

The author is not a native speaker of Spanish. I apologize for any misspellings, typos, incorrect grammar or just plain wrong use of the Spanish language. I do recognize that there are many words throughout this document that do not have the proper diacritical marks (í, é, ó, ñ, etc.) This is a function of Microsoft Word's inability to deal with Unicode and the difficulty of adding each mark in, one by one. If anyone wants to send me corrections of my Spanish, it would be greatly appreciated and taken as a kindness and professional courtesy. –Carol Hixson

Slide 1

Gracias por la invitación de hablar aquí. Voy a hablar en español pero pido vuestra paciencia. No conozco bien el vocabulario de las bibliotecas digitales. Y también debo decir que he estudiado el español, el francés, el portugués, y el italiano y algunas veces las palabras que salen de mi boca son palabras mezcladas de los 4 idiomas. Espero que me perdonéis.

Mi tema es la realidad y la ilusión, un tema que encontré hace muchos años en mis estudios de Cervantes, Fuente Ovejuna, Buero Vallejo, Unamuno y muchos otros autores españoles de la edad de oro y de la edad moderna.

Una realidad es que no soy experta. Es una honra estar aquí con los verdaderos expertos. La realidad es que soy una bibliotecaria que está inventándose, que está descubriendo cada día que no sabe la mitad de lo que debe saber para hacer lo que está haciendo.

Voy a hablar de algunas otras realidades y creo que podréis ver las ilusiones claramente.

La URL que podéis ver aquí es la dirección de esta presentación. Siempre cuando hago una presentación, la deposito en nuestro archivo institucional. Se puede encontrar la presentación completa. También se puede encontrar allá un documento con referencias a todos los documentos, sitios, y estándares de que voy a hablar hoy.

Slide 2 UO

Sé que los europeos conocen la geografía mejor que los americanos. Yo vivo en el estado de Oregon, bien al oeste de los Estados Unidos, norte del estado de California. La mayoría de los americanos no saben este detalle. Una vez yo tenía una discusión con una mujer en Chicago quien insistía que Oregon fuera parte de Canada.

La Universidad de Oregon es la mayor Universidad en el estado y la única del primer nivel. Cuenta con más de dieciocho mil (18,861) estudiantes y más de ciento veinte (129) programas de estudio. Hay siete bibliotecas en el sistema de la Universidad de Oregon. Trabajan en las siete bibliotecas más de 50 bibliotecarios, casi 100 otros empleados, y más de 60 estudiantes

Slide 3 MDLS

Mi departamento tiene 30 empleados, incluyendo siete bibliotecarios. Nosotros tenemos la responsabilidad por todo tipo de colección digital, incluyendo el repositorio institucional, y también la catalogación y la preservación de los recursos tradicionales (libros, periódicos), y el servicio de microfilm.

Pues una realidad es que estamos construyendo una biblioteca digital sin ser expertos, como ya he dicho. Los expertos son los otros y nosotros intentamos de enseñarnos, de aprender los estándares, los procesos, la tecnología al mismo tiempo que construimos las colecciones.

Otra realidad es que las colecciones digitales son solo una responsabilidad entre muchas otras responsabilidades que tenemos. Para hacerlo bien, creo que se necesita especializar. Se necesita tener tiempo, dinero, conocimiento, flexibilidad, y energía. Nosotros no tenemos una cantidad suficiente de ninguno de estos elementos.

Slide 4 digcol home

Ese es el sitio web de todas nuestras colecciones digitales. La mayoría usan el programario de CONTENTdm. Ese es un sitio colectivo para todas las colecciones, sin respeto al programario que utilizan. Fue el primer esfuerzo de coordinar todas las colecciones digitales que tenemos. Tenemos colecciones de imágenes, de cartas, de documentos, de recursos en chinos, japones y coreano, de medios por caudales, de colecciones de EAD (encoded archival

description) para describir las colecciones de archivos y manuscritos. Un poco de todo.

En CONTENTdm tenemos ahora dieciseis mil quinientos objetos digitales. La mayoría son JPGs de fotografías, de documentos,

Slide 5 Scholars' bank

Una de nuestras colecciones digitales se llama Scholars' Bank. Es un archivo digital donde coleccionamos los artículos, papeles, presentaciones, periódicos, revistas, y mucho más de nuestros profesores, estudiantes, y otros miembros de la universidad. Con este archivo – como con todas nuestras colecciones digitales – hay acceso abierto a través de internet.

Como muchas otras instituciones, nuestro esfuerzo fue una respuesta al abusivo aumento de los precios de las suscripciones de los periódicos. Nuestro objetivo al principio era el de ser una parte del movimiento mundial tendiente a reformar o cambiar el modelo de comunicación académica. Nuestro objetivo a largo plazo es sólo facilitar el conocimiento - la ciencia- a través del establecimiento de una infraestructura adecuada que puedan utilizar los académicos directamente, esto es, sin intermediación de terceros. No voy a hablar más de Scholars' Bank hoy. Esta es otra presentación completa que se puede encontrar en el archivo mismo – una presentación que hice el año pasado en Barcelona.

Slide 6 IR presentation

Aqui es la información de la presentación en español sobre nuestro repositorio institucional, Scholars' Bank, si teneis interes.

