

RedCLARA



Un proyecto presentado por RedCLARA. Financiado por el Fondo de Bienes Públicos Regional del Banco Interamericano de Desarrollo (BID). Con la participación de Argentina, Brasil, Chile, Colombia, Ecuador, México, Perú y Venezuela

Acta acuerdo comité técnico.

5ta Reunión LaReferencia.

Plan Piloto Componente III

Bogotá. 27 y 28 de Agosto de 2012.

Tabla de contenidos

[Acerca del esquema de metadatos acordado. Rol del validador y filtro del agregador.](#)
[Acerca de los metadatos obligatorios y adaptaciones automáticas a desarrollar ante la falta de alguno de ellos.](#)

[Acerca del metadato dc: type](#)

[Acerca del metadato dc:rights](#)

[Acerca del metadato dc:language](#)

[Acerca del metadato dc:description](#)

[Acerca del metadato dc:date](#)

[Acerca del metadato dc:format](#)

[Acerca del metadato dc:identifier](#)

[Acerca del metadato dc:source](#)

[Acerca del Workflow de registro, autenticación y cosecha de las redes nacionales.](#)

[Requerimientos deseables para el portal web de acceso público. Módulo Agregador.](#)

[Acerca de la modalidad de cosecha por sobre las redes nacionales.](#)

[Acerca de los TdRs especificados en el marco de la Consultoría del Piloto. Aclaraciones y especificación detallada de los mismos.](#)

[Diagnóstico en relación a las directrices. Estado técnico de las redes nacionales en vistas a la red regional](#)

[Acerca de la agenda recomendada de ejecución del Plan Piloto.](#)

[Conclusiones del equipo técnico de la revisión de algunas experiencias similares.](#)

[Experiencias de las redes nacionales en cuanto a agregadores / cosechadores.](#)

[Conexión con las otras componentes. Componente I del proyecto.](#)

[Documentación mínima exigible de los desarrollos a realizarse en el marco del plan piloto.](#)

1. Acerca del esquema de metadatos acordado. Rol del validador y filtro del agregador.

Siendo que el espíritu del acuerdo técnico debe basarse en la colaboración con las redes nacionales para la inclusión de la mayor cantidad de objetos digitales que resulten relevantes para LaReferencia, durante la ejecución del plan piloto es que se intentará salvar las diferencias con las directrices DRIVER en forma automática en los casos técnicamente posibles para aquellas redes que aún no las cumplan, destacando que el objetivo final es que las redes nacionales adopten la norma establecida y firmada en el acuerdo técnico de Bogotá. Por tanto, las redes nacionales deberán realizar estas adaptaciones (cuando las hubiera) antes del 30 de Marzo de 2013, resultando prioritaria la adopción de la normativa DRIVER para el metadato dc:type, dada la complejidad necesaria para su adaptación por parte de la red.

A fin de conseguir que las redes nacionales cumplan en un mediano plazo con los estándares de metadatos establecidos por la Componente, es que este Comité Técnico recomienda al Comité Estratégico implementar las acciones que se consideren necesarias en un corto plazo (firma de acuerdos, difusión, capacitación, etc.)

2. Acerca de los metadatos obligatorios y adaptaciones automáticas a desarrollar ante la falta de alguno de ellos.

Se acuerda contemplar los siguientes campos como obligatorios, de acuerdo a las pautas DRIVER 2.0 y a lo establecido en el Acta de Acuerdos con las siguientes consideraciones:

1. dc.title
2. dc.creator
3. dc.type
4. dc.date
5. dc.identifier

Se acuerda que en el caso de faltar uno o más de los campos precedentes, el registro resultará desechado del proceso de cosecha.

Acerca del metadato dc: type

En el caso especial de dc:type se acuerda que las redes nacionales proveerán los siguientes tipos documentales de acuerdo al acta suscripta en México por los Socios.

