# REPORT: Survey of I-D Authors

Introduction

Part 1 - Formats and what drives their uptake

Part 2 - Supporting tools

Part 3 - How people run their tools

Part 4 - v3 XML adoption

Part 5 - Future planning of tools overall

## Introduction

This report of analysis and recommendations from the IETF Executive Director and the Temporary RFC Series Project Manager is drawn from the results of the survey "I-D Authors: Formats and Tools" active from October-November 2020. This survey was issued to a total of 6037 email addresses, with that list built from the addresses of everyone who submitted an I-D, or was listed as an author, in the last five years. 718 responses were received, giving a margin of error of +/- 3.43%.

The results are available as a standalone web site. Some custom graphics have been produced for this report that are not otherwise published.

The recommendations in this report are written in the passive voice as further work is required to understand to whom each recommendation is made.

# Part 1 - Formats and what drives their uptake

**Q3 - How often have you used the following document formats and associated tools (however those tools are hosted), when authoring an I-D? (check all that apply)**

XML is by far the most used format for authoring as expected given that XML is now the format used for canonical and archival copies of RFCs.

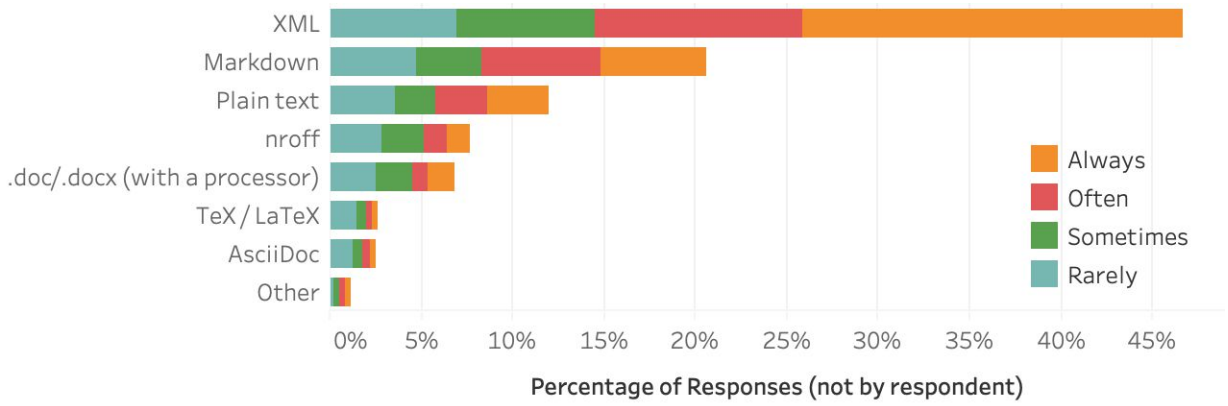


Figure 1: Q3 - non-text responses clustered by format

The surprising part comes when we look at the different ways in which the formats are used, with the survey appearing to show that the use of XML for authoring but then plain text for submission, comes out on top:

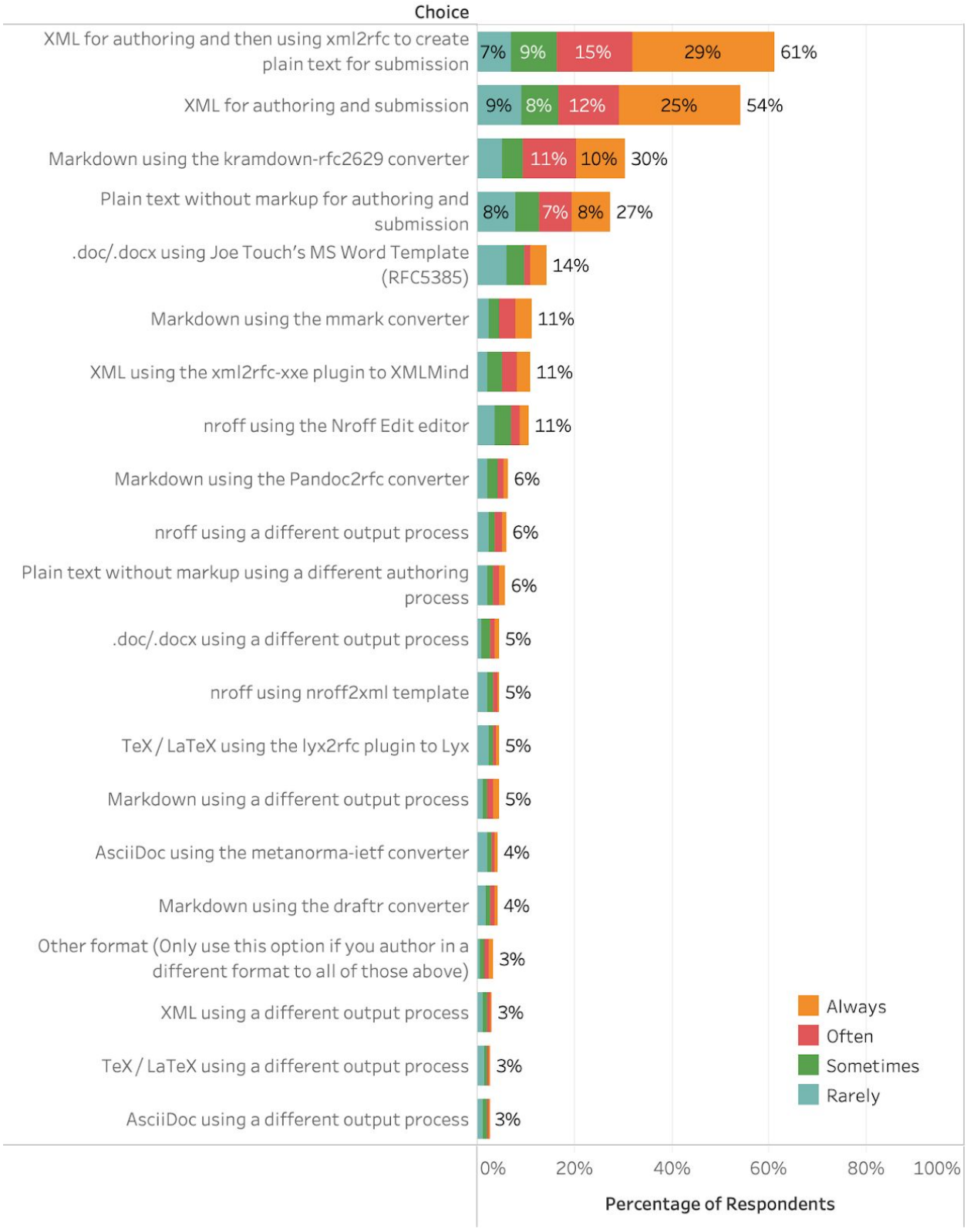


Figure 2: Q3 - individual non-text responses

The current submission process allows people to submit multiple files at once, so it is not clear if this result is because people who submit both XML and TXT have answered as if they submit just the TXT.

**Recommendation 1**: *Adjust the I-D submission interface to nudge people to submit the XML where they have it and to be clear that any files submitted in addition to the XML are secondary and not taken as authoritative.*

This leads to a broader concern that we don't have good enough data on what tools are used and what formats they author in. Some tools (e.g. id2xml) identify themselves but in some cases that information is stripped out during processing and in some cases the RPC have found that information to be incorrect.

