a Convenção de Santa Fé para a Iniciativa de Arquivos Abertos

Lançado oficialmente em 15 de fevereiro de 2000

A Convenção de Santa Fé é descontinuada . Em vez disso , use o protocolo Open Archives Initiative para coleta de metadados .

Bem-vindo à Convenção de Santa Fé. Esta convenção é o resultado de uma reunião da Open Archives Initiative que foi realizada em Santa Fé, Novo México, de 21 a 22 de outubro de 1999. Esta convenção foi endossada por unanimidade por todos os participantes da reunião, que representaram organizações que mantêm ou planejam Arquivos de impressão eletrônica destinados ao acesso aberto e organizações interessadas em fornecer serviços, como interfaces de pesquisa ou links de citações, com base nos dados desses arquivos. A convenção apresenta uma estrutura técnica e organizacional simples para apoiar a interoperabilidade básica entre arquivos de impressão eletrônica. Os participantes manifestaram a intenção de implementar este quadro para permitir experiências de interoperabilidade no decurso do ano 2000. Os mantenedores de arquivos de impressão electrónica existentes ou futuros que não estiveram representados na reunião são fortemente encorajados a juntarem-se a este esforço, implementando o quadro para seus arquivos.

Objetivo

Autores acadêmicos podem disponibilizar documentos eletrônicos para um público global, enviando-os para arquivos impressos eletronicamente. No momento em que esta convenção foi escrita (janeiro de 2000), o número de arquivos de impressão eletrônica estabelecidos era pequeno. Um deles é o arXiv.org, o arquivo eletrônico de Los Alamos criado por Paul Ginsparg, que se tornou um centro crucial para a comunicação de resultados de pesquisas em física. Prevemos a criação de muito mais arquivos de impressão eletrônica nos próximos anos. Esses arquivos serão distribuídos pela Internet; alguns serão orientados para disciplinas acadêmicas específicas, outros baseados em afiliação institucional. Para que os arquivos impressos eletrónicos se tornem um mecanismo estabelecido para a comunicação académica, é necessário apoiar algum nível de interoperabilidade entre eles. Esta convenção é um primeiro passo para essa interoperabilidade. À medida que a nossa experiência com o assunto se desenvolve, mais questões poderão ser abordadas, novas convenções poderão ser criadas e as existentes poderão ser melhoradas. A Open Archives Initiative fornece um processo de crescimento e desenvolvimento.

A Convenção de Santa Fé apresenta uma estrutura técnica e organizacional projetada para facilitar a descoberta de conteúdo armazenado em arquivos distribuídos de impressão eletrônica. Ele faz recomendações técnicas fáceis de implementar para arquivos que – quando implementadas – permitirão que os dados dos arquivos de impressão eletrônica se tornem amplamente disponíveis por meio de sua inclusão em uma variedade de serviços ao usuário final, como mecanismos de pesquisa, serviços de recomendação e sistemas de interligação. documentos. Além disso, a convenção introduz um quadro organizacional para disponibilizar informações sobre arquivos que aderem às recomendações técnicas desta convenção e sobre partes confiáveis que criam serviços de usuário final para dados provenientes de tais arquivos.

Convidamos você a implementar as recomendações feitas nesta convenção para que seu arquivo contribua para a criação de um mecanismo de comunicação científica mais eficaz.

Documentos que compõem a Convenção de Santa Fé

A Convenção de Santa Fé incorpora os seguintes documentos:

- o documento central da Convenção de Santa Fé [http://www.openarchives.org/sfc/sfc.htm]
- o conjunto de metadados de arquivos abertos [http://www.openarchives.org/sfc/sfc_oams.htm]

- o subconjunto Dienst de arquivos abertos [http://www.openarchives.org/sfc/sfc_dienst.htm]
- o modelo a ser usado pelos provedores de dados para se registrar como um arquivo compatível com Santa Fe [http://www.openarchives.org/sfc/data_provider_template.htm]
- o modelo a ser usado pelos provedores de serviços para se registrar como um serviço compatível com Santa Fé [http://www.openarchives.org/sfc/service_provider_template.htm]]
- a lista de formatos de metadados usados no contexto do Open Archives [http://www.openarchives.org/sfc/sfc_metadata.htm]
- a lista de provedores de dados Open Archives [http://www.openarchives.org/sfc/sfc_archives.htm]
- a lista de provedores de serviços de Open Archives [http://www.openarchives.org/sfc/sfc_services.htm]
- a página introdutória da Convenção de Santa Fé - este documento [http://www.openarchives.org/sfc/sfc_entry.htm]



entre em contato com a Open Archives Initiative entrando em contato com openarchives@openarchives.org