1. Artículos de Publicaciones Periódicas
2. Tesis doctorales
3. Tesis de postgrado

4. Informes Técnicos (de acuerdo al Acta de México, 14/11/2011)

Temporalmente y nivel de la red, se admitirán los mapeos automáticos (adaptaciones) para el metadato dc:type, evaluando para los casos apropiados

| Metadato | DRIVER | Términos admitidos |
|----------|---|--|
| dc:type | info:eu-repo/semantics/article | Artículo Artigo Journal Article Artículo de Revista |
| dc:type | info:eu-repo/semantics/ doctoralThesis | Tesis de Doctorado Tesis Doctoral Tese (Red Brasil) |
| dc:type | info:eu-repo/semantics/ masterThesis | Trabajo de grado (Red Venezuela) Tesis de Maestría Dissertacao (Red Brasil) |
| dc:type | info:eu-repo/semantics/report | Informe Técnico Relatório Técnico |

En caso de no incluir una segunda ocurrencia de dc:type, se asumirá el valor info-eu:repo/semantics/published.

Todos los registros que resulten cosechados y alguna ocurrencia del campo dc:type no se corresponda con la forma normalizada o alguno de los términos de las variantes, será desechado, quedando bitácora de este proceso la cual podrá ser revisada por el perfil gestor de red nacional.

Acerca del metadato dc:rights

Al igual que DRIVER 2.0, tiene carácter de recomendable, y se especifica adicionalmente que los valores del campo, respeten las directrices OpenAire para este metadato:

- info:eu-repo/semantics/closedAccess
- info:eu-repo/semantics/embargoedAccess
- info:eu-repo/semantics/restrictedAccess
- info:eu-repo/semantics/openAccess

Se recomienda, asimismo, el uso de la segunda ocurrencia del metadato especificando la URL de la licencia, como en el caso de creative commons, o cualquiera de los ejemplos de DRIVER 2.0 en relación al uso de la misma.

Acerca del metadato dc:language

Conserva el carácter Recomendado, como en las directrices DRIVER, indicándose adicionalmente que se aceptarán codificaciones ISO con dos o tres caracteres y siempre resultará convertido a ISO-639-3 caracteres como las directrices recomiendan por el software agregador de los registros cosechados. Se eliminarán las ocurrencias de dc:language que no correspondan con valores admitidos de ISO-639-2, imposibles de mapear con ISO-639-3 a partir del 30/3/2013.

Acerca del metadato dc:description

En el caso de dc:description, el mismo no será controlado pero resulta recomendable que las redes nacionales envíen los resúmenes en el idioma original e inglés, siendo que será responsabilidad de la componente agregador de LaReferencia, exponer el mismo en inglés en el orden especificado por las directrices DRIVER 2.0.

Acerca del metadato dc:date

Primer paso: Para todas las fechas se verificará el cumplimiento de la norma ISO-88601. Las fechas que no superen dicha verificación serán eliminadas del registro a incluir en el cosechador.

Segundo paso: Si luego del primer paso, el registro aún conserva fechas, se obtendrá la menor de ellas y se asumirá la misma como fecha de publicación del objeto digital. Si el registro no conserva fechas, no resultará cosechado.

Para ambos pasos de procesamiento quedará bitácora de las decisiones tomadas por el proceso.

Acerca del metadato dc:format

Se acuerda que a partir del uso del metadato dc:format, las soluciones a seleccionar pueden brindar servicios relacionados con la problemática de preservación digital. En consecuencia, se

recomendará su uso y cuando esté presente, su adhesión a lo indicado por las directrices. En caso de no estar incluido en el vocabulario controlado, se emitirá una alerta y se eliminará la ocurrencia.