Slide 7 que, por que, como

- Primero, voy a dar una mirada de algunos objetos en nuestras colecciones
- Voy a presentar alguna información sobre:
- Las razones por que hemos comenzado de crear las colecciones digitales
- Porque escogimos el programario CONTENTdm para la mayoría de nuestras colecciones
- Algunos puntos de nuestro proceso de implementación
- Como se usan las colecciones
- Los retos
- Y los planos para el futuro

Slide 8 descriptions

Actualmente tenemos 9 colecciones digitales publicas que usan el programario de CONTENTdm. Hay algunas otras en progreso. Aqui se explican estas colecciones y algunas otras. Intentamos dar un contexto para nuestro programa digital – para nosotros mismos y para nuestros usuarios. Explicamos lo que todavia hemos hecho, lo que estamos haciendo, y lo que esperamos hacer.

Slide 9 tour

Hemos creado una exhibición de varios ejemplos de nuestras colecciones digitales. La imagen cambia automaticamente o se puede adelantar a mano. Estamos en vias de revisar nuestros sitios web para nuestras colecciones digitales. Por primera vez, tengo una persona que se preocupa completamente con las colecciones digitales. Ella está examinando todos los sitios para identificar problemas, falta de consistencia, y mas. Espero que dentro de 4 o 5 meses la interfaz sea mucho mejor.

Slide 10 Tour

Esta es una imagen historicamente significativa de Alaska de 1906 del fotografo Clarence Andrews. Como he dicho tenemos mas de cuatro ciento mil fotografias en nuestras colecciones. Solo hemos digitalizado entre 7000 y 8000 mil.

Es dificil decidir cuales de los fotografias merecen ser digitalizadas. Intentamos no seleccionar varias vistas del mismo lugar o de la misma persona, generalmente.

Slide 11 reasons - access

Voy a hablar un poco de las razones porque nosotros comenzamos a crear las colecciones digitales.

Una razon principal era de proveer acceso a los materiales poco utilizados (como los materiales raros o unicos en nuestro departamento de Colecciones Especiales.

Slide 12 access

Hay muchos materiales especiales en nuestras Special Collections, como los 400,000 fotografias. Tambien tenemos manuscritos medievales, manuscritos unicos como los diarios del "Oregon Trail" o los papeles de los presidentes de la Universidad.

Con los recursos originales, un usuario tiene que visitar un lugar con horas limitadas, tiene que registrar, y tiene que esperar la entrega de los materiales para poder utilizarlos. La mayoría de las colecciones allí también no tienen ningún acceso bibliográfico. Pues es muy difícil descubrir que hay esos materiales y aún más difícil utilizarlos.

Con las colecciones digitales nosotros anunciamos la existencia de las colecciones ocultas. El mundo entero puede descubrir la riqueza de nuestras colecciones. Recibimos muchas peticiones de escolares, investigadores, editores, autores, y otros que quieren utilizar nuestras imágenes en su trabajo. Muchos de ellos aprenden que tenemos estas colecciones por medio de nuestras colecciones digitales.

Slide 13 reasons – broaden access

Otra razón era de aumentar el acceso a los materiales físicamente frágiles, como los manuscritos.

Slide 14 broaden access

Por ejemplo, como este manuscrito medieval. Con las imágenes digitales el mundo entero puede utilizar estos materiales frágiles. Con un objeto tan raro y tan frágil, pocas personas reciben permiso para utilizar el original. Con una colección digital, el mundo entero puede verlo y estudiarlo.

Slide 15 reasons - preserve

También queremos preservar algunos materiales que están en peligro de desaparecer.

Slide 16 preserve

Por ejemplo, el original de esta imagen existe en vidrio. Es un negativo fotografico en vidrio. Puesto que nuestra region es una zona de terremoto, siempre hay la posibilidad de perder toda esta colección fotografica. Ahora con algunas imágenes digitales, podemos preservar por lo menos una parte de la colección.

Slide 17 reasons - partnerships

Otra razon es que queremos establecer y fortalecer las asociaciones o colaboraciones con las otras instituciones del patrimonio

Como los museos, otras bibliotecas, y mas

Slide 18 partnerships

Por ejemplo, esta imagen es parte de una colección desarrollada con la ayuda de las tribus de indios que viven en otra parte de Oregon. Voy a hablar mas de esta colección.

Fue nuestra primera colección digital usando CONTENTdm y fue el resultado de una colaboración entre el Instituto Cultural Tamastlikt de las Tribus Confederadas de las Umatillas (TCI). Recibimos una subvención pequeña para establecer la colaboración e desarrollar la colección digital. Fue una de las mejores experiencias de mi vida profesional. Todavía estoy en contacto con los miembros de las tribus e llego a otras oportunidades de colaborar. Eso es un resultado típico – una colección digital llega a otras colaboraciones y proyectos. Eso es por mi lo más interesante e importante de mi trabajo.

Slide 19 partnerships

Esta imagen representa otro proyecto colaborativo entre la biblioteca y el museo de la historia natural y cultural de la universidad.

Ellos tienen imágenes digitales de una colección de cestas hechas por los indios de nuestra región. Nosotros en la biblioteca estamos ayudando en la creación de una colección digital, con metadatos y una presentación pública donde se puede ver las cestas de varios puntos de vista. La biblioteca va a hospedar la colección digital en nuestros servidores. Espero que tendremos una colección pública dentro de 4 o 5 meses.

Slide 20 reasons – new delivery mech

queremos investigar nuevos modos de presentar el contenido de los documentos, como la búsqueda del texto completo, la indización del texto completo

Slide 21 new delivery

Al mismo tiempo se puede ver el libro y leer las paginas completas y se puede buscar una frase o palabra y leer solo una parte especifica.

Es algo que no se puede hacer con el libro original. Eso es nuestra razon principal por digitalizar textos – este nuevo modo de utilizar el texto.

