

QGIS Application - Bug report #2151

"Save as shapefile..." truncates field names, sometimes causing duplicates

2009-11-25 09:04 AM - dmorissette -

Status: Closed	
Priority: Low	
Assignee: nobody -	
Category: Vectors	
Affected QGIS version:	Regression?: No
Operating System: Debian	Easy fix?: No
Pull Request or Patch supplied:	Resolution: fixed
Crashes QGIS or corrupts data:	Copied to github as #: 12211

Description

Since the DBF format limits field/attribute names to 10 chars, the "Save as shapefile" command (from the right-click menu on vector layers in legend) has to truncate the name of fields longer than 10 chars. However, this is done without checking for potential duplicates field names.

For instance, if a dataset contains fields "mylongattr1" and "mylongattr2", they are both written as "mylongattr". We should change the dups to have a unique name, presumably by adding a unique number at the end.

I believe QGIS uses OGR for this export (am I right?). If that's the case then the source of the problem is in the OGR shapefile driver where the field name truncation happens. I have filed an OGR ticket about this with a testcase to reproduce, and once it's fixed QGIS should automatically benefit from the fix:

<http://trac.osgeo.org/gdal/ticket/3247>

However, QGIS could possibly have some smarts to better handle this in the meantime. I'll let you decide if you wait for OGR or implement a fix in QGIS.

Associated revisions

Revision af24fd4e - 2009-11-29 12:20 AM - Jürgen Fischer

fix #2151

git-svn-id: <http://svn.osgeo.org/qgis/trunk/qgis@12280> c8812cc2-4d05-0410-92ff-de0c093fc19c

Revision 29ca1a23 - 2009-11-29 12:20 AM - Jürgen Fischer

fix #2151

git-svn-id: <http://svn.osgeo.org/qgis/trunk@12280> c8812cc2-4d05-0410-92ff-de0c093fc19c

History

#1 - 2009-11-28 03:21 PM - Jürgen Fischer

- Resolution set to fixed

- Status changed from Open to Closed

fixed in commit:29ca1a23 (SVN r12281).

#2 - 2009-11-30 07:55 AM - dmorissette -

I see that you produce a fatal error if duplicates are found, which is way better than silently producing an invalid output. However, the day OGR is fixed to automatically rename the fields, QGIS won't be able to benefit from that fix...

If I'm not mistaken (and I'd love to be proven wrong), the current fix (ErrAttributeCreationFailed) forces users to rename the fields in the source before the conversion if they want to export to shapefile, which is not always possible if they connect to a DB or other data source with a schema that they do not control. I believe in the long run it would be better to either support renaming the fields automatically in QGIS (e.g. add a unique number at the end of dups to make them unique), or to be able to let OGR do it for us once <http://trac.osgeo.org/gdal/ticket/3247> is fixed.