Acerca del metadato dc:identifier

En cuanto a dc:identifier, se acuerda verificar que alguna de sus ocurrencias contenga una URL bien formada, en caso contrario, el registro será descartado. En caso de hallar alguna URL bien formada, la misma será verificada invocando esa URL y si no devuelve un código HTTP 200 se generará una alerta de validación. Este proceso de validación no garantiza la existencia real del objeto digital, el cual se asume como responsabilidad de las redes nacionales y los repositorios que las componen.

Acerca del metadato dc:source

Las redes nacionales enviarán una identificación del repositorio y la institución nacional al cual pertenece cada registro y el portal web público de LaReferencia posibilitará la búsqueda, navegación y faceta por repositorio. En el caso que la red no envíe la información del repositorio, la faceta presentará el total por país.

La información del nombre del repositorio y de la institución, será enviada a través del campo dc source de la siguiente forma:

`<dc:source>instname:Nombre de la Institución</dc:source>`

`<dc:source>reponame:Nombre del repositorio</dc:source>`

Se recomienda registrar las distintas instancias institucionales de las cuales dependa el Repositorio respetando la jerarquía de mayor a menor, separando cada instancia por un punto seguido de un espacio.

3. Acerca del Workflow de registro, autenticación y cosecha de las redes nacionales.

El perfil gestor de la red nacional solicitará su registro en la solución Validador de la Red La referencia. Una vez aprobado el mismo, podrá ingresar con una clave y password que será enviada, usando la misma podrá acceder a la plataforma Validador de la red LaReferencia y recuperar los informes generados por el mismo.

Cuando se considere adecuado se solicitará la incorporación de la red nacional a la solución Agregador de la red LaReferencia.

Una vez registrada la red nacional en la plataforma del agregador, la misma resultará cosechada y listada en el portal web de LaReferencia entre las redes nacionales cooperantes, aún cuando su número de registros validados luego de la aplicación de los filtros sea 0 registros.

4. Requerimientos deseables para el portal web de acceso público. Módulo Agregador.

Los siguientes requerimientos serán considerados por el Consultor para su desarrollo, en caso de no resultar posible concretarlos durante el plazo de la presente consultoría, el Comité Técnico considera que no deben ser desechadas como ideas para el futuro del portal.

- En función de mejorar la experiencia del usuario ofrecida por el servicio del portal web de LaReferencia, sería deseable que exista la posibilidad de verificar la disponibilidad del recurso digital presentado como resultado de la consulta para filtrar o señalar al usuario final aquellos registros cuyo objeto digital no esté disponible en el momento de la consulta (por ejemplo ante la caída de un servidor local).
- Resulta de interés la generación de un directorio de Repositorios a nivel regional, el cual podría ser referenciado (a través de una identificación única) en los metadatos que son cosechados. Este directorio podría brindar diferentes servicios orientados a los perfiles de gestor de repositorios, tomadores de decisión y usuarios finales.
- En caso que el registro contenga una segunda ocurrencia de dc:type con valor info:eu-repo/semantics/published y la primer ocurrencia sea info:eu-repo/semantics/article el campo dc:relation resulta mandatorio, de no cumplirse esta regla se generará una alerta del validador.
- Resulta deseable contar con una aplicación Facebook la cual consuma un webservice de consulta de la plataforma agregador y permita revisar en forma paginada y sin facetado los resultados de una consulta. Estudiar las posibilidades de usar los restantes servicios de Facebook, por ejemplo la posibilidad de publicar en un muro una referencia a un objeto digital, o por ejemplo indicar Me Gusta o publicar en el muro de otro usuario un resultado específico de una consulta.

5. Acerca de la modalidad de cosecha por sobre las redes nacionales.

El cosechador de la red verificará si el set LaReferencia (setSpec lareferencia) existe en la red nacional. En ese caso realizará cosecha sobre el mismo, caso contrario lo hará sobre el agregador nacional completo aplicando las reglas de filtrado de tipos de material (ver punto 2).