**Recommendation 2**: *Implement a standard mechanism for an authoring tool to identify itself and the original format that a document started in, simple enough for someone to code by hand if they wish, and which is preserved through the entire authoring process.*

**Q4 - How did you choose the document format(s) and associated tool(s) that you use? (check all that apply)**

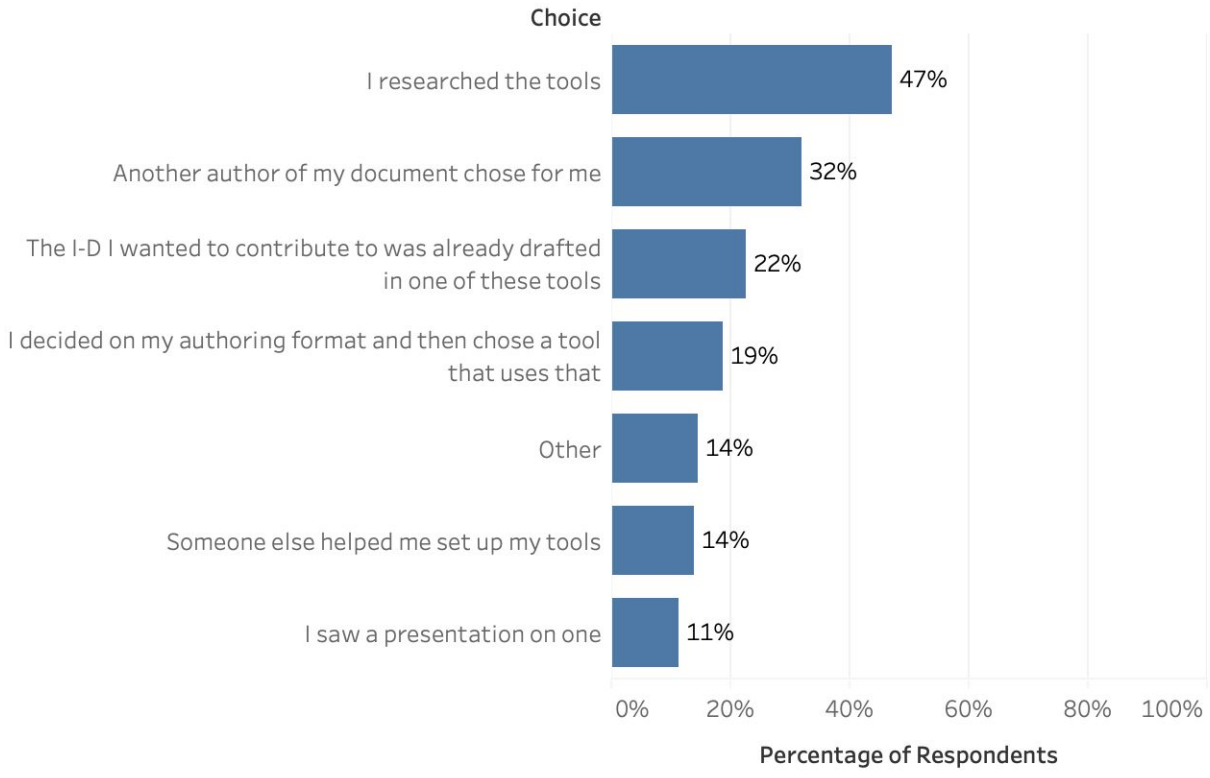


Figure 3: Q4 - individual responses

Approximately 50% of the responses show that people have used information we provide on the tools to make their choice, highlighting just how important it is that we provide high quality, accurate and educational information. Currently information on authoring is available in very different places:

1. On the front page of tools.ietf.org. This is a dense page of tools with a short paragraph by each one. There is no distinction between tools that are in regular usage and tools that are now largely defunct. Nor is there any overall explanation of the authoring workflow and an explanation of how the various tools fit into that.
2. The RFC Editor site has a page that's a bit tricky to find, with a subset of tools and details about them.
3. There are a variety of presentations and videos linked to from the IETF site that have been used at IETF meetings to introduce people to tools.
4. The authors of the various tools provide their own documentation.

**Recommendation 3**: *Create a high quality, accurate and educational micro-site on the authoring process including available tools and formats, with guidance for participants at a variety of levels of process knowledge and technical expertise. Then point all other existing sources under IETF control to this site.*

**Q5 - How often have you used the following template(s) when drafting an I-D? (check all that apply)**

The stand out answer here is that people reuse existing I-Ds to create new I-Ds.