última atualização em 20 de janeiro de 2001

a Convenção de Santa Fé : o documento central

A Convenção de Santa Fé é descontinuada . Em vez disso , use o protocolo Open Archives Initiative para coleta de metadados .

Este documento principal apresenta uma abordagem passo a passo para fazer com que seu arquivo de impressão eletrônica ou seu serviço esteja em conformidade com a Convenção de Santa Fé. Para esclarecer as diretrizes de Santa Fé, alguns conceitos subjacentes são introduzidos primeiro. Os detalhes técnicos estão em documentos separados, cujos links são fornecidos no texto e em uma lista no final deste documento .

Contorno

- Conceitos introdutórios
  - Arquivos e arquivos de impressão eletrônica abertos
  - Arquivos de impressão eletrônica gerenciados
  - Provedores de dados e provedores de serviços
  - Dados em um arquivo de impressão eletrônica
- Para o provedor de dados: como fazer com que seu arquivo e-print cumpra a Convenção de Santa Fé?
  - Etapa 1: Escolha um identificador exclusivo para seu arquivo de impressão eletrônica
  - Etapa 2: use identificadores persistentes exclusivos para dados no arquivo
  - Etapa 3: Implementar o conjunto de metadados de arquivos abertos
  - Passo 4: Implemente e documente outros formatos de metadados suportados pelo seu arquivo e-print
  - Etapa 5: Implementar a interface de colheita Dienst
  - Etapa 6: Informe a iniciativa Open Archives de que seu arquivo de impressão eletrônica está aberto
- Para o prestador de serviços: como fazer com que seus serviços cumpram a Convenção de Santa Fé?
  - Etapa 1: mantenha os identificadores originais em seus serviços
  - Passo 2: Cumpra as restrições de uso especificadas pelos provedores de dados
  - Passo 3: Informe a iniciativa Open Archives que você desenvolveu um serviço baseado em dados de arquivos abertos

Conceitos introdutórios

Arquivos e arquivos de impressão eletrônica abertos

Consideramos os seguintes componentes cruciais de um arquivo de impressão eletrônica:

- Um mecanismo de submissão;
- Um sistema de armazenamento de longo prazo.

Além disso, consideramos crucial que um arquivo de impressão eletrónica seja aberto, incorporando um mecanismo que permita a terceiros recolher dados do arquivo. Tal mecanismo permite que terceiros criem serviços de usuário final que apoiam a descoberta, apresentação e análise de dados no arquivo. Reconhecemos que a maioria dos arquivos de impressão eletrônica também fornecerá serviços ao usuário final. Contudo, consideramos que facilitar a ampla disseminação de dados de arquivo através de serviços de terceiros é uma característica crucial de um arquivo de impressão eletrónica.

Arquivos de impressão eletrônica gerenciados

Também assumimos que os arquivos de impressão eletrônica são gerenciados. Isto significa que têm algum tipo de política no que diz respeito à apresentação de documentos e também uma política no que diz respeito à preservação e retenção de documentos.

Provedores de dados e provedores de serviços

Consistente com o objetivo da Convenção de Santa Fé e com a identificação das funções cruciais de um arquivo de impressão eletrônica, fazemos uma distinção entre provedores de dados e prestadores de serviços.

- Um *provedor de dados* é o gerente de um arquivo de impressão eletrônica, agindo em nome dos autores que enviam documentos para o arquivo. Tal como referido acima, o fornecedor de dados de um arquivo aberto irá, pelo menos, fornecer um mecanismo de submissão, um sistema de armazenamento a longo prazo e um mecanismo que permita a terceiros recolher dados do arquivo.
- Um *provedor de serviços* é um terceiro que cria serviços para o usuário final com base em dados armazenados em arquivos de impressão eletrônica. Por exemplo, um prestador de serviços poderia implementar um motor de busca para impressões electrónicas matemáticas armazenadas em arquivos em todo o mundo.

