

Metacat - Bug #3527

getprincipals action returns invalid XML document

10/13/2008 05:11 PM - Chris Jones

Status:	Resolved	Start date:	10/13/2008
Priority:	Normal	Due date:	
Assignee:	Jing Tao	% Done:	0%
Category:	metacat	Estimated time:	0.00 hour
Target version:	1.9	Spent time:	0.00 hour
Bugzilla-Id:	3527		

Description

When calling action=getprincipals, metacat returns the list of users and groups as an XML document with a character encoding set to "US-ASCII". However, users register from around the world, and many have registered with usernames that use characters beyond ASCII character set. For instance:

```
<user>
  <username>uid=redbc,o=unaffiliated,dc=ecoinformatics,dc=org</username>
  <name>Zulay Míndez</name>
  <organization>unaffiliated</organization>
  <organizationUnitName>unaffiliated</organizationUnitName>
  <email>zulay@ula.ve</email>
</user>
```

The above user's name (Mendez?) includes a non ASCII character.

Client side parsers will not parse the output from getprincipals because the XML document is not valid.

We need to either:

- 1) Change the output's character encoding to an all-inclusive one (UTF-8?), or
- 2) Change the registration CGI such that it only accepts character input in the ASCII range.

Also, there are a half-dozen LDAP entries that include non-ASCII characters, and so in option #2, these entries would have to be modified.

History

#1 - 10/14/2008 10:14 AM - Chris Jones

After modifying the output XML to include:

```
encoding="UTF-8"
```

client parsers parsed the XML just fine. This may be a matter of changing AuthLdap.java to use this as the default encoding in the getPrincipals() method.

#2 - 11/05/2008 04:08 PM - Jing Tao

Chris and I tried encoding="UTF-8" and it didn't work. Both firefox and chris's matlab code didn't this encoding.

So we tried encoding="iso-8859-1" and it works on both firefox and matlab.

iso-8859-1 is better than US-ASCII, but we should keep an eye on it down the road.

#3 - 03/27/2013 02:23 PM - Redmine Admin

Original Bugzilla ID was 3527