GdataServer FAQ

GData-Server FAQ

General

What is GData?

The Google data APIs ("GData" for short) provide a simple standard protocol for reading and writing data on the web. GData combines common XML-based syndication formats (Atom and RSS) with a feed-publishing system based on the Atom publishing protocol, plus some extensions for handling queries. For detailed information see the Google data API overview.

What can Gdata do for you?

Do you want to publish information any kind on the web, in your intranet or any other HTTP compatible network gdata offers you a xml format based on either Atom 1.0 or RSS 2.0. That enables you to make your information universially accessible via a webbrowser or any other kind of application or API can handle HTTP requests and parse XML data. Basically any information which can be displayed and/or organized as xml can be served on a gdata server. The client API provides a highly extendable model to customize and extend the api to your suites. You can create feed and entry extensions programmatically for your client application and use it on the fly on the server by just providing the classes in the classpath.

The range of operational area is huge, you can already use the feed / entry types included in the client api to build your own basic feed server, organize your mailing list via "MessageFeeds", create a distributed calendar or even organize your own website content on the gdata server. Additional to just have the posssibility to provide this kind of service the gdata server makes all the content searchable with customized index / search schema.

Generally spoken GData provides a general interface to make information available even beyond a browser context.

Does the server supports all of the client api features?

Yes it does. All the features described in the gdata protocol are supported. There is also an extended search query language based on the lucene query syntax.

Does the server has an web based adminstration interface?

Just a simple answer, not yet. But that is on top of the todo list.

All feed, entry and user administartion is based on a REST approach and has to be done programmatically via a http client. You can find some examples in the Getting started section of this wiki.

Can I provide more than one feed format on a GData - Server instance?

Yes, you can provide as many different feed formates as you want to. A certain feed type as "Feed", "EventFeed" or "MessageFeed" is defined by creating a service via the *gdata-config.xml* configuration file. A service constists of a feed class, an entry class, an extension profile and an index schema to specify how the feed typ e.g. it's entries should be indexed and be available for searching.

How can I configure the GData - Server?

All configuration for the server is located in the *gdata-config.xml* file. This file must be available in the classpath at server startup. Current location is */WEB-INF/classes/gdata-config.xml*.

What is a server component and can I use my own implementation?

The GData - server is build on a component model to provide a flexible server architecture. Currently the server consists there are six different server component which can be configured at startup via the *gdata-config.xml* file. Each component has to implement a defined interface and will be loaded as a singleton at server startup. All classes need to access a component use a lookup service to get the configured implementation. If one of the component fail at startup time due to some error the server will not startup an releases all already loaded components.

If you want to implement your own component for example a new indexing or storage implementation you just need to provide the defined interface to do that.

To get more information on that issue see Getting started.

Are modification reflected to the search index straight away?

It depends, if you use a timed indexer implementation updated, deleted or new added entries will be available within the defined idle time. With other implementation the indexer will refresh searchers after a definded amount of changes to the index.

Timed indexer tasks have an idle time and call a commit if no modification occured for the defined time.

How does the server index entries?

First you have to configure a index schema for your service which will be loaded at startup. This schema configures the associated service indexer which extracts the content of incoming entries and writes it into the index. Field names, analyzers, store and index type will be defined in the index schema. The GData - Server provides several default types of *element content* like html or html content to extract the indexable text out of an xml element. More information on that will be available in the HowTo section of this wiki.

Each defined field will be searchable via the HTTP interface using the fields name as a GET parameter.

How can I rebuild an index for a specific service?

1. Shut down your gdata server 2. go to the directory you configured as "indexlocation" in the `gdata-config.xml` file. 3. go into the directory named after the service you want to reindex 4. delete all index files 5. restart you server (the server will detect that there is no existing index and creates a new one) 6. update your entries you want to index.

Can I create my own user account?

Yes you can, the GData - Server has a simple user role model with 4 different roles.

- Standard user - can alter his own feed.
- User administrator - can create user accounts.
- Entry administartor - can post, delete and update entries one every feed and service.
- Feed administrator - can create new feeds on the server for every entry.

You can also combine the user roles to create a user who can create feed and user or even more. If you deploy your server the first time there will be already an admin user with all roles available. This user is set to the username *administrator* with the default password *password* you should update this user to set your own admin password.

Can I provide my server via HTTPS?

Yes, check your servlet container documentation how to configure it to use https for the url pattern you want to be available via https. Apache Tomcat SSL HowTo

Is there an alternative to Lucene GData-Server

There might be several syndicate servers arround on the web creating gdata feed but as far as I know there is no server implementation avaiable right now.

Where can I ask questions?

All you need to do is to subscribe to the lucene java-user / java-dev mailing list and ask your question.

How is the server released?

GData-Server comes as a war file incuding all requiered jars and configuration files. To alter configuration you should create your own war file via the source distribtion. Currently the server is only available via the source repository.

Is there an ant task to create the war file?

Certainly there is, see build.xml.

What Java is requiered to run GData-Server?

GData-Server will only run with and compile with JDK >= 1.5 as it makes extensive use of 1.5 features.

What kind of server is requiered to run the GData - Server?

The server runs in any JDK 1.5 supporting servlet container. We suggest to use the latest version of Apache Tomcat.

Which version of lucene does the server use?

Development is based on the current trunk of Apache Lucene. As the project is a contrib project it will also be released with lucene. The War file will contain the current development lucene jar or the realeas jar if released as a version.

Can I contribute in the GData-Server?

Certainly you can just see these steps to submit a Lucene patch and start coding.