Dados em um arquivo de impressão eletrônica

Um arquivo pode armazenar metadados que descrevem o conteúdo completo sem armazenar o conteúdo completo em si. Neste caso, consideramos os metadados como um registro. No entanto, assumimos que se o conteúdo completo for armazenado, sempre haverá metadados associados armazenados no arquivo, bem como um mecanismo para unir metadados e conteúdo. Neste caso consideramos a combinação de metadados e conteúdo completo como um *registro* .

Nesta convenção, portanto, definimos um arquivo como uma coleção de registros. Esses registros possuem as seguintes propriedades:

- Um registro em um arquivo e-print contém, pelo menos, metadados que descrevem o conteúdo completo;
- Um registro em um arquivo e-print também pode conter conteúdo completo, como um artigo de pesquisa, um conjunto de dados, software, etc., ou um pacote destes.

Para o provedor de dados: como fazer com que seu arquivo e-print cumpra a Convenção de Santa Fé?

Etapa 1: Escolha um identificador exclusivo para seu arquivo de impressão eletrônica

Para apoiar a interoperabilidade, cada arquivo deve ter um *identificador de arquivo exclusivo* . Este identificador refere-se à autoridade que gere o arquivo ou à iniciativa de arquivo. Escolha um identificador que consista em caracteres alfanuméricos [az, AZ, 0-9]. Para ter certeza de que o identificador do seu arquivo não coincide com o de outro arquivo, verifique a lista de identificadores existentes em http://www.openarchives.org/sfc/sfc_archives.htm . Formalmente, o caso dos caracteres no identificador do arquivo é significativo; no entanto, os identificadores devem ser selecionados para serem distintos, independentemente do caso.

- *Ao criar um arquivo aberto para a Universidade de Spa na Bélgica, pode-se escolher BESPA como identificador do arquivo.*

- *A iniciativa NCSTRL existente poderia escolher o NCSTRL, enquanto a iniciativa RePEc poderia optar pelo RePEc.*

Etapa 2: Use identificadores persistentes exclusivos para registros no arquivo

Os registros no seu arquivo devem ter identificadores persistentes exclusivos. Cabe a você tomar decisões acertadas sobre sua estrutura e geração.

Quando você combina um *identificador de arquivo exclusivo* e um *identificador de registro exclusivo* para um registro no seu arquivo, o resultado é um *identificador completo* para um registro no seu arquivo, que nunca coincidirá com um identificador completo de um registro em outro arquivo. Isso facilita muito o trabalho dos prestadores de serviços. Escolha um caractere não alfanumérico imprimível como *separador* para delimitar o identificador de arquivo do identificador de registro.

- *O arquivo da Universidade de Spa atribuirá identificadores significativos aos metadados submetidos ao arquivo. Suponhamos que esses identificadores consistam no corpo docente do primeiro autor seguido por um número de sequência que contém informações de data. Portanto, um identificador de registro pode ser MEDICINE/19991104/012. Se este arquivo escolher um hífen para delimitar o identificador do arquivo do identificador do registo, o identificador completo do registo seria: BESPA-MEDICINE/19991104/012.*

- *arXiv.org usa identificadores de registro que começam com o nome de um subarquivo seguido por um número de sequência sensível à data. Por exemplo, fisica/9811004 e hep-th/9909044 são identificadores válidos neste arquivo. Nesse caso, usando dois pontos como delimitador, os identificadores completos se tornariam arXiv:physics/9811004 e arXiv:hep-th/9909044.*

O identificador de um registro no seu arquivo – seja o identificador completo ou o identificador do registro (sem seu identificador de arquivo principal) – será a chave crucial para extrair os metadados de um registro. Em alguns arquivos, também pode ser a chave para se chegar ao conteúdo completo do registro. Em outros arquivos, outros elementos de metadados dentro de um registro apontarão para o conteúdo completo.

