DSpace-Fedora Integration FAQ

This page is outdated / obsolete!

This page details an ideal from Open Repositories 2010 which was loosely called "DSpace with Fedora inside". Since that point, this concept never gained leverage or momentum in the DSpace or Fedora communities. It is now considered outdated/obsolete based on the latest technology roadmaps for each platform. At this time, there are no immediate plans to merge DSpace and Fedora or rebuild DSpace as a Fedora-based application. For more information on the direction of both DSpace and Fedora, please consult their respective technology roadmaps.

- DSpace RoadMap
- Fedora Roadmap

At the Open Repositories 2010 conference, DuraSpace recommended that the DSpace community begin to work more closely with the Fedora community to allow DSpace software to run "on top of" the Fedora platform. Essentially, this would mean that the DSpace application would have an option to install it with "Fedora Inside". What follows is an initial FAQ around this proposal.

- DSpace & Fedora Integration Questions
  - What is the DuraSpace recommendation?
  - What does integration mean for me as a DSpace user? Why should I be interested?
  - What is the proposed end result of this integration?
  - Will the end result still look like DSpace and be called DSpace?
  - What are the perceived benefits? Why do this integration?
  - Will I still be able to use DSpace without Fedora?
  - What are the technical plans for integrating the DSpace application with the Fedora repository platform?
  - How is this work being supported?
  - Who will do this work? When will it be complete?
  - When will this work begin?
  - Will new features (or bug fixes) be put on hold until Fedora is integrated with DSpace?
  - How can I help? How do I volunteer?
  - Once this work is complete, how hard will the upgrade be for current DSpace users? How much expertise will I need to perform this upgrade?
  - Will I also need to learn how to install/upgrade Fedora in order to install/upgrade DSpace?
  - Will my current DSpace User Interface / Add-On customizations still work?
  - What happened to the DSpace 2.0 work? Is this replacing it?

DSpace & Fedora Integration Questions

What is the DuraSpace recommendation?

DuraSpace recommends that the DSpace and Fedora communities investigate strategies to allow DSpace software to run "on top of" the Fedora platform. Essentially, this would mean that the DSpace application would have a new option to install it with "Fedora Inside".

DSpace with "Fedora Inside" would be a version of the DSpace application which can store its contents (files and/or metadata) in Fedora. The goal is to retain the out-of-the-box experience that is DSpace, while also enabling the extra features that Fedora provides (e.g. versioning, relationships between objects, flexible architecture).

We are looking for volunteers from either community to help make this concept a reality. This work will happen in stages, over several releases of DSpace software. DuraSpace is already volunteering some staff time to help make this happen.

What does integration mean for me as a DSpace user? Why should I be interested?

The hope is that DSpace with "Fedora Inside" will enable DSpace to support extra features already provided by the Fedora platform. Some of these features may include:

- Versioning – DSpace may be able to support Bitstream/Item/Collection/Community versioning
- Relationships between Objects – DSpace may be able to support more complex relationships between Items. For example, relating together all articles in a journal issue, or all journal issues of a specific journal. This may even help enable Journal "views" or Table of Contents "views" within DSpace.
- Flexible Architecture – DSpace may no longer be limited to the current Community -> Collection -> Item hierarchy.

That's just a small sample of Fedora features, and how they may eventually map into DSpace features. We hope that this integration will be advantageous to both platforms, and promote more sharing and collaboration between these communities.

What is the proposed end result of this integration?

The end result would be a family of packages that you would install together for the full DSpace experience. One of these packages would offer a Fedora-based storage layer. So, when you install DSpace, there would be an option to use Fedora as the storage layer for all files/metadata.

Essentially, DSpace will still retain very similar interfaces and workflows, but it will now have "Fedora Inside".

Will the end result still look like DSpace and be called DSpace?

Yes. We are committed to maintaining the DSpace out-of-the-box experience. The end result of this integration may be an evolved version of DSpace (with new or slightly changed features), but it will still provide the same out-of-the-box institutional repository software.

To put it another way, the goals of DSpace software will not change. DSpace will continue to provide the same experiences you've come to expect. However, we hope the end result will provide even more useful features and will help to evolve the DSpace platform.