Las redes nacionales serán cosechadas preferentemente en forma incremental de acuerdo a lo especificado en el protocolo OAI-PMH a lo largo del proyecto. En caso de interrupciones en el proceso de cosecha por algún tipo de error HTTP, se realizarán tres intentos de restablecimiento, separados entre ellos por 2 minutos. Al finalizar el tercer intento sin éxito, es que se marcará la cosecha como fallida y se conservará el conjunto de datos de la red a como estaba con anterioridad a la cosecha.

Los agregadores de las redes nacionales respetarán en forma progresiva las pautas especificadas por DRIVER 2.0 en lo relacionado con el tamaño de paginación de los resultados, los elementos que componen el resumptionToken y lo relacionado con los registros eliminados, posibilitando la cosecha incremental.

El agregador de LaReferencia expondrá los metadatos en DRIVER 2.0, organizados de acuerdo a un set por red nacional (el setSpec será el nombre del país con los espacios reemplazados por guión bajo, sin acentos y en minúsculas). El agregador de la red regional respetará las pautas DRIVER 2.0 en lo relacionado a tamaño de paginación de los resultados, registros eliminados y elementos que componen el resumption token. A su vez, implementará los elementos friends en el verbo Identify, apuntando a cada una de las redes nacionales.

6. Acerca de los TdRs especificados en el marco de la Consultoría del Piloto. Aclaraciones y especificación detallada de los mismos.

- La funcionalidad especificada en los TdRs relacionada con la búsqueda semántica se especifica como deseable y se refiere, como una de sus posibilidades al proyecto Intelligo. Como resultó acordado en la reunión, la representante técnica de Argentina (Paola Azrilevich) ha mantenido una reunión con los responsables del proyecto intelligo (de la Organización de Estados Iberoamericanos-OEI) de la cual se puede desprender lo siguiente:
 - La Organización de Estados Iberoamericanos para la Educación, la Ciencia y la Cultura (OEI) se halla interesada en colaborar con La Referencia. A fin de establecer un lazo formal con la Red, se propone la firma de un acuerdo para el uso del producto de software intelligo, incorporando los metadatos producto de cosecha de la red, sin costo económico para LaReferencia, siempre que no se requiera el desarrollo de nuevas funcionalidades sobre el mismo. Podría realizarse la firma de este acuerdo en la Reunión de Altas Autoridades de noviembre.
 - El plan piloto de la componente III puede generar los insumos para la producción de los mapas conceptuales de intelligo. Es de destacar que solo serán incluidos en el procesamiento realizado por intelligo aquellos registros recuperados por los procesos de cosecha que incluyan en sus metadatos dc:description.
 - El consultor de la componente III se compromete a obtener en la nueva corrida de diagnósticos, a partir de la fecha límite de designación de repositorios muestra, una proporción de los registros que poseen ese metadato (dc:description) faltante.
 - Los métodos semánticos de intelligo solo resultan viables ante una muestra de volumen de importancia superior a 10.000 registros, esto implica que los mapas conceptuales sólo podrán ser discriminados por país.
 - La implementación de intelligo dependerá fuertemente de la configuración del hosting que el proyecto pueda disponer para el mismo (principalmente memoria y espacio en disco del servidor) de modo de posibilitar el procesamiento y la carga ante consultas de volumen medio-alto.
 - En caso de resultar deseable realizar adaptaciones y/o mejoras al desarrollo de intelligo, específicas para LaReferencia, habría que analizar su costo y posibilidad de financiamiento por parte de la Red.
- El apartado relacionado con la generación de Comunidades se aclara que consiste en una funcionalidad requerida para la plataforma agregador que permita la creación de usuarios y grupos de interés. Los usuarios pueden pertenecer a varios grupos de interés. Los grupos pueden ser creados libremente por los mismos usuarios. Se solicitará el nombre del grupo, su descripción y se podrán asociar tags, bajo el mismo

concepto de tagging de los registros cosechados.