Etapa 3: Implementar o conjunto de metadados de arquivos abertos

Reconhecemos que os arquivos utilizarão conjuntos e formatos de metadados específicos que atendam às necessidades das suas comunidades e aos tipos de dados que manuseiam. No entanto, a interoperabilidade depende de um formato partilhado para a troca de metadados e, portanto, os arquivos devem implementar o conjunto básico de metadados de arquivos abertos (oams), descrito em http://www.openarchives.org/sfc/sfc_oams.htm .

Passo 4: Implemente e documente outros formatos de metadados suportados pelo seu arquivo e-print

Os provedores de serviços poderão fornecer serviços mais poderosos aos usuários se um formato de metadados mais rico do que os arquivos básicos puder ser obtido do seu arquivo. Encorajamos os fornecedores de dados a fornecer acesso a toda a riqueza dos metadados disponíveis para apoiar a descoberta e recuperação de registos nos seus arquivos, de preferência através da adopção de um formato de intercâmbio já utilizado por outros arquivos de impressão electrónica. Para ajudá-lo a determinar se um formato existente pode atender às suas necessidades, mantemos uma lista desses formatos de metadados em http://www.openarchives.org/sfc/sfc_metadata.htm .

Se nenhum formato existente atender às suas necessidades, você deverá assumir a difícil tarefa de compilar seu próprio formato. Para conformidade com a interface de colheita apresentada na Etapa 5, você também terá que definir uma representação XML para ela. É importante que você documente totalmente o formato dos seus metadados e disponibilize a documentação aos provedores de serviços. Novamente, isso tornará seu trabalho muito mais fácil. Se você informar a iniciativa Open Archives sobre o formato alternativo de metadados para o seu arquivo, iremos incluí-lo na lista OA de formatos de metadados em http://www.openarchives.org/sfc/sfc_metadata.htm , para que outros possam se beneficiar. seu trabalho.

- *Os formatos que estão em uso ou em implementação pelos arquivos que contribuem para a compilação desta convenção são - no momento da escrita - os seguintes: MARC, ReDIF, Dublin Core, REFER, RFC 1807, Open Archives Metadata Set.*

Etapa 5: Implementar a interface de colheita Dienst

Depois que o seu arquivo tiver identificadores e suportar um ou mais formatos de metadados, a próxima etapa transforma-o em um arquivo aberto, garantindo que os provedores de serviços possam acessar os dados do seu arquivo. Como prevemos a criação de muitos arquivos abertos e porque queremos que os serviços nesses arquivos sejam construídos num período de tempo curto a médio, recomendamos fortemente que todos os arquivos implementem a mesma interface de colheita. Essa interface de coleta deve permitir que terceiros escrevam software que colete dados seletivamente dos arquivos abertos. Propomos uma interface de coleta que esteja em conformidade com o subconjunto de arquivos abertos do protocolo Dienst, descrito em detalhes em http://www.openarchives.org/sfc/sfc_dienst.htm . Para apoiar uma melhor compreensão das recomendações desta convenção, fornecemos aqui uma breve descrição do subconjunto do protocolo.

O protocolo Dienst possui uma implementação baseada em HTTP. Seu subconjunto Arquivos Abertos define um procedimento de comunicação, bem como a sintaxe para as mensagens e respostas correspondentes, que permitirá aos provedores de serviços coletar metadados seletivamente de arquivos abertos que cumpram a Convenção de Santa Fé. O procedimento tem três etapas:

- **Ação 1: um arquivo pode ser pesquisado para obter as seguintes informações sobre o arquivo:**
  - As partições lógicas nas quais os registros são agrupados no arquivo;
  - Os formatos de metadados suportados para entrega de metadados de arquivo em resposta a uma solicitação de coleta.

  O subconjunto Open Archives do protocolo Dienst define como tais solicitações de polling devem ser enviadas e a sintaxe usada por um arquivo para responder a tais solicitações. Não define o conjunto de respostas válidas à solicitação de formato de metadados. Mas no contexto dos Arquivos Abertos, a lista de formatos válidos e seus identificadores está disponível em http://www.openarchives.org/sfc/sfc_metadata.htm .

  Os arquivos devem suportar vários critérios para dividir seu conteúdo em partições lógicas que sejam reconhecidas pelo Open Archives Dienst Subset. Recomendamos especialmente o particionamento orientado por assunto, bem como o particionamento baseado na afiliação do autor.