What are the perceived benefits? Why do this integration?

We feel this tighter integration between the two platforms will provide the following benefits:

- Less redundancy between DSpace & Fedora code and development.
- Ability to combine development work/projects – a larger pool of developers working on common technologies.
- DSpace will also be able to take advantage of features offered by Fedora (which are not available in DSpace). Some examples may include:
  - File versioning, which DSpace could utilize to support versioning of DSpace Objects.
  - More flexible architecture, which may allow DSpace to eventually support hierarchies other than "Community -> Collection -> Item"
  - Flexible relationships model, which may allow for better defined relationships between DSpace Objects (e.g. linking together all articles within a Journal Issue, or all Issues within a Journal)

Will I still be able to use DSpace without Fedora?

Initially, the plan is that Fedora will be one of several storage layer options available. So, we anticipate that other storage options will still exist, and you could utilize DSpace without requiring that you use Fedora. Likely, in the future, Fedora will become the default storage layer.

Whether all storage layers support the same features is another question. It may be possible that only certain storage layers will be able to support certain features (like file versioning).

What are the technical plans for integrating the DSpace application with the Fedora repository platform?

There is already work underway that may eventually enable an option to run the DSpace application on top of a Fedora repository. To achieve this would require a series of technological developments:

1. **A way to extract the items from DSpace's current internal store, along with all of their relationships (communities, collections, and Epeople).**
   - **Released in DSpace 1.7.0**: DuraSpace (with help/support from MIT & IUPUI) developed this as a general, backup/restore AIP capability for DSpace.
2. **An implementation of the DSpace data model using the Fedora repository.**
   - Related Project(s): There is prior art (which would need to be updated) for this from Google Summer of Code 2008 and Google Summer of Code 2009. There is also a current GSoC 2010 project to backport DSpace 2.x Storage Services to 1.x – this work may enable the ability to use prior Fedora integrations in the DSpace 1.x platform.
3. **A conversion, crosswalking, or import utility for the above two items.**
   - Related Project(s): At this time, there is neither a schedule nor a commitment for this work.
4. **Designing an integration of the DSpace business logic, User authentication and authorization with Fedora.**
   - Related Project(s): At this time, there is neither a schedule nor a commitment for this work.
5. **Porting the DSpace user, administrative, and web services interface applications to run against the resulting environment.**
   - Related Project(s): At this time, there is neither a schedule nor a commitment for this work.

Not all of these steps need to be completed for some models of operation involving running DSpace and Fedora as distinct, cooperating applications.

How is this work being supported?

DuraSpace cannot support this work on its own. That's why we are making a strategic recommendation that the DSpace and Fedora communities work towards tighter integration between the software platforms.

DuraSpace is committed to helping with strategic development work that serves to benefit this initiative. For example, we feel our current AIP Backup /Restore work will be beneficial when upgrading existing DSpace instances to an eventual DSpace with "Fedora Inside".

Who will do this work? When will it be complete?

We currently don't have an exact timeframe for this work to be complete. We fully expect it will be spread out through several releases of DSpace software.

As stated, DuraSpace does not have the technical staff to do this work alone, nor would we want to. As with any open source development initiative, the commitment and support of the broader community will build a better 'end product', than if one institution were to implement it on their own. We are currently looking for DSpace and Fedora community volunteers to help make this initiative happen.

When will this work begin?

In fact, this work has already begun through various projects and initiatives over the last few years. This DuraSpace recommendation is just serving to formalize our support of this effort and broaden community interest in the endeavor. Although they've been mentioned elsewhere already, here are several ongoing and past projects which relate to this initiative:

- DSpace 2.0 prototype work – provided a basis for modularizing DSpace, allowing it to support different storage layers (like Fedora)

- GSoC 2008 - Fedora Integration Project - An initial implementation of a prototype Fedora storage layer for DSpace 1.x
- GSoC 2009 - Fedora Integration Project - An improved implementation of a prototype Fedora storage layer for DSpace 2.x
- GSoC 2010 - DSpace 2.x Storage Services to 1.x - A new effort to backport the DSpace 2.x storage layer to DSpace 1.x. This may enable one of the above Fedora Integration projects to be utilized with DSpace 1.x.
- **Released in DSpace 1.7.0:** AIP Backup and Restore - An effort to allow DSpace to export and import AIPs (Archival Information Packages). These AIPs could provide an easier upgrade path from one underlying data model to another.