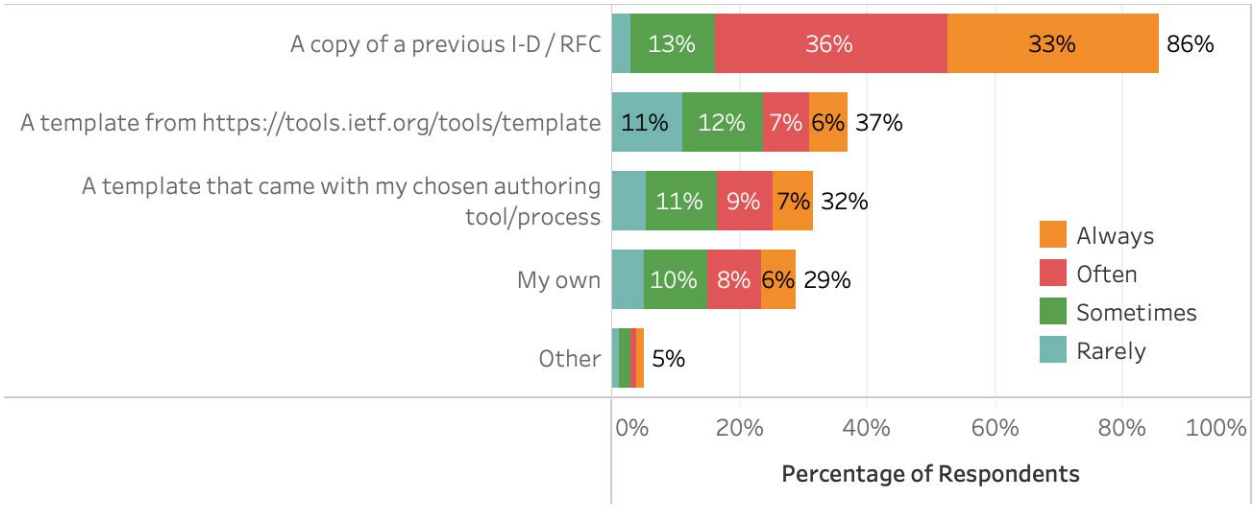


Figure 4: Q5 - individual responses

While we didn't ask why people do that, we can reasonably assume that it is because those are easy to find and are pre-populated with text that would otherwise have to

be added. The problem we have with this though is that changes to the I-D structure or syntax will be slow to propagate as people are not going back to the authoritative source for new versions. However, our authoritative source of templates is sub-optimal:

- The templates are available in a raw directory with no annotation.
- Only one of these templates is for v3 XML and not clearly named as such though it is directly referenced on the front page of tools.ietf.org.
- There is an XML2RFC Internet Draft Creation Wizard which builds a template based on user input. Unfortunately this is v2 XML only and has a very old fashioned UI, but is otherwise an excellent tool.
- The RFC Editor also makes available a directory of unprepped RFCs, which can be used as templates or references for more complex work, though this directory is not advertised.

**Recommendation 4**: *Create a new microsite for I-D templates that is both high quality and educational. Include in that a modern tool for creating customised I-D templates.*

# Part 2 - Supporting tools

**Q6 - How often do you use the following checking tools (however hosted)?**

The answers to this question are pretty much as expected.