Slide 22 new delivery

Aquí se puede ver como funciona la indización del texto completo en CONTENTdm. Las paginas en que aparece el termino buscado aparecen aquí con color rojo. El termino buscado aparece con letra negrita.

.

Slide 23 reasons – shape digital landscape

Otra razon es que queremos participar en el desarrollo de los estándares.

Slide 24 shape digital landscape

Por ejemplo, esta colección representa el trabajo de un consorcio de 12 bibliotecas. Cada biblioteca tiene su propia colección de materiales y contribuimos los metadatos a una colección central. Con los otros bibliotecas estamos desarrollando y poniendo a prueba una gran variedad de estándares. Hay muchos estandares que necesitan ser probados y a nosotros nos gusta ser una parte de la investigación e expansion de los estándares.

Slide 25 reasons – support instruction

También queremos apoyar los programas de instrucción de la universidad. La misión de la universidad es enseñar – los estudiantes de la universidad y el público de nuestro estado. Todo lo que hacemos en la biblioteca tiene el objeto de apoyar la misión de la universidad.

Slide 26 support instruction

Esta imagen es parte de una colección que creamos para una clase de historia del arte. El profesor seleccionó las diapositivas que quería utilizar en sus clases. Nosotros creamos las imágenes digitales y proveemos acceso a ellas antes de cada clase en que él quería utilizarlas. Fue un reto para nosotros porque muchas veces él seleccionó las diapositivas solo un día antes de querer utilizarlos en su clase y fue difícilísimo digitalizarlos y presentarlos en solo 24 horas.

Slide 27 support instruction

Esta colección fue el resultado de un proyecto de una clase de honra que estudiaba los años de la guerra de Vietnam por medio de nuestra colección de los papeles de los presidentes de la Universidad. Los estudiantes seleccionaron varios documentos y mi departamento los digitalizó dentro de 48 horas y los hicimos accesibles en esta colección digital. El gran reto de esta colección fue la necesidad de crear texto completo de documentos viejos, muchas veces escritos de mano. Voy a hablar un poco más de esto en algunos minutos.

Slide 28 support instruction

También creamos una colección en nuestro archivo digital Scholars' Bank para los papeles de los estudiantes. Las dos colecciones están relacionadas. Aquí se ve el "hiperenlace"

Slide 29 blog

Este blog forma parte de una clase que examina como la tecnología está cambiando el modo de aprender y enseñar. Fue el resultado de una colaboración entre la biblioteca y dos profesores creando una colección digital. Los profesores quieren que yo os invite de leer el blog y participar en la discusión. Seria un gran placer para ellos recibir sus comentarios.

Continued

Acabamos de tener una jornada sobre los efectos de la tecnología sobre la indagación academica y luego podrais encontrar todas 4 horas de la jornada en nuestro archivo digital, Scholars' Bank. Os invito a mirar los ficheros de video por caudeles en el archivo.

Slide 30 why content

He dicho que utilizamos el programario, el sistema de CONTENTdm. Estamos mas o menos satisfechos con el sistema.

Algunas de las razones porque lo escogimos son:

- El programario acepta una variedad de formatos digitales, una variedad de ficheros.
- Permite la adquisición de objetos uno tras uno o de un proceso de “trabajo en serie” – un grupo de objetos al mismo tiempo
- La interfaz puede ser modificada fácilmente
- Hay muchos usuarios del programario en nuestra region, pues podemos pedir ayuda a colegas y podemos colaborar con otros mas facilmente
- Los metadatos son de Dublin Core – apoya Dublin Core simples y calificado – eso es muy importante para poder distribuir los metadatos y colaborar
- Se puede utilizar un vocabulario controlado para cualquier “campo de datos” – eso ayuda bastante en la descripción de los objetos

Slide 31 customizable

Aqui se puede ver la administración tecnica de una colección en CONTENTdm. Es muy facil crear nuevos campos de datos, modificar las etiquetas, determinar si van a utilizar un vocabulario controlada, determinar si el publico va a poder buscar este “campo de datos” o verlo. Y es muy facil cambiarlo despues de hacer la colección accessible al publico. Algunas veces creo que es demasiado facil cambiarlo porque es posible gastar mucho tiempo modificando lo que todavia se ha hecho.

Slide 32 mapping to DC

Aquí se ve que se puede escoger el campo de datos que quiere mapear a un elemento de Dublin Core. Los administradores de una colección digital necesitan solo algunos segundos para cambiar el mapping. Al mismo tiempo pueden cambiar la etiqueta del campo de datos.

Slide 33 OAI registries

Puesto que utilizamos Dublin Core metadatos, podemos hacer conexiones con otras colecciones que también los usa. Por ejemplo, OAIster es un repositorio que cosecha los metadatos de varios archivos y colecciones que usan Dublin Core.

OAIster es un proyecto de la Universidad de Michigan, del su Servicio de Producción de Bibliotecas Digitales. La OAI quiere decir “Open Access Initiative” o Iniciativa del Acceso Abierto.

Continued

OAIster y muchos otros registros semejantes se conforman con los estándares del Open Archives Initiative-Protocolo para el Cosechamiento de Metadatos (OAI-PMH). Una búsqueda en OAIster revela ciento noventa objetos digitales en nuestras colecciones con el término de “cayuse” – que es un término para un tribu de indios. Nosotros registramos nuestras colecciones digitales con registros semejantes para multiplicar la posibilidad del descubrimiento de nuestras colecciones.

Slide 34 controlled vocabs

Aquí se ve la administración del vocabulario controlado. El programario apoya TGM – Thesaurus for Graphical Materials – un vocabulario que utiliza la Biblioteca del Congreso de los EEUU para describir imágenes o recursos gráficos.