Los grupos poseen como servicios, la recepción de un alerta de material enviado por un usuario, la inclusión y exclusión de usuarios, es decir que un usuario puede suscribirse a un grupo o des-suscribirse libremente. Todos los integrantes de un grupo recibirán un correo electrónico ante el evento de alerta (novedad).

- La frase de los TdRs referida a:

“...Debe permitir compartir metadatos por correo electrónico a usuarios registrados en la plataforma. Las citas serán generadas a partir de los metadatos en los formatos Dublin Core XML, BibTex, otros...”

Se refiere a que solo un usuario registrado puede enviar los metadatos a un destinatario de correo (registrado o no, solo se ingresa la dirección de correo) a través formulario de la plataforma. El envío de dicha recomendación será registrado y contabilizado, siendo que entre los requerimientos para la infraestructura de IT se contará la inclusión de un mailer.

- Un usuario registrado podrá enviar registros a modo de novedad a uno o mas de los grupos a los que se halla suscripto. Asimismo un usuario puede suscribirse a un subject en particular y con esa operación recibirá por parte de la plataforma un alerta por correo electrónico cuando ingrese un material (metadatos) que lo incluya.
- Respecto al apartado que indica Servicios web 2.0, se acuerda someter las implicancias potenciales del uso de estas herramientas al grupo de gestión.
 - Los comentarios sobre los registros serán realizados sólo por los usuarios registrados en el portal web público. El consultor se compromete a revisar herramientas que puedan filtrar en forma automática comentarios inapropiados.

Caso que dichas herramientas no resulten eficaces, el grupo técnico presenta dos propuestas:

- Cada red se haga cargo de la moderación de los mismos.
- Eliminar los comentarios como funcionalidad.

Los comentarios brindan beneficios directos relacionados con determinar las opiniones y valoraciones de otros usuarios sobre los documentos.

Otros beneficios posteriores se relacionan con la aplicación de tecnologías de text-mining para la generación de servicios que, por ejemplo, permitan extraer

conocimiento implícito de las comunidades de usuarios. Siendo que esta última se halla fuera del alcance del presente plan piloto.

- Los tags debe ofertarse con una modalidad de autocomplete para intentar unificar los mismos y generar una taxonomía lo mas controlada posible.
- El front-end público no considerará generar un wiki interno a la plataforma.
- En cuanto al apartado de Estadísticas.
 - Las estadísticas / indicadores que serán provistas se diferencian por perfil de usuario

- Caso perfil de usuario público

Indicadores

- Número de redes involucradas / años
- Número de repositorios / año
- Número de objetos digitales en general / años.
- Número de objetos digitales por type (acorde DRIVER) / años
- Nube de subjects (hasta 100) / años
- Distribución de dc:rights primer ocurrencia.
- Licencias más frecuentes acorde a dc:rights segunda occ.

Estadísticas

- dc:type Material mas consultado por tipo / años. Definiéndose consultado como el accionamiento de la url por parte del usuario.
- Exploración por registros mas consultados.
- Exploración por registros mas posteados en redes sociales.
- Distribución de usuarios por país que consultan los objetos digitales identificados por IP / años.
- Distribución de países cuyos objetos digitales resultan consultados / años.

- Caso Perfil gestor de redes.

Indicadores

- Número de cosechas exitosas / Número de cosechas fallidas / años. Solo esta red.
- Número de registros filtrados general por agregador / años. Solo esta red.

- Número de registros filtrados por metadato que no cumplimenta norma / años. Solo esta red.
- Número de registros automáticamente mapeados general / años. Solo esta red.
- Número de registros automáticamente mapeados por metadato / años. Solo esta red.

Estadísticas

- dc:type Material mas consultado solo de esta red por tipo / años
- Exploración por registros mas consultados solo de esta red
- Exploración por registros mas posteados en redes sociales solo de la red.
- Distribución de usuarios por país que consultan los objetos digitales identificados por IP / años. Solo esta red.
- Distribución de países cuyos objetos digitales resultan consultados / años. Solo esta red.

