

## NAME

xml2json - converts an XML document to JSON format

## SYNOPSIS

```
xml2json [-help] [-version] [-array <csv>] [-prefix <s>] [-text <s>] <xml-file>
```

## DESCRIPTION

The akiwi xml2json command converts a UTF-8 encoded XML document to a UTF-8 encoded JSON representation and writes the result to stdout.

The output format is heirarchical, making it easier for humans to read. If you need a more compact format, then please use another utility to strip out the excess whitespace afterwards.

The ordering of the JSON output is identical to the XML input. i.e. the XML structure is mirrored exactly within the JSON output document.

To work-around any possible issues with naming conflicts, there are options to add a prefix to all XML attributes and to define the name used for the XML text attribute.

Without parsing and utilising an XML DTD, it's impossible to know which XML elements should be contained within a JSON array and which should not, if there is only one such XML element present. An additional option is provided to force a specified list of XML element names to always be output as JSON array elements if there is only one occurrence.

This command also has a couple of [important limitations](#) that users should be aware of beforehand.

Firstly, the XML parser is a tiny homegrown recursive descent parser that only supports the essential aspects of XML. If your documents have processing instructions or other complex DOCTYPE declarations, it's likely to choke on them. The XML parser automatically recognises all predefined HTML 4 entity references.

Secondly, it is assumed that XML elements of the same name at the same level are grouped together consecutively so that they are all placed within the same JSON array. This is typically the case anyway, but just be aware that name conflicts may occur if this is not the case.

## OPTIONS

The option flags may occur in any order on the command line, but only once. For option flags with arguments, there must be at least one space between the option flag and its argument.

- help Prints a short summary of the command line syntax and options.
- version Prints program version and copyright information.
- array Specifies a comma separated list of element names that should be forced to be contained within a JSON array. e.g, the option '-array var,product' forces any single XML elements named 'var' or 'product' to be contained within a JSON array. There should be no whitespace between the element names in this list. The default is an empty list.
- prefix Specifies a string to be prefixed to all XML attribute names. e.g., the option '-prefix @' would rename the XML attribute 'valid' to '@valid'. The default is no prefix.
- text Specifies the name of the attribute used to store XML text (or chardata). The default is 'content', the Yahoo convention, but google uses '\$t' and others use '#text'. So for example '<name>Steve</name>' with '-text \$t' would produce:-

```
"name": { "$t": "Steve" }
```

## EXAMPLES

### Example 1

```
xml2json -array email,tel -prefix @ -text $t foo.xml > bar.json
```

Convert the UTF-8 encoded XML file 'foo.xml' to JSON format and pipe the output to a file named 'bar.json'. Any single elements named 'email' or 'tel' are forced to be contained within a JSON array. All XML attribute names are prefixed with an '@' character. Any XML text is stored in an attribute named '\$t'.

## COPYRIGHT

Copyright (C) 2013. By akiwi. All Rights Reserved.