Pero también el programario acepta cualquier otro vocabulario que se quiere utilizar. Para nuestras colecciones digitales hemos desarrollado una variación del vocabulario de la biblioteca del congreso de los EEUU.

Otras colecciones, como el primero con las imágenes de los indios, tiene varios otros vocabularios que desarrollamos nosotros mismos.

Slide 35 highlights

Cuando comenzamos, tuvimos un grupo llamado el Metadata Implementation Group. Este equipo ahora funciona un poco como el comité gobernador (steering comité) siendo yo la directora de las colecciones digitales. Tenemos una lista electrónica para discutir toda situación nueva que se presente. En el comienzo, nuestro equipo reunió regularmente cada dos semanas. Hoy es raro que nos reunamos. Al contrario, normalmente nosotros estamos acostumbrados de tomar decisiones sobre la lista

Slide 36 first collection

Nuestra primera colección con los indios fue bien complicada y el grupo de Metadata Implementation group fue muy activo en el desarrollo de la colección.

Ahora, mi departamento trabajo directamente con varios asociados o conservadores de colecciones para determinar las políticas y procesos.

Slide 37 use

Para manejar y promover nuestras colecciones, queremos saber una variedad de cosas, como: quién visita el repositorio, cuántas veces se visita, cuáles ficheros se utilizan, como han descubierto el repositorio, de dónde son los usuarios, y más.

CONTENTdm tiene algunas herramientas administrativas para proporcionarnos alguna información, pero no la mayoría de la información que queremos saber. Nosotros utilizamos un programario llamado Webalizer para recoger y analizar varias estadísticas, como por ejemplo las estadísticas sobre el uso del repositorio. Webalizer analiza los registros estándares de

un servidor. Nos da la información básica sobre quién visita nuestro sitio de Web, que tipo de navegador utilizan, de dónde son referidos, cuáles páginas ellos ven, etc.

Es probable que sigamos utilizando una variedad de herramientas para descubrir y analizar las estadísticas.

Este es un resultado de Webalizer, mostrando un resumen de las estadísticas – Aquí se pueden ver los golpes o consultas al servidor. Webalizer también nos da las estadísticas de manera más detalladas, como el uso diario.

Ese es un diapositiva que muestra el uso de nuestras colecciones. Se puede ver que se han visitado nuestras colecciones digitales mas de 4 millones de veces en un año. Eso no es el resultado de publicidad, mas bien es porque utilizamos los metadatos de Dublin Core que facilita el descubrimiento de las colecciones.

Slide 38 use - searches

Aquí se puede ver una selección de las búsquedas que se hacen en nuestras colecciones digitales. Es posible ver el uso hora por hora.

Webalizer también nos da información sobre los países de origen de nuestros usuarios. Hasta hoy, hemos recibido visitas de más de 70 países. Webalizer nos proporciona la misma información en forma gráfica.

Slide 39 ordering

También vendemos copias de nuestra imágenes y tenemos una forma para facilitar la compra. Lo que cobramos depende del usuario y del uso que quiere hacer de la imagen. No vendemos ficheros de documentos que digitalizamos.

Slide 40 comment form

También tenemos una forma para permitir los comentarios sobre las colecciones o una imagen específica. Recibo comentarios de todas partes y son muchas veces muy interesantes y algunas veces informativas.

Slide 41 comment received

Ese es parte de un comentario que recibimos. La persona se interesó mucho en el perro que aparece en la foto. (Como he dicho, los comentarios son muy interesantes – y reveladoras algunas veces.)

Slide 42 comment on image

Aquí se ve la foto - y el perro durmiendo.

Slide 43 challenges - selection

Ahora voy a hablar un poco de los retos

El primer reto con cada colección es la selección de materiales. No es posible para nosotros digitalizar todos los recursos que tenemos.

Slide 44 selection

Algunas colecciones tienen una política escrita para la selección. Otras no. La selección depende del público primario por lo cual hacemos la colección. Con esta colección, un fotógrafo escogió algunos ejemplares de varias colecciones de fotografías históricas que tenemos para dar una representación de la riqueza de nuestras colecciones. **CONT**

No fue posible crear imágenes digitales de todas las fotos – fue necesario seleccionar cuidadosamente, según principios establecidos por los conservadores de las colecciones originales.

Es un reto porque muchas veces los conservadores de las colecciones no tienen el tiempo de seleccionar cuando nosotros tenemos el tiempo y el personal para digitalizar. Muchas veces, ellos seleccionan poco antes de querer acceso completo a la versión digital. Es un problema entre la realidad y la ilusión.

Slide 45 challenges - metadata

Un reto enorme son los metadatos. Cuales son los estándares que vais a utilizar? Tenéis el tiempo de proveer todos los metadatos?

Teneis el tiempo de revisar y corregir los metadatos? Porque puedo aseguraros que vais a querer revisarlos.

Slide 46 metadata - cayuse

Con esta colección tuvimos grandes retos con los metadatos. Queríamos dar a las tribus la oportunidad de describir las imágenes de sus antepasados en sus propias palabras. Al mismo tiempo, queríamos crear una colección para nuestros estudiantes y profesores. Utilizamos “campos de datos” distintos de metadatos para las descripciones de las tribus y para nuestras descripciones- pero son mapeados al mismo elemento de Dublin Core.

Slide 47 poker jim

Esta es una de las imágenes de esta colección. Cuando nosotros, los académicos blancos, miramos esta imagen, vemos algo distinto de lo que ven las tribus. Por ejemplo, los indios ven una foto falsificado – el “headdress” de plumas no se pone correctamente y muchos otros detalles que solo los miembros de las tribus saben.