El perfil de administrador de la plataforma contará con las estadísticas públicas, las que permiten generar la visión de red. Sin embargo podrá producir informes específicos relacionados con:

- Número de cosechas exitosas / Número de cosechas fallidas / años. Todas las redes. Selección de una red en particular.
 - Número de registros filtrados general por agregador / años. Todas las redes. Selección de una red en particular.
 - Número de registros filtrados por metadato que no cumplimenta norma / años. Todas las redes. Selección de una red en particular.
 - Número de registros automáticamente mapeados general / años. Todas las redes. Selección de una red en particular.
 - Número de registros automáticamente mapeados por metadato / años. Todas las redes. Selección de una red en particular.
- El validador para las redes nacionales será accesible a través de la herramienta de back-end, es decir que la red nacional se autentica en dicha plataforma y solicita el proceso de validación, el cual adicionalmente al cumplimiento de las directrices DRIVER 2.0 indicará las adaptaciones (mapeos automáticos) convenidos en el primer punto del presente documento. Esta herramienta de validación será instanciada en la plataforma de IT dispuesta por la RedCLARA para el presente Piloto.

- En el apartado que se indica:
“...Opcionalmente debe ofrecer la posibilidad de realizar autoregistro de repositorios previa aprobación del administrador de la plataforma.
Proceso de subida/bajada de la plataforma automática. Procesos de backup/restore del contenido del portal...”

Se ha considerado que dado que lo que el Piloto administrará son redes nacionales, no resulta necesaria la funcionalidad de auto-registro. El perfil de usuario gestor de la plataforma realizará el proceso de alta de la red a partir de una comunicación de la misma. Podrá definir a partir de que momento, la misma ingresa al proceso de cosecha del agregador, pero independientemente de dicho proceso, la red contará con las facilidades de autenticación para usar el servicio del validador de la Red.

Cuando la red considere oportuna su inclusión en el proceso de cosecha (acorde a los resultados del proceso de validación), así lo comunicará al administrador de la plataforma, quien la incluirá en el proceso de cosecha, publicación de sus metadatos en el portal público y transformación/exposición a DRIVER 2.0 en la modalidad OAI-Data Provider.

En cuanto a los procesos de backup-restore del contenidos del portal, la factibilidad de los mismos será acorde a las tecnologías que se usen, pero no se considerará la posibilidad de realizar una descarga a través de un cliente web del contenido del portal, ni el upload/restore de una imagen a través de un cliente web. Dichas operaciones deberán ser cronológicas, automáticas e inscriptas en la gestión tecnológica que la plataforma de IT provista se halla incluida.

7. Diagnóstico en relación a las directrices. Estado técnico de las redes nacionales en vistas a la red regional

Las redes nacionales tendrán como plazo máximo el 7 de Setiembre de 2012 informar la fecha en la cual el agregador a nivel nacional será cosechable, el cual tiene como plazo máximo de 60 (sesenta) días.

Inmediatamente y en función de posibilitar la generación de los diagnósticos en relación al cumplimiento de las directrices, las redes informarán un máximo de 3 (tres) repositorios considerados como muestra válida solo hasta que el agregador a nivel nacional resulte cosechable. El plazo máximo para enviar esta información al Consultor de la componente III es el 7 de Setiembre de 2012.

A la fecha de cierre de la presente acta se ha recibido información para continuar con los procesos diagnósticos de las siguientes redes nacionales:

- Argentina
- Brasil
- Colombia
- Ecuador
- El Salvador
- México
- Perú
- Venezuela

8. Acerca de la agenda recomendada de ejecución del Plan Piloto.

Se demarca la necesidad que el desarrollo/implementación comience con la implementación / customización del módulo de validación de las directrices orientado a las redes nacionales.