Will new features (or bug fixes) be put on hold until Fedora is integrated with DSpace?

No. We do not want this initiative to hinder the current development process. Upcoming releases of DSpace software will continue to include new features and resolve bugs/issues. Chances are much of the code changes coming out of this integration effort will be rolled out over several releases (initially many may be architectural, "under the hood" changes).

Think of this initiative as a separately organized project, whose end goal is better collaboration between our communities and better integration between the software. DSpace software itself will continue to improve and change as this effort goes along. We fully expect there to be several releases of DSpace software before we are able to release a DSpace with "Fedora Inside".

How can I help? How do I volunteer?

The best way to help out is to volunteer to lead or help with one (or more) of the development projects listed under the *DSPACE:What are the Plans for Integrating the DSpace Application with the Fedora Repository Platform?* question.

As you can tell, several initiatives have already begun. If you want to help out with one of the existing initiatives, contact the person(s) in charge or send an email to the dspace-devel listserv. If you'd like to lead a new initiative (which hasn't yet started), please announce your intentions on the dspace-devel listserv so that we can send other volunteers your way!

At this time, we are mostly looking for volunteer developers/architects. As the initiative gets further along, we will put out additional calls for volunteer testers, etc.

Once this work is complete, how hard will the upgrade be for current DSpace users? How much expertise will I need to perform this upgrade?

At this time, we don't know the complete answer.

To be clear, we fully expect there to be several releases of DSpace software before we are able to release a DSpace with "Fedora Inside". Most upcoming releases should have normal upgrade processes (similar to upgrading from 1.5 to 1.6). The only upgrade that may contain extra steps is when DSpace is first released with "Fedora Inside". Our intention is the make this upgrade as easy as possible and find common solutions that everyone in the community can utilize.

One of these common tools is the aforementioned AIP Backup/Restore work. This tool will provide you with a way to export all your DSpace contents (Communities/Collections/Items) into a series of AIPs (Archival Information Packages), and import or restore them to any DSpace instance. We hope that this tool could be used to ease the migration of your DSpace contents from the current DSpace data model into a Fedora-based data model.

We plan to build other common tools that will allow everyone to easily upgrade their DSpace when the time comes.

Will I also need to learn how to install/upgrade Fedora in order to install/upgrade DSpace?

At this time, we don't know the complete answer.

To be clear, we fully expect there to be several releases of DSpace software before we are able to release a DSpace with "Fedora Inside". You will not need to install Fedora until DSpace is initially released with "Fedora Inside".

In the first few "Fedora Inside" releases, you may need to perform some basic steps to install or upgrade Fedora whenever you install or upgrade DSpace. However, the Fedora component should never be more difficult to manage than the existing relational database (PostgreSQL) component.

We anticipate that in the future we will be able to simplify this process. The final goal is to be able to install DSpace with "Fedora Inside" without too much effort (i.e. hopefully less than or equal to the effort required to install DSpace 1.6).

Will my current DSpace User Interface / Add-On customizations still work?

In the near term, all your customizations should continue to work (or require minor updates during upcoming releases). Again, we fully expect there to be several releases of DSpace software before we are able to release a DSpace with "Fedora Inside".

We currently don't know which customizations will continue to work once DSpace is released with "Fedora Inside". We hope that basic user interface customizations (CSS / Javascript / etc) will work immediately. However, any customizations to the internal API or complex add-ons will likely need to be rebuilt or refactored. As we move closer towards figuring out the internal API changes required for this integration, we will be able to better answer this question. We also will report which APIs will change, so that you can more easily determine whether your add-ons will require updates.

What happened to the DSpace 2.0 work? Is this replacing it?

This initiative is actually an extension of the DSpace 2.0 investigation work that took place in 2009.

The goal of the DSpace 2 work was primarily to modularize DSpace, and Fedora as one of those modules just makes sense. The DSpace 2 work, along with related Google Summer of Code projects (in 2008 and 2009), has helped us to more closely realize how DSpace and Fedora could be integrated.