Slide 48 TCI metadata

Aquí se ven los metadatos dados por las tribus. TCI es una abreviación del nombre de la confederación de tribus con que trabajamos. Queríamos proveer una versión de los metadatos que presentaría el punto de vista de las tribus.

Slide 49 complete descriptive

Aqui se ven todos los metadatos descriptivos – los nuestros y los de las tribus. Desarrollamos un metodo de esconder o mostrar – este boton que dice Toggle TCI View

Slide 50 technical metadata

Los metadatos tecnicos son muy dificiles. Hay gran variedad en los estándares en el mundo. Los metadatos tecnicos son los que ayudan a manejar y preservar las colecciones digitales. Referimos a varios estándares, como se ve aqui. Nos referimos al trabajo de RLG/OCLC, de la Coalición de la Preservación Digital, de la federación de bibliotecas digitales, y otros

Slide 51 view of technical

Desarrollamos un metodo de esconder o mostrar los metadatos tecnicos. El publico general no tiene interes en tales metadatos. Pero otras bibliotecas o museos o instituciones culturales tienen interes en las decisiones que hacen los otros. Es por eso que nosotros tenemos este mecanismo para mostrar esta información si alguien tiene interes.

Slide 52 challenges - interface

El diseño del sitio web, la interfaz del usuario, es un reto para nosotros. Estamos aprendiendo al mismo tiempo que creamos las colecciones. En mi departamento no tenemos empleados que se especializan en los sitios de web – es solo una responsabilidad entre muchas otras. Hay un grupo en la biblioteca que se especializan en el diseño de sitios web pero cobran mucho para sus servicios – mismo dentro de la biblioteca.

Slide 53 browsable subject lists

Esta es una manera en que intentamos de proveer acceso a nuestras colecciones digitales – con las listas de términos utilizados en la colección. No tenemos estas listas en cada colección y esto es una cosa que queremos hacer uniforme.

Slide 54 drop-down nav

Lo normal de nuestra interfaz es un “cuadro de búsqueda” para buscar la colección completa.

Otro ejemplo de nuestra interfaz – son las selecciones en menús plegables. Las selecciones en cada menú depende de la colección, el tipo de material, los usuarios, y más. Aquí las selecciones son una lista de todas las imágenes, una lista de nombres, de lugares, de sujetos, y una búsqueda avanzada.

Slide 55 advanced search

Aqui se ve la busqueda avanzada que desarrollamos. El programario tiene otra interfaz mas simples para este tipo de busqueda.

Slide 56 browse by coll

Aqui es un ejemplo de una interfaz donde se puede buscar colecciones distintas dentro de una colección mas grande.

Slide 57 photographer

Otro tipo de interfaz – buscar por medio del fotografo.

Slide 58 challenges – context

El contexto de una colección digital es un reto enorme. Son los conservadores de los materiales originales que conocen las colecciones y son ellos los que tienen que preparar el contexto. Pero para nosotros es difícil – ellos no tienen el tiempo de proveer todo el contexto cuando comenzamos a presentar la colección digital al publico. Siempre hay una lucha entre la realidad del esfuerzo necesario y la ilusión – los conservadores, los administradores, quieren las colecciones pero no se dan cuenta del esfuerzo para crearlas ni tampoco del tiempo necesario.

Slide 59 context moorhouse

Por ejemplo, esta pagina representa solo una parte pequena del contexto para esta colección. Podeis ver que representa mucho trabajo, mucha investigación.

Slide 60 context – office of pres

Sin contexto, es dificil interpretar una imagen o objeto digital. Por ejemplo con esta colección los estudiantes seleccionan varios documentos. Pero muchos veces hay documentos originales que existen en la misma carpeta que explicarian mejor el contexto.

Cuando se selecciona solo un documento de una carpeta, se pierde este contexto tan importante. Una manera en que intentamos proveer un poco de este contexto es con algunas busquedas pre-coordinadoras de temas principales.

Slide 61 context – hist photos

Otro ejemplo del contexto que proveemos para nuestra colecciones digitales – esta vez es información sobre el fotografo. Tenemos mucha información sobre el fotografo pero también se puede buscar – hojear – la colección por medio de un clic aqui.

Slide 62 challenges – multiple collections

Es difícil crear y desarrollar varias colecciones simultáneamente. Pero hay tanto interés en las colecciones digitales que siempre estamos desarrollando por lo menos 3 colecciones al mismo tiempo. Todo mundo tiene interés en estas colecciones – nuestro público, nuestra administración, nuestros donadores.

Slide 63 multiple collections

Manejar los proyectos – y cada colección es distinta y representa un proyecto – es uno de los grandes retos para nosotros.

Cada colección necesita una persona que se preocupe con los detalles – la selección de los materiales originales, donde están los ficheros, hemos acabado con los metadatos, quien va a crear las páginas de web para el público? Y mucho más.

Slide 64 challenges - preservation

Pero es la preservación de los materiales digitales que yo opino ser el mayor reto.

La preservación constituye el problema más serio que tenemos en la actualidad. Vamos a preservar sólo los dígitos binarios? O bien vamos a preservar el contenido? Tuvimos una discusión muy larga en nuestro equipo sobre este **CONT**

particular. La decisión que tomamos finalmente es que intentaremos preservar los dos – los dígitos y el contenido. Hay varias gestiones que seguimos para intentar de proteger la integridad y la utilidad de los ficheros y en esta página de nuestro sitio las explicamos.

