Autor firma del Orientador / Director

Lugar_____ Fecha_____/_____/_____

_____ Fecha_____/_____/_____
firma entidad patrocinante

(1) AUTOR: *en caso de tesis "subsidiadas", la presente AUTORIZACION deberá ser otorgada con la conformidad de la entidad patrocinante.*

***Aclaración**: se nos ha sugerido que:*
1.- para las BECAS UBA el Consejo Superior autorice a cada área responsable de tesis en las facultades a firmar en su nombre; respecto de otro tipo de becas externas, que sea cada Consejo Directivo quien defina cómo se obtiene la autorización.

2.- que el AUTOR sea notificado fehacientemente al INICIO DE SU EXPEDIENTE , de que , una vez APROBADO su trabajo, deberá firmar por SI ó por NO el presente formulario, a fin de que desde un principio tanto el autor como el director sepan de la eventual publicación del texto completo en la red mundial de computadoras.

[1] El SISBI aconseja archivo de estructura abierta (no propietario)

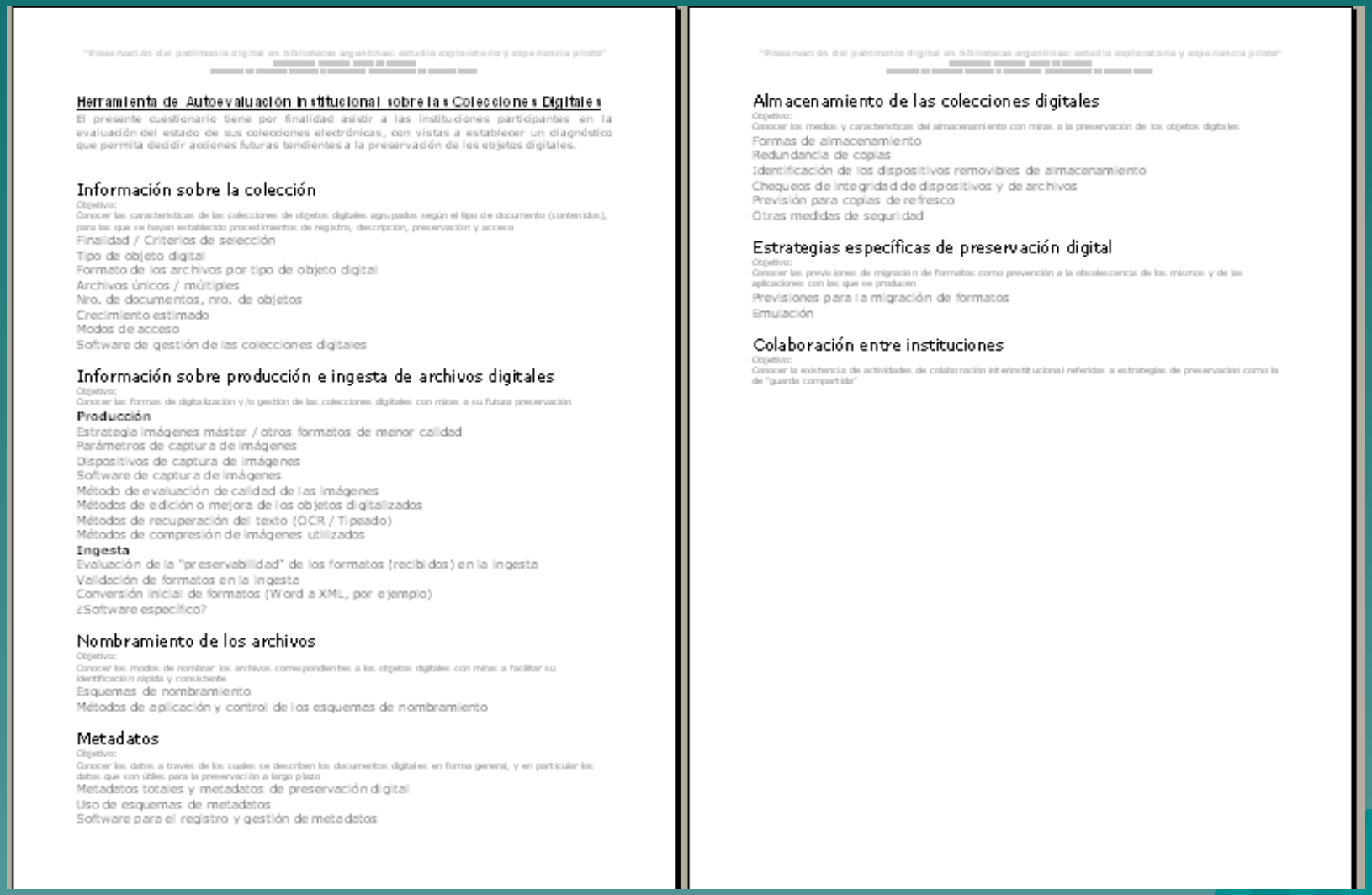
OAIS - Aplicación del modelo a la Biblioteca Digital de la FCEN

