VOTable 1.6: Progress and Status

Mark Taylor (Bristol)
with input from many members of Apps WG

Applications WG,
IVOA Interop,

Boorloo/Perth
IVOA Interop
6 November 2026

$Id: vot.tex,v 1.5 2026/11/06 05:46:21 mbt Exp $

Why VOTable 1.6?

Some things in VOTable 1.5 need updating/improving/modernising: (see also Görlitz, Strasbourg talks)

- **Unicode handling is inadequate/outdated** (Issues #55, #69):
  - `datatype="char"` restricted to 7-bit ASCII
  - `datatype="unicodeChar"` restricted to *Basic Multilingual Plane*, also inefficient `BINARY` serialization
    → No emoji possible in VOTable data 😢
  - VOTable standard discusses Unicode in outdated terms (using the obsolete "UCS-2" encoding)
  - Non-ASCII content is technically illegal in `INFO` elements
- Media type parameters (Issues #26, #15):
  - Need to support usage like `application/x-votable+xml; content=datalink` used in other standards
- MIVOT/MANGO references (Issue #85):
  - Update in view of recent standards progress
- Some minor editorial things
  - XSD file has lots of commented-out historical material
  - Some formatting issues, minor errors, outdated URLs, ...
- Appendix A *"Possible VOTable Extensions"* (Issue #53):
  - Untouched since v1.10 (2004)
  - Contains several sections that are obsolete/of unclear status (in many cases ideas never taken up)

Why VOTable 1.6?

Some things in VOTable 1.5 need updating/improving/modernising: (see also Görlitz, Strasbourg talks)

- **Unicode handling is inadequate/outdated** (Issues #55, #69):
  - `datatype="char"` restricted to 7-bit ASCII
  - `datatype="unicodeChar"` restricted to *Basic Multilingual Plane*, also inefficient `BINARY` serialization
    → No emoji possible in VOTable data 😢
  - VOTable standard discusses Unicode in outdated terms (using the obsolete "UCS-2" encoding)
  - Non-ASCII content is technically illegal in `INFO` elements
- Media type parameters (Issues #26, #15):
  - Need to support usage like `application/x-votable+xml; content=datalink` used in other standards
- MIVOT/MANGO references (Issue #85):
  - Update in view of recent standards progress
- Some minor editorial things
  - XSD file has lots of commented-out historical material
  - Some formatting issues, minor errors, outdated URLs, ...
- Appendix A *"Possible VOTable Extensions"* (Issue #53):
  - Untouched since v1.10 (2004)
  - Contains several sections that are obsolete/of unclear status

A Possible VOTable extensions **36**
A.1 VOTable LINK substitutions . . . 36
A.2 VOTable Query Extension . . . 36
A.3 Arrays of Variable-Length Strings . . . 38
A.4 FIELDs as Data Pointers . . . 38
A.5 Encoding Individual Table Cells . . . 39
A.6 Very Large Arrays . . . 40
A.7 Additional Descriptions and Titles . . . 40
A.8 A New `XMLDATA` Serialization . . . 40

Recent Progress 1

Status just before Strasbourg Interop: WD-VOTable-1.6-20260413

- ☑ Unicode handling
  - ▷ `datatype="char"` now represents UTF-8 bytes not ASCII characters
  - ▷ `arraysize` attribute now counts bytes (*UTF-8 code units*) not characters
  - ▷ `datatype="unicodeChar"` deprecated
- ☑ Media type
  - ▷ `content` parameter defined for `application/x-votable+xml`
  - ▷ "MIME Type" reworded as "Media Type"
- ☑ Editorial
  - ▷ XSD file tidied up
  - ▷ A few minor mistakes and outdated terms cleared up
- ☑ MIVOT/MANGO references
  - ▷ Laurent authored ~~PR #88~~, merged and will be in next draft
- ☐ Appendix A *"Possible VOTable Extensions"*
  - ▷ Not modified yet

Recent Progress 2

Discussions at Strasbourg Interop

- ☑ Media type
- ☑ Editorial
- ☑ MIVOT/MANGO references
- STOP Unicode updates
  - ▷ Objection from François Ochsenbein: there is no longer an ASCII-only data type
- STOP Delete Appendix A
  - ▷ Concerns from several CDS people about removal of *templated* `LINK` mention in Appendix A.1
  - ▷ Related: `LINK content-role` attribute

Discussions since Strasbourg on contentious issues:

- Some messaging by email in recent months
- Face-to-face (+Zoom) splinter during ADASS 2 days ago between interested parties (Mark, Pierre, Laurent, Gilles, Tom, Adrian, Marco, Brigitta)

Resolution: Unicode

Objection to Unicode changes:

- Issue: it is not possible in general (though in some cases yes) to tell whether a FIELD contains pure ASCII or not
  - ▷ François Ochsenbein doesn't like it (apps email 10 Jul 2026):

    > *"the question of *how to specify that a* `<FIELD>` *is made of ascii characters only* has no answer in the current VOTable draft... But apparently nobody cares ?"*