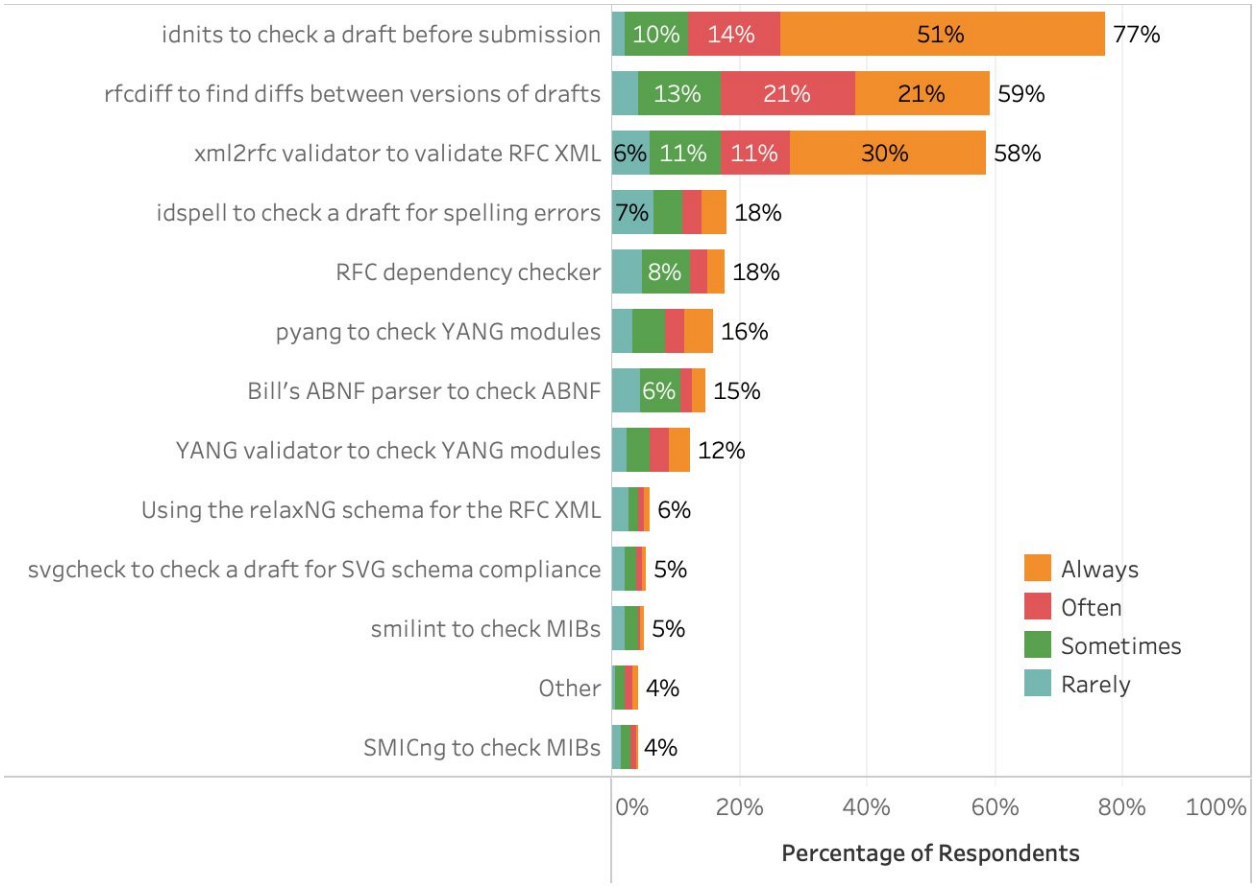


Figure 5: Q6 - individual responses

One tool was left off this list, *xmlrfcdiff*, because this was still under development at the time and now following the untimely death of the author, its status is unclear. Also, that tool was developed solely for the RPC with no mechanism for incorporating community requests. The RPC in the meantime are comfortable using *rfcdiff* with appropriate settings for the length of XML 'lines'.

This is related to the more controversial issue of what format is used for author review during the RPC editing process and what tools authors use to do that. Very few authors review the XML diffs preferring the plain text instead, but as the plain text is not authoritative and does not include the semantic information that the XML does, they may miss vital information this way. Those that do use the XML have until recently (after the survey) had problems with unnecessary line breaks being added that mess up the working of diff tools. This has now been fixed with a change to the configuration of the editing tools used by the RPC.

**Recommendation 5**: *The process by which authors review RPC changes should be examined with a view towards understanding what tooling is required if this process is to use XML and not plain text as the common format.*

**Q8 - How often do you use the following conversion tools (however those tools are hosted)?**

The dominant conversion tool is as expected *xml2rfc*, however the other tools are still well used and need to be maintained.