Slide 65 DCC

Hace casi dos años que tenemos un grupo dedicado al estudio de la preservación digital. Preservar los materiales digitales es muy complejo y tiene muchos componentes. En los últimos dos años, asistí a tres seminarios para aprender los principios de la preservación digital – cada uno duro una semana completa. Pues no tengo el tiempo aquí de explicar bien la complejidad de este sujeto. Practicamente es un proceso de aminorar los riesgos

Slide 66 Links to digpres resources

Aquí se puede ver algunos de los estándares que consultamos para la preservación digital. Es difícil leer y entender todo lo necesario en este campo porque la tecnología y los estándares siempre cambian.

Slide 67 digpres

La preservación digital es muy complicada. Es una serie de acciones manejadas para asegurar la utilidad y accesibilidad de los ficheros. Una parte consiste en asegurar el mantenimiento del dígito binario. Una acción necesaria es hacer copias de los objetos digitales y guardarlas en otro sitio. Otra acción es examinar los ficheros periódicamente para averiguar que no haya ningún menoscabo.

Slide 68 digpres

Pero también se necesita asegurar que el contenido de los ficheros siga siendo accesible a largo plazo. Es la información intacta? Se puede leerla con un ordenador?

Los seres humanos pueden verla y comprenderla?
Se necesita asegurar el contexto de los ficheros también.

Slide 69 digpres

Hay muchas estrategias específicas para asegurar la preservación de los materiales digitales. Estas son algunas que se puede utilizar. No tengo ni el tiempo ni el vocabulario en español para explicarlas.

CONTINUED ...

Pero la estrategia que escogéis va a depender en los que queréis realizar y también en vuestros recursos – el dinero, el personal, la infraestructura técnica que teneis.

Slide 70 components

Aquí hay algunos componentes de un programa de preservación digital. El registro de los metadatos, el registro de los formatos, la verificación del checksum, los procesos de respaldo de los ficheros, un sistema de identificación persistente, como los URLs

Slide 71 metadata registry

Ese es un registro de metadatos que existe en el programario de DSpace que utilizamos para nuestro archivo digital, Scholars' Bank.

Slide 72 metadata registry

CONTENTdm tiene una interfaz administrativa donde se puede controlar los metadatos de cada colección. Puede funcionar como un registro de los metadatos para cada colección.

Slide 73 Format registry

El registro de los formatos de bitstream constituye una parte de un programa de preservación. Con un registro no se pierde la vista de los formatos en vuestras colecciones y se puede determinar los procesos necesarios para asegurar los ficheros.

Sin embargo, se necesita establecer políticas para tener un programa sólido de preservación. Nosotros tenemos un grupo llamado Los Coordinadores del Contenido Digital que establece políticas sobre la preservación de todo tipo de contenido digital que creemos o obtenemos en la biblioteca.

Ese es el registro de formatos de DSpace.

Slide 74 checksum

Otra parte de un programa de preservación lo constituye la verificación de los checksums. Los checksums son una herramienta que ayuda con la verificación de la estabilidad de un fichero. Al punto de adquirir un fichero, es bueno notar esa formula para poder investigar el menoscabo o deterioración mas tarde. Nosotros todavía no tenemos un programa automático para comprobar los checksums – esto es, para verificar que el fichero efectivamente no tiene defectos. Hemos hablado de la necesidad de comenzar un programa. Ahora lo hacemos de vez en cuando, pero esto es completamente insuficiente.

Slide 75 backup

Es bueno establecer los procesos de respaldo y documentar los procesos. Eso es una parte necesaria para ser un archivo digital fiado.

Una parte muy importante es el backup – el respaldo. Nuestro respaldo utiliza la cinta. Nuestro sistema es el de respaldar el base de datos en cinta con esta frecuencia:

- Hace una semana
- Hace dos semanas
- Hace tres meses
- El diez de julio del previo año fiscal

Slide 76 backup

Se necesita explicar las políticas de respaldo y como funcionan.

También hacemos vaciado de datos que residen en el espacio de respaldo en el servidor mismo,

Cada cinta de respaldo contempla también los vaciados de datos, tan bien como la copia dinámica de la memoria del servidor completo en el día del respaldo. Esto incluye también los ficheros.

No se puede normalizar o restaurar CONTENTdm de la cinta. Por eso también hacemos los vaciados de datos. El horario del vaciado de datos es con esta frecuencia:

CONTINUED...

- Cada día
- Los últimos siete días
- El primer día de cada mes
- Los tenemos desde el noviembre de 2004

Slide 77 TDR

Digo que estamos intentando hacerse un archivo digital fiado. Un archivo digital fiado es uno cuya misión es proveer acceso seguro a largo plazo a los materiales digitales para una comunidad específica. Sus características son:

- Responsabilidad administrativa
 - Eso quiere decir un compromiso explícito a cumplir con los estándares, políticas, y procesos que se pueden medir y revisar
- La viabilidad organizadora
 - Eso quiere decir que la misión de la organización, su posición legal, y sus operaciones son suficientes para sostener un TDR
- La sostenibilidad financiera
- La aptitud tecnológica y procedural
 - Para poder evitar la tendencia de la infraestructura técnica de caer en desuso
- La seguridad del sistema
 - Tener procesos adecuados de seguridad
- La contabilidad de procesos
 - Documentación de procesos y políticas

- Cumplir con el modelo de referencis de OAIS - explicado en el diapositiva siguiente

Slide 78 OAIS reference model

El modelo de referencia del sistema de información para los archivos abiertos fue desarrollado por NASA. El objeto fue determinar los papeles y los componentes necesarios para preservar la información digital. Muchas otras organizaciones trabajando en la preservación digital utilizan este modelo de referencia para asegurar que todo mundo habla del mismo proceso con los mismos terminos.