- Acciones que nos acercaron al modelo OAIS
  - Capacitación y Formación
    - Proyecto Preservación digital: Consultoría en preservación digital, Talleres de intercambio, creación de un grupo Yahoo sobre PD, Taller de Archivística Aplicada
    - Cooperación interinstitucional

OAIS - Aplicación del modelo a la Biblioteca Digital de la FCEN

- Generación de herramientas por parte de las instituciones del proyecto PDBA
  - Cuestionario para Autodiagnóstico
  - Software para marcar textos en xml/Greenstone (CLACSO)
  - Software para la generación de copias de uso de imágenes, por lote (Diapoteca FADU-UBA)
  - Guía para la digitalización (Biblioteca FHCE-UNLP)

Cuestionario de autoevaluación para colecciones digitales



"Preservación del patrimonio digital en bibliotecas argentinas: estudio exploratorio y experiencia piloto"

Herramienta de Autoevaluación Institucional sobre las Colecciones Digitales

El presente cuestionario tiene por finalidad asistir a las instituciones participantes en la evaluación del estado de sus colecciones electrónicas, con vistas a establecer un diagnóstico que permita decidir acciones futuras tendientes a la preservación de los objetos digitales.

Información sobre la colección

Objetivo:
Conocer las características de las colecciones de objetos digitales agrupados según el tipo de documento (contenidos), para las que se hayan establecido procedimientos de registro, descripción, preservación y acceso

Finalidad / Criterios de selección
Tipo de objeto digital
Formato de los archivos por tipo de objeto digital
Archivos únicos / múltiples
Nro. de documentos, nro. de objetos
Crecimiento estimado
Modos de acceso
Software de gestión de las colecciones digitales

Información sobre producción e ingesta de archivos digitales

Objetivo:
Conocer las formas de digitalización y/o gestión de las colecciones digitales con miras a su futura preservación

Producción
Estrategia imágenes máster / otros formatos de menor calidad
Parámetros de captura de imágenes
Dispositivos de captura de imágenes
Software de captura de imágenes
Método de evaluación de calidad de las imágenes
Métodos de edición o mejora de los objetos digitalizados
Métodos de recuperación del texto (OCR / Tipeado)
Métodos de compresión de imágenes utilizados

Ingesta
Evaluación de la "preservabilidad" de los formatos (recibidos) en la ingesta
Validación de formatos en la ingesta
Conversión inicial de formatos (Word a XML, por ejemplo)
¿Software específico?

Nombramiento de los archivos

Objetivo:
Conocer los modos de nombrar los archivos correspondientes a los objetos digitales con miras a facilitar su identificación rápida y consistente

Esquemas de nombramiento
Métodos de aplicación y control de los esquemas de nombramiento

Metadatos

Objetivo:
Conocer los datos a través de los cuales se describen los documentos digitales en forma general, y en particular los datos que son útiles para la preservación a largo plazo

Metadatos totales y metadatos de preservación digital
Uso de esquemas de metadatos
Software para el registro y gestión de metadatos

"Preservación del patrimonio digital en bibliotecas argentinas: estudio exploratorio y experiencia piloto"

Almacenamiento de las colecciones digitales

Objetivo:
Conocer los medios y características del almacenamiento con miras a la preservación de los objetos digitales

Formas de almacenamiento
Redundancia de copias
Identificación de los dispositivos removibles de almacenamiento
Chequeos de integridad de dispositivos y de archivos
Previsión para copias de refresco
Otras medidas de seguridad

Estrategias específicas de preservación digital

Objetivo:
Conocer las previsiones de migración de formatos como prevención a la obsolescencia de los mismos y de las aplicaciones con las que se producen

Previsiones para la migración de formatos
Emulación

Colaboración entre instituciones

Objetivo:
Conocer la existencia de actividades de colaboración interinstitucional referidas a estrategias de preservación como la de "guarda compartida"

¡Muchas gracias!