  - ▷ I encouraged followups to this in apps email 13 July 2026
  - ▷ No response ... apparently nobody else is worried by this

- **Decision: Go ahead with current draft text about Unicode updates**
  - ▷ Weight of opinion agrees that the benefits outweigh the disadvantage of no pure-ASCII type
  - ▷ Will note when PR is passed to Exec that François made an objection

Resolution: Appendix A/Templated LINK

```
<LINK href="http://ivoa.net/lookup?Galaxy=${Name}&amp;RA=${RA}&amp;DE=${DE}"/>
```

Templated `LINK`:

- CDS have used syntax like this for a long time, with benefit to CDS-friendly clients
  - Usage has never been properly documented, but Appendix A.1 documents similar syntax
  - Provides functionality *somewhat* overlapping with DataLink Service Descriptors
  - But can do some substitutions that Service Descriptors can't (non-`key=value` request parameters)
- ... so there has been some resistance to removing Appendix A.1
  - Write templated `LINK` docs in VOTable document? Or separate Note?

Decision:

- Delete Appendix A
- No changes to `LINK` section in VOTable 1.6
- Investigate upgrade of DataLink Service Descriptor mechanism to use more powerful templating
  - This will probably be based on RFC 6570 "URI Template"
  - Mark and Laurent to work on this
- Until upgraded DataLink is available, CDS will continue to use their convention as before

Resolution: `LINK content-role` attribute

`content-role` attribute of LINK indicates *"The role and the nature of the link"*

- At VOTable 1.5 there are two options:

  `"doc"`: *"indicates documentation"*
  - Straightforward

  `"type"`: *"indicates a type-like relationship"*
  - Intended for semantics (SKOS), introduced at VOTable 1.3 by Norman Gray
  - Little/never used since then?
  - But not doing any harm
  - Pierre suggested to remove

Decision: No changes

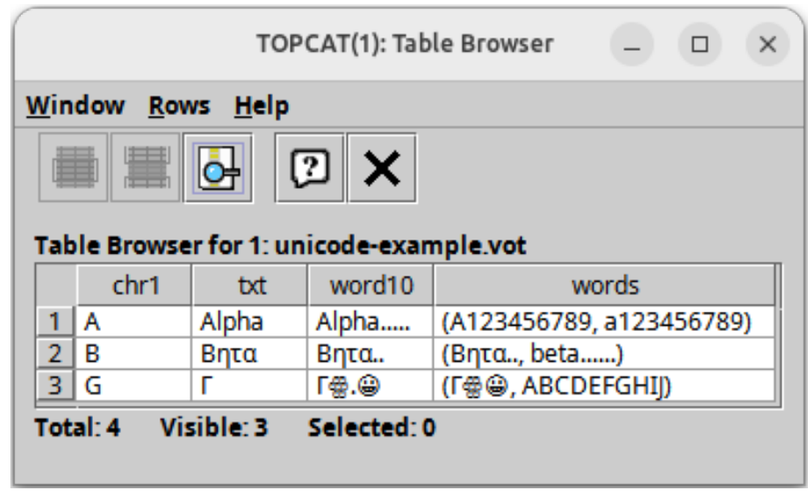
- This change is not pressing, and now no other changes to `LINK` section

Remove Appendix A! 

Remove Appendix A! 

⇒ should be ready for PR → RFC

Implementation Status



We have ≥2 interoperable implementations and a validator!

- STIL/STILTS/TOPCAT (Mark T)
  - Recent release (TOPCAT v4.10-6, STILTS v3.5-5, May 2026) has prototype support for WD-VOTable-1.6-20260413
  - Reads UTF-8 `char` data from VOTable 1.6
  - Writes UTF-8 `char` data on request (`ofmt="votable(version=1.6)"`)
  - Validation: `votlint` (hence `taplint`) checks new constraints in VOTable 1.6 `char`/`unicodeChar` data
  - No behaviour change so far for VOTable ≤1.5, which is/was a bit sloppy about character data
- cds-votable-rust (F-X Pineau)
  - cds-votable-rust v0.7.0 (Dec 2025) supports VOTable 1.6 I/O
- Astropy ready to go? (PR #18987, Stelios)
- Others?

Implementation support:

- Notes on implementation requirements, see MT presentation from Görlitz
- Example Unicode-heavy VOTable available at PR #75 (unicode-example.vot)

Next Steps

Any objections?

- I am hoping not

Actions:

- Remove Appendix A and references to it from the text
- Produce and publish Proposed Recommendation (mid November)
- Circulate test Unicode-heavy VOTable example (to Astropy, others?)
- Start RFC (end 2026 or very start of 2027)

Next Steps

Any objections?

- I am hoping not

Actions:

- Remove Appendix A and references to it from the text
- Produce and publish Proposed Recommendation (mid November)
- Circulate test Unicode-heavy VOTable example (to Astropy, others?)
- Start RFC (end 2026 or very start of 2027)
- All WG chairs review and approve quickly!