- **Ação 2: pode ser solicitada uma lista de identificadores para registros em um arquivo.**

  Esperamos que esses identificadores sejam os identificadores persistentes exclusivos dos registros nos arquivos. O subconjunto Open Archives Dienst define a sintaxe a ser solicitada:

  - Uma lista de identificadores para todos os registros do arquivo;
  - Uma lista de identificadores para os registros em uma partição do arquivo;
  - Uma lista de identificadores de registros que ficaram disponíveis no arquivo após uma data especificada;
  - Uma lista de identificadores de registros que ficaram disponíveis em uma partição de arquivo morto após uma data especificada.

  O Open Archives Dienst Subset também define a forma como a lista de identificadores será retornada.

- **Ação 3: dada uma lista de identificadores (obtidos pela Ação 2) e um formato de metadados suportado (identificado pela Ação 1), uma solicitação para coletar metadados pode ser enviada.**

  O subconjunto Open Archives Dienst define a sintaxe para a solicitação de coleta. No subconjunto, a troca de metadados é sempre em XML, qualquer que seja o conjunto de metadados escolhido. Uma lista de formatos e detalhes sobre sua interpretação e representação XML está disponível em http://www.openarchives.org/sfc/sfc_metadata.htm . Um arquivo deve responder a uma solicitação de metadados em um formato especificado, retornando dados renderizados de acordo com o formato de troca correspondente.

Etapa 6: Informe a iniciativa Open Archives de que seu arquivo de impressão eletrônica está aberto

The last step is to join the community of open e-print archives by informing the Open Archives initiative of the details for your open archive. You should use the template we provide for that purpose at http://www.openarchives.org/sfc/data_provider_template.htm to do so. Fill out the appropriate information, install it on your archive server and send us an e-mail at openarchives@openarchives.org to inform us of its URL. We will include your archive in the list of Santa Fe compliant archives that we maintain at http://www.openarchives.org/sfc/sfc_archives.htm , thus helping to make your archive conveniently visible to service providers. You will see that the form also includes a possibility to inform service providers of restrictions that apply in the usage of your data.

You may also want to insert the Open Archives logo in the main entry point of your archive to indicate that you have joined the initiative.

For the service provider: how to make your services comply with the Santa Fe Convention?

As you can see from the recommendations we make to data providers, our aim is to make it easy for you - the service provider - to create services based on open archive data. In return, we ask you to comply with the following:

Step 1: Retain the original identifiers in your services

When you create services based on records originating from open archives, keep the original full identifiers associated with the data as a means of indicating the provenance of records.

Step 2: Comply with the usage restrictions specified by the data providers

With regard to the usage of data from the open archives, respect the restrictions that data providers mention in the form describing their archive. For each open archive, this form is accessible via http://www.openarchives.org/sfc/sfc_archives.htm.

Step 3: Let the Open Archives initiative know that you have developed a service based on open archives data

Inform the Open Archives initiative and thus the data providers about the data that you are harvesting and about the use you make of it. To facilitate this, we provide a template at http://www.openarchives.org/sfc/service_provider_template.htm that you are encouraged to use and fill out. Install it on your server and inform us of its URL by sending us e-mail at openarchives@openarchives.org. By doing so, you are joining the community of Open Archives. We will list all information on service providers at http://www.openarchives.org/sfc/sfc_services.htm.

You may also want to insert the Open Archives logo in the main entry point of your service to indicate that you have joined the initiative.

Supporting information is available at:

- the Santa Fe Convention
- o conjunto de metadados de arquivos abertos
- o subconjunto Open Archives Dienst
- a lista de provedores de dados Open Archives , incluindo seus identificadores exclusivos de arquivo
- a lista de provedores de serviços de Open Archives
- a lista de formatos de metadados usados no contexto Open Archives
- o modelo a ser usado pelos provedores de dados para se registrar como um arquivo compatível com Santa Fé

- o modelo a ser usado pelos provedores de serviços para se registrar como um serviço compatível com Santa Fe

---



entre em contato com a iniciativa Open Archives entrando em contato com openarchives@openarchives.org

última atualização em 20 de janeiro de 2001