

QGIS Application - Bug report #21375

Order of attributes in QGS XML is non-stable

2019-02-25 01:54 PM - Adam Liddell

Status:	Open	
Priority:	Normal	
Assignee:		
Category:	Project Loading/Saving	
Affected QGIS version:	3.6.0	Regression?: No
Operating System:	Windows 7	Easy fix?: No
Pull Request or Patch supplied:	No	Resolution:
Crashes QGIS or corrupts data:	No	Copied to github as #: 29192
Description		
<p>When saving a project, the order of the attributes on the XML elements in the saved QGS file appears to be non-stable. When trying to version control a project, this causes massive diff 'noise' from attributes being reordered, even when the actual data in the attributes is identical.</p> <p>For example, on one save, an element may be written as:</p> <pre><layer-tree-layer source="?query=SELECT..." id="some_layer" providerKey="virtual" expanded="1" name=some_layer" checked="Qt::Unchecked"></pre> <p>The next time the file is saved, the order of the attributes may change for the same element:</p> <pre><layer-tree-layer providerKey="virtual" expanded="1" id="some_layer" name="some_layer" source="?query=SELECT..." checked="Qt::Unchecked"></pre> <p>Simply resaving a file immediately after a previous save does not appear to trigger this, so there is some 'change' in the project that causes the attributes to reorder.</p> <p>As a proposed solution, the attributes on the elements should be ordered stably, perhaps alphabetically?</p> <p>The order of the XML elements within the document appears to be stable.</p> <p>Tested in 3.4.4 and 3.6.0.</p>		

History

#1 - 2019-02-25 02:06 PM - Alessandro Pasotti

IMO this is a won't fix, the attribute order is random due to QT internal implementation <https://doc.qt.io/qt-5/qhash.html#algorithmic-complexity-attacks>

You can control with an env var as explained in the docs.

#2 - 2019-02-25 02:34 PM - Adam Liddell

Ok, I see the issue with using QDomDocument in QT5 to get stable ordering. However, whilst using QT_HASH_SEED would work on one machine, the output is supposedly not necessarily stable across different machines or versions even when set to 0, which doesn't solve the (perhaps minor) problem of using shared version control on a project.

Would there be benefit to post-processing the generated XML into canonical form (https://en.wikipedia.org/wiki/Canonical_XML), as suggested here: <https://geidav.wordpress.com/2013/02/27/deterministic-attribute-order-in-xml-documents-in-qt-5/>? Perhaps as an option if you don't want to pay the price for every project.