Slide 79 challenges - scanning

El conocimiento tecnico para crear estas colecciones es un reto principal. Para nosotros, estamos aprendiendo cada dia. Algunas de las cosas que hicimos al principio no haríamos hoy. Pero no tenemos tiempo de revisar el pasado.

Tuvimos que aprender los procesos de escanear, modificar las imágenes, y crear el texto completo y los procesos presentan grandes retos.

Slide 80 scanning

La digitalización de texto presenta retos específicos, especialmente cuando se quiere crear texto para poder buscar todas las palabras.

Slide 81 scanning

Esta diapositiva muestra un ejemplo del texto que el programario crea de una imagen de una página – se puede ver que hay muchos errores con las palabras. Para poder utilizar una búsqueda del texto completo, es algunas veces necesario que un ser humano corrija el texto – palabra por palabra

Slide 82 image adjustment

También hay retos con la fidelidad a la imagen original. El original de esta imagen era una diapositiva en vidrio con los colores pintados de mano. Era una diapositiva para un farol.

Con esta colección era muy difícil reproducir los colores del original. Una parte de la dificultad era probar de reproducir la intención del artista con la proyección con farol.

Slide 83 challenges – levels of access

Otra complejidad es como manejar derecho de copia y acceso restringido.

Slide 84 restricted - IP

Una manera en que manejamos acceso restringido es por medio de la dirección IP – la dirección del ordenador.

Slide 85 restricted – sagrada familia

Por ejemplo, esta es una imagen de la Catedral de la Sagrada Familia en Barcelona – es una imagen que hemos comprado y que presentamos al público por medio del programario de CONTENTdm. El acceso es limitado al campus porque no es nuestra imagen – pagamos por el privilegio de presentar esta colección a nuestros estudiantes y profesores.

Slide 86 restricted - password

Otra manera de limitar acceso es con un “contraseña”, como con esta colección creado para una clase de la historia del arte.

Slide 87 password control

Se puede ver las imágenes si se tiene el “contraseña” y después de hacer el login. Esta diapositiva muestra solo una lista de los resultados de una búsqueda. Cuando se escoge una imagen, se ve la imagen completa y todos los metadatos.

Slide 88 without password

Sin la contraseña, solo se ven los metadatos y una representación de la imagen.

Slide 89 copyright

También es un reto determinar los mecanismos de proteger nuestro contenido. Con las imágenes de nuestras fotos, utilizamos un “brand” en la presentación.

También es difícil establecer quien es el dueño de los materiales originales algunas veces y si tenemos el derecho de crear una colección digital de cierto contenido.

Slide 90 copyright

También explicamos nuestras políticas sobre el uso de las imágenes. El derecho de copia con los materiales digitales es muy complicado – mucho más que con los materiales originales. Yo podría dar una presentación completa sobre los retos con copyright. Un reto es que cada país tiene una interpretación distinta de los derechos de autores o de creadores de contenido. Creo que la Europa juega un rol muy importante con todas las iniciativas de la Unión Europea.

Slide 91 challenges - partners

Otro reto es trabajar con los asociados, los colaboradores

Slide 92 partners – western waters

Conozco a una administradora de una biblioteca que dijo que la colaboración conserva el tiempo. No estoy de acuerdo. Por lo menos, no al principio de un proyecto. Un ejemplo es nuestra colección sobre las aguas del oeste de los Estados Unidos. Es una colección de un consorcio de 12 bibliotecas.

Slide 93 WWDL

Este es un “screen” de la colección integrada de las 12 colecciones.

Slide 94 WWDL

Aqui se ven los resultados de una busqueda

Slide 95 special challenges

Creo que hay retos especiales con las colaboraciones

- Las decisiones sobre metadatos son mucho mas complicadas
- Las busqueda integrada de colecciones distintas presenta problemas unicas
- La comunicación entre colaboradores es un reto enorme
- La tecnologia es un reto
- Y el financiamiento de una colección sigue siendo un reto casi imposible

Slide 96 metadata challenges

Los participantes de este proyecto aceptaron utilizar los mismo estándares de metadatos. Utilizamos Dublin Core con algunas políticas mas específicas. Sin embargo, hay gran variedad de interpretación de los estándares.

Al mismo tiempo, el consorcio deo entrar algunas colecciones que existian antes del proyecto y que no utilizaban los mismos estándares.

Slide 97 DC mapping

Aqui se puede ver un problema con las busquedas en un sitio central. Nosotros tenemos cinco campos de datos mapeados al elemento de sujeto. En nuestro propio sitio podemos construir una interfaz para buscar cada campo de dato distintamente.

Slide 98 local customized

Localmente, podemos crear las interfaces para poder buscar los terminos de distintos vocabularios. Pues, el hecho de que tenemos terminos de la biblioteca de congreso y terminos de TGM no es un problema para nosotros en nuestro sitio local. Tenemos distintos campos de datos para cada vocabulario.

Slide 99 Local customized

Aquí se ve una interfaz para buscar los términos de TGM. En el sitio del consorcio, este tipo de búsqueda no es posible porque todos los elementos mapeados al elemento de sujeto en Dublin Core se buscan juntos.

Slide 100 inconsistent search results

Además de los problemas con vocabularios distintos, hubo problemas cuando los asociados no interpretaron los estándares de la misma manera. Hubo resultados no consistentes por causa de la variedad de estándares y de la interpretación de estándares. Por ejemplo, aquí la búsqueda era solo para imágenes – pero se puede ver que la búsqueda muestra un texto. Eso fue a causa de la falta de estándares comunes entre los colaboradores. Gastamos mucho tiempo en entender, explicar, y corregir el problema.