De acuerdo a los hitos marcados en la propuesta aprobada:

1. Relevamiento de requerimientos. 40 (cuarenta) días a partir del inicio de la propuesta. Finalizaría este plazo el 10 de Setiembre de 2012.
2. Evaluar y obtener infraestructura de IT apta considerando los máximos requerimientos de las soluciones disponibles. 15 (quince) días a partir de 1. Este plazo finalizaría el 25 de Setiembre de 2012.
3. Evaluación y selección de las soluciones posibles a resultar implementadas. 40 (cuarenta) días a partir de 1. Este plazo finalizaría el 20 de Octubre de 2012.
4. Conformación del equipo de desarrollo. Esto involucra confeccionar los TdRs, convocar al llamado y poner en producción al/los desarrolladores seleccionados. Se destaca que el grupo Técnico ya ha colaborado en los comentarios de las versiones preliminares de los mismos.

9. Conclusiones del equipo técnico de la revisión de algunas experiencias similares. Experiencias de las redes nacionales en cuanto a agregadores / cosechadores.

En base a las revisiones de casos recabados por el Consultor, considerando la infraestructura tecnológica en cuanto a servidores disponibles, el grupo Técnico aconseja:

1. En cuanto al licenciamiento: Se recomienda que los desarrollos a realizarse en el marco del presente proyecto, se enmarquen en un esquema de licenciamiento de software libre y el código fuente resulte accesible a través de repositorios públicos tales como github o sourceforge.
2. En cuanto a replicación/colaboración: Los esfuerzos de desarrollo del agregador/ cosechador deben resultar replicables para las redes nacionales, de forma tal de facilitar el proceso de creación de la misma. La plataforma debe exponer en su rol de agregador los metadatos cosechados en DRIVER 2.0. En base a esto y en el proceso intermedio de los desarrollos nacionales el consultor se compromete a mantener un seguimiento del avance de los desarrollos, los cuales son volcados en el presente documento y serán mantenidos/actualizados en la wiki, en función de poder contribuir para brindar una herramienta a la conformación de futuras redes, reaprovechar las experiencias y tendencias en el uso de tecnologías de cada país.
3. Los desarrollos que implementan los procesos de cosecha a nivel nacional en cada uno de los países, al cierre de la presente acta, son los siguientes:
 - a. Argentina. DRIVER D-NET¹. Comunidad de desarrollo:CEE (Comunidad Económica Europea. DRIVER). Tecnologías: Java, Arquitectura SOA, NOSQL Db (Couch DB), Apache Derby, Apache Solr (indexador y buscador).
 - b. Brasil. Portal OASIS. Portal agregador de redes brasileiras de acceso abierto. Implementada con el software libre REPOX². Base de datos PostGreSQL.
 - c. Colombia. Desarrollo propio BDCOL. Desarrollo realizado en Java + XML Db (Exist)
 - d. Ecuador. Uso de OpenHarvester (PKP).
 - e. México. Desarrollo propio denominado REMERI. realizado en Java + XML Db (Exist). Tecnología de indexación Apache Lucene.

¹ D-NET. Consorcio de desarrollo herramienta D-NET: <http://www.d-net.research-infrastructures.eu/>. Portal de backend/público/cosechador/agregador/mapeo de metadatos.

² REPOX. <http://rebox.ist.utl.pt/index.html>. Framework con funcionalidades de cosechador/agregador y mapeo de metadatos.

f. Venezuela. Uso de OpenHarvester (Public Knowledge Foundation)³.
PHP+MySql/PostgreSQL.

4. Dado que los casos de desarrollos propios cuentan en cada uno de los países con equipo de RRHH propio en función del mantenimiento de los mismos, y dado que el agregador/cosechador regional solo contará con esfuerzos destinados a su mantenimiento (al menos en esta instancia), es que el grupo técnico recomienda la implementación de las modificaciones requeridas sobre soluciones pre-existentes, basado en proyectos de software libre con un buen nivel de actividad y detallar en forma específica las adaptaciones realizadas (ver documentación) con la finalidad de posibilitar/facilitar la actualización de las versiones y minimizar el costo de traslado de las modificaciones realizadas entre versiones sucesivas de la plataforma.

³ PKF. Open Harvester. <http://pkp.sfu.ca/?q=harvester>. Desarrollo de portal e interfaz de back-end con funcionalidad de cosechador/agregador/portal web público.

10. Conexión con las otras componentes. Componente I del proyecto.

En relación a la integración con la componente I, en el punto 3.3.2 de los TdRs del piloto se especifica:

3.3.2 Diseño de la plataforma de información y de la estrategia de desarrollo de esta.

* Definir el diseño tecnológico que permita la Adaptación y desarrollo de los requerimientos funcionales y no funcionales de la plataforma de integración de la red federada.

* Integración con el portal Web del proyecto producto del componente I.

* Generación de manuales de: instalación, administración, uso y de arquitectura técnica del software.

Será entonces objetivo de este ítem de especificación de requerimientos, definir cual es el ámbito del término integración a partir del conocimiento que se tiene de la herramienta de base para el sitio de LaReferencia, el gestor de CMS, Drupal.

La propuesta presentada y redactada por el consultor del plan piloto especifica

“...La integración de las plataformas web de ambas componentes podrían acoplarse a través de un micrositio descriptivo y link que permita navegar de uno al otro portal, o podría buscarse una integración mas compleja basada en el consumo de una API o un protocolo de consulta tal como Open Search, de forma tal que pudieran incluirse en los resultados de búsqueda del portal CMS, los resultados de la misma consulta sobre la interfaz web del agregador de objetos digitales de las redes de repositorios...”

La inclusión de los objetos digitales cosechados en la búsqueda del sitio Drupal de LaReferencia podrá implementarse, en el caso de resultar considerado de utilidad por los grupos de Gestión o Técnico a través del consumo del protocolo de interoperabilidad para búsquedas OpenSearch. Será misión del componente III, preparar y dejar implementado dicho protocolo, sin acción específica relacionada con su consumo desde el portal de LaReferencia.

11. Documentación mínima exigible de los desarrollos a realizarse en el marco del plan piloto.

Documentación técnica

- Arquitectura de la solución. Principales módulos y su descripción.
- Especificación detallada de las modificaciones realizadas (en caso apropiado) sobre la/s solución/es estándar descargable. Informe de contacto con el equipo técnico de la misma para integraciones.
- Diagramas UML. Casos de uso de la solución considerados.
- Diagramas UML. Esquema de clases de los principales módulos.
- Diagramas E-R en los casos de almacenamientos en bases de datos relacionales.
- javadoc en caso que la solución esté realizada en Java.
- Wiki en la comunidad La Referencia relacionado con el proceso de inception de los desarrolladores en el mismo. Ejercicios, casos comunes de resolución.
- LDAP en el caso de ser usado. Versión y tecnología. Campos usados en el mismo.

Deploy y Administración

- Instrucciones de instalación en plataformas Linux mas comunes. Wiki de instalación de la plataforma y sus productos de base.
- Interrupción del servicio. Casos hallados en la puesta en producción y las soluciones halladas.
- Políticas de backup tomadas por la instanciación de producción. Casos en los que ha sido necesario recurrir a ellos.
- Fine-tuning de las herramientas de base, ej: optimización de Apache Solr, compresión y/o mantenimiento de las bases de datos, etc.
- Buenas prácticas recomendadas.

Manual del Administrador de la plataforma

- Procesos de creación/mantenimiento de redes objeto de cosecha.
- Configuración de la cosecha sobre cada red, completa, incremental. Lanzar manualmente una cosecha.
- Definición de usuarios / passwd y derechos de acceso a información de una red en particular.
- Definición de mapeo automático de valores de metadatos.
- Opciones de reportes.