Slide 101 challenges - staff

Otro reto es el entrenamiento o instrucción de nuestro personal. Ellos no sabían nada de colecciones digitales y nosotros aprendimos solo un poco antes de enseñarles. Sigue siendo un reto enorme para nosotros.

Slide 102 Name and subject

Cuando comenzamos a trabajar con imágenes, nos damos cuenta de que no sabemos nada de como describir imágenes. Utilizamos varios estándares y especialmente un artículo escrito por una colega que explica la complejidad de describir una imagen. Ahora tenemos nuestras políticas escritas para el entrenamiento de nuestro personal.

Slide 103 scanning procs

Aqui hay una pagina donde tenemos algunos procesos locales por la creación de imágenes digitales. Forma parte de nuestra interfaz publica también, porque queremos establecer confianza en nuestros procesos.

Slide 104 Digcol procs

Tenemos un capitulo de nuestro manual para el personal que contiene una variedad de políticas y procesos relativos a las colecciones digitales.

Slide 105 other challenges

Hay muchos otros retos en crear y mantener las colecciones digitales. No hay tiempo para examinarlos todos. Algunos otros seguir siendo informado de los cambios de la tecnología y de los estándares y los recursos – dinero, personal, equipo.

Slide 106 future

Nuestros planos para el futuro consiste en:

- Continuar con la creación de nuevas colecciones
- Entrenar mas personal en la digitalización y la colección de los metadatos
- Asegurar que tenemos mas personas que puedan manejar los proyectos
- Refinar nuestras interfaces para el publico
- Revisar y corregir los metadatos
- Mejorar la documentación de nuestros estándares y procesos
- Desarrollar un plano mejor para los fondos
- Llegar a ser un archivo digital fiado

Slide 107 spanish images

Algunos ejemplos de imágenes que tienen algo que ver con Espana en nuestras colecciones

Slide 108 Merchants' bank

Un edificio en una ciudad pequeña de los EEUU. Esta colección restringida tiene las imágenes del arte y de la arquitectura.

Slide 109 Prefontaine

El atleta mas famoso de la Universidad de Oregon.

Slide 110 warner

Una imagen de la colección de Gertrude Bass Warner.

Slide 111 furlong

Otra colección de fotos historicas.

Slide 112 ulmann

Otra colección com imágenes historicamente significante.

Slide 113 athletics and academy

Una foto de nuestra colección sobre el programa atletico de la universidad.

Llamamos esta foto - “The Meeting of the Minds” el encuentro de los espíritus

Slide 114 pres papers

Esta es una imagen de una manifestación estudiantil durante los años de la guerra en Vietnam. Forma parte de nuestra colección de los papeles de los presidentes de la universidad.

Slide 115 pres papers

Otra imagen de la misma colección.

Slide 116 hist photos

Una imagen de nuestra colección de fotos históricas.

El fotógrafo se llama Clarence Andrews, que murió en 1948. Esta foto es una que más me gusta porque es tan hermosa y también porque muestra claramente algo de una época particular.

Slide 117 hist photos

Muchos de los materiales que escogimos representan una época específica de nuestra historia local.

Aquí se ve a los indios de nuestra región pescando. Este sitio ahora está debajo del agua y no existe más. Nuestras colecciones nos permiten preservar una parte de nuestra historia y educar a nuestros ciudadanos al mismo tiempo.

Slide 118 athletics

Un cartel de las pruebas olímpicas para la pista en nuestra ciudad (es una ciudad famosa por los deportes de pista). Hay gente que dice que el jogging comenzó en la ciudad de Eugene.

Slide 119 athletics

Otro cartel de la misma colección. Vamos a tener las pruebas olímpicas para los deportes de pista otra vez – por eso hemos escogido estas imágenes.

Slide 120 western waters

Una foto de nuestra colección de las aguas del oeste. Tenemos muchas fotos hermosísimas y históricamente importante en esta colección.

Slide 121 realidad ilusión

En el mundo de las colecciones digitales que es la realidad y que es la ilusión? Aquí hay algunas ideas populares de estas colecciones.

Los administradores de mi biblioteca también creen en algunas de estas ideas.

- Muchas personas creen que son fáciles – y no son.
- Muchas personas creen que son baratas – y seguramente no son.
- Muchas creen que van a durar – y no duraran sin gran esfuerzo.
- Mucha gente cree que las colecciones digitales son completas o comprensivas – y solo Google tiene los recursos para crear una colección comprensiva.
- Mucha gente sigue creyendo que las colecciones digitales van a reemplazar los materiales originales. Lo dudo.

Lo que yo aprendí después de 3 años trabajando intensivamente con las colecciones digitales es que consisten en mucho más que tecnología.

CONT ...

Mas bien, consisten en las colaboraciones, en las colecciones de materiales significantes e interesantes, y en la gente que las utiliza. La tecnologia es solo un instrumento y nunca debe ser la fuerza que dirige el desarrollo de las colecciones digitales.

Slide 122 ack

Quería dar los agradecimientos a mi departamento y los otros departamentos que ayudan tanto con nuestras colecciones y algunas personas de mi departamento que juegan un rol tan importante en nuestras colecciones.

Slide 123 contact

Aquí es mi información de contacto. Me gusta mucho recibir preguntas y comentarios de colegas de todas partes.

Slide 124 viva espana

Ah, como podeis ver, a mi marido y a mi nos gusta mucho la España. El modo de vivir, la gente, el clima – todo.