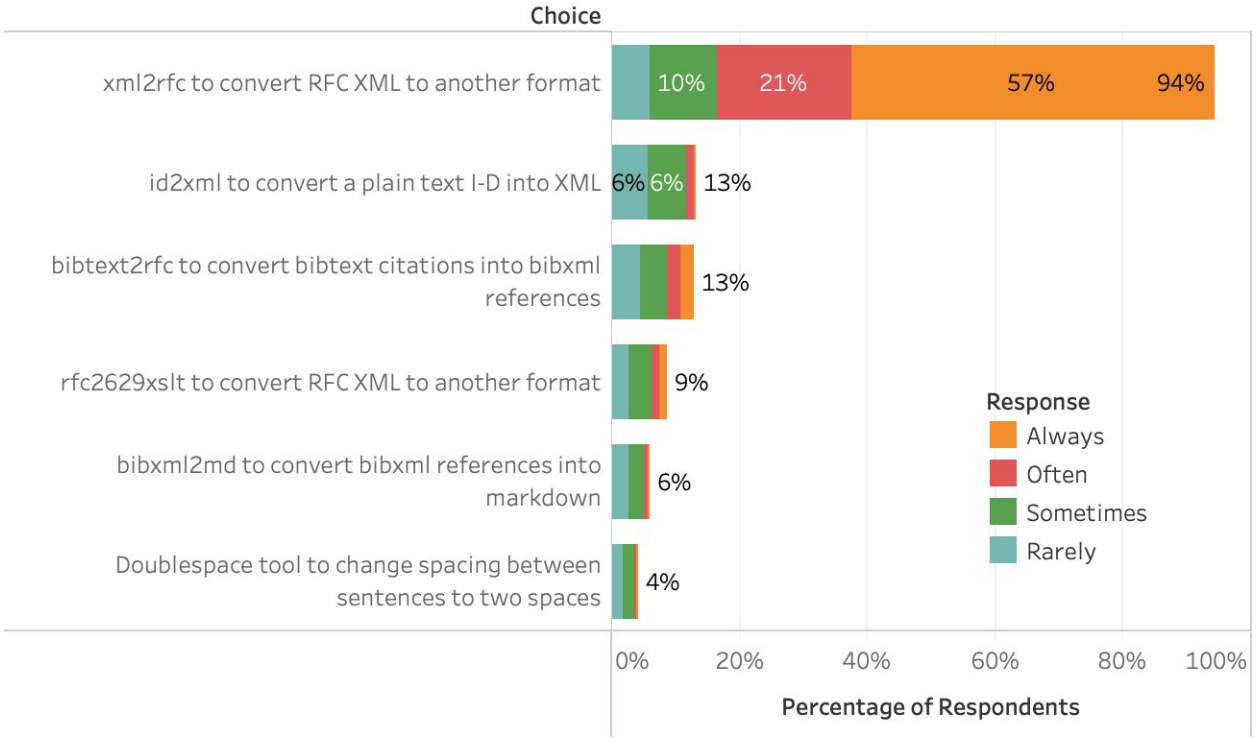


Figure 6: Q8 - individual responses

One interesting result is how often *id2xml* is used, which indicates that some of those who author in one of the non-XML formats that do not convert directly into XML as say markdown does,  are still interested in how that translates into XML.

# Part 3 - How people run their tools

**Q7 - How do you ensure that your draft is correctly formatted/validated when you submit it? (check all that apply)**

While the strong preference is for people to run the tools prior to submission, though a sufficient number use the submission form as their error checking:

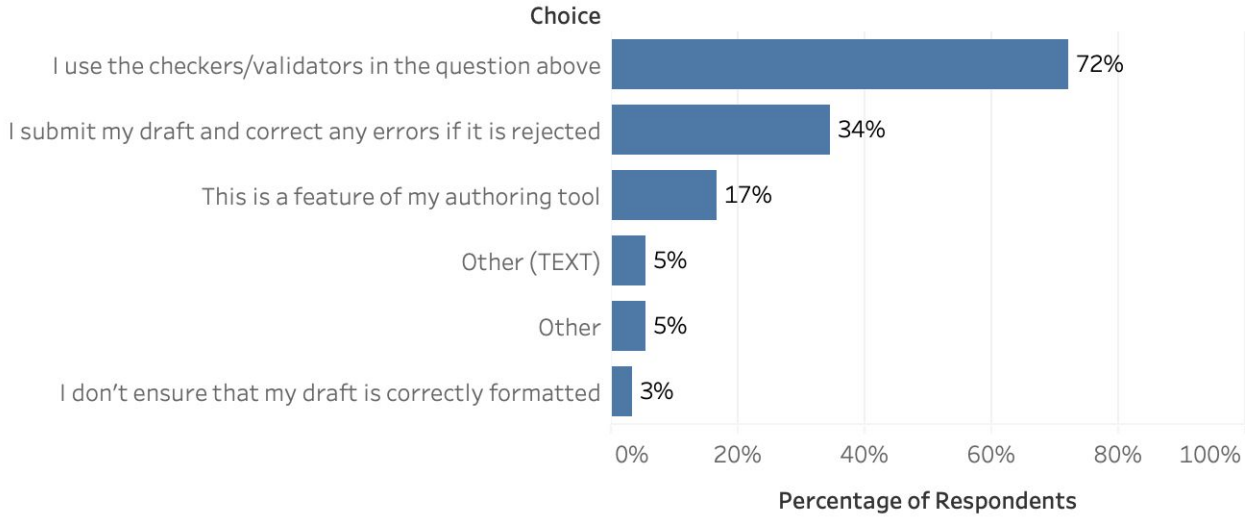


Figure 7: Q7 - individual responses

**Q9 - How do you run your tools? (check all that apply)**

This tallies with the answer above as the majority of respondents run their tools locally.

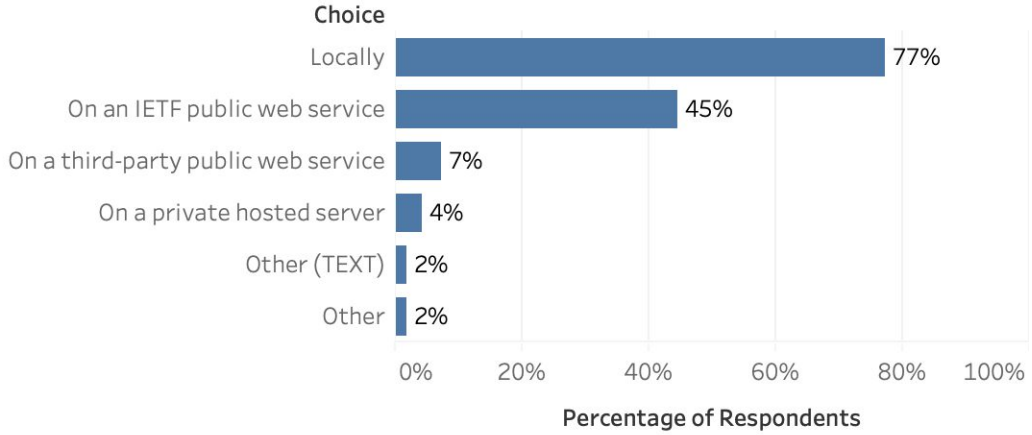


Figure 8: Q9 - individual responses

**Q10 - Do you run an automated build process for producing I-Ds?**

While the majority of respondents say no, enough do that we need to be sure we understand how our tools are used in an automated build process and ensure that they are designed to properly support this usage.

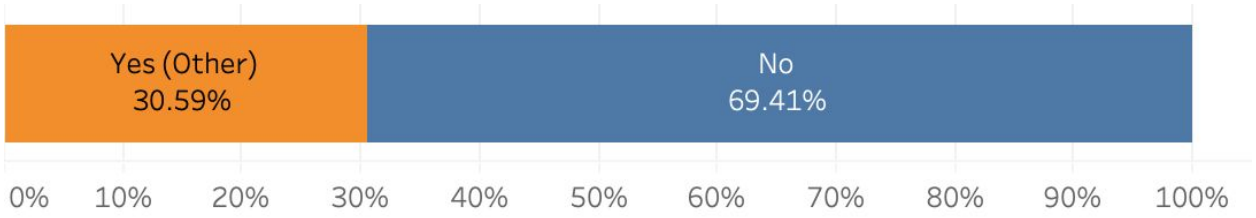


Figure 9: Q10 - individual responses

**Q11 - Please list any other tools that you use and the purpose of those tools**

This was a free text question, with the following key points noted:

- git and GitHub appear multiple times.
- Multiple people use relatively unsophisticated editors (e.g. notepad)
- Only a few people use modern XML editing tools such as Oxygen.
- A number of people use Emacs Org Mode, which is a form of markup not identified before the survey and so not included as a specific choice.

**Recommendation 6**: *This survey has identified four broad categories of client that are in use: Editor, Build System, Web Service and Command Line Tool; and four broad categories of functionality provided by the tools: Editing, Boilerplate, Validation and Format Conversion. An architectural model is needed that sits above and incorporates these categories of clients and functionality, and determines how new and existing services/tools integrate, easily and efficiently.*

# Part 4 - v3 XML adoption

**Q12 - How do you rate your knowledge of the v3 RFC XML format?**

The majority (54%) rate their knowledge of the v3 XML as Poor or None, with only 20% rating it either Good or Excellent.

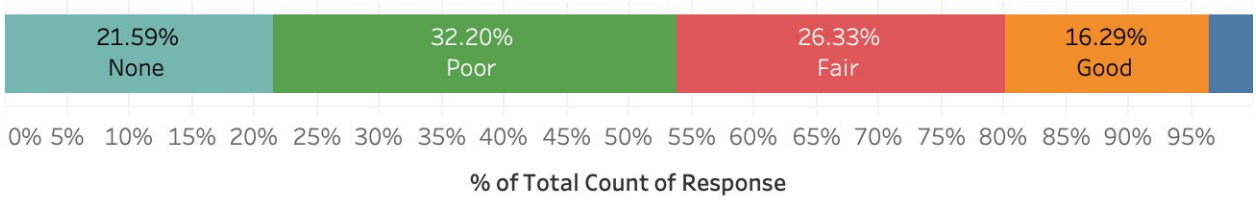


Figure 10: Q12 - Individual responses

It seems therefore that while many people know and use RFC XML, that knowledge is more around the v2 XML and than the v3.

**Recommendation 7**: *Produce specific education materials around the changes between v2 and v3 XML and include in the micro-site of recommendation 3 and also proactively push to authors and tools developers.*

**Q13 - How satisfied are you with XML as an authoring format?**

Despite the underwhelming level of knowledge of v3 XML, satisfaction is still high with only 21% either dissatisfied or very dissatisfied. A large number of people (34%) remain undecided, and they will need to be tracked over time to see if they move into either satisfied or dissatisfied.

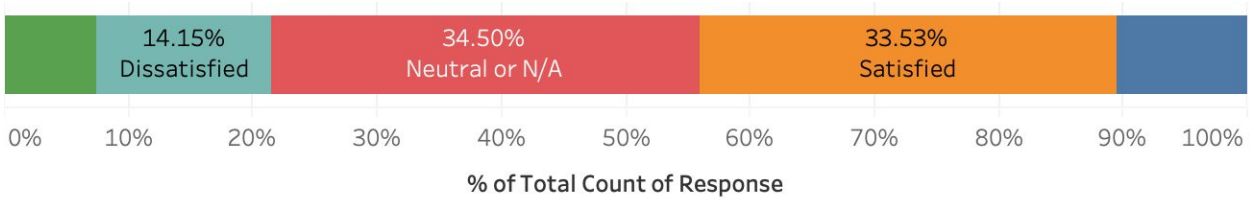


Figure 11: Q13 - individual responses

The message here appears to be that adopting XML as the canonical format switching to XML was not a mistake, though some remain to be convinced, and those for whom XML is not acceptable are a relatively small proportion.

**Recommendation 8**: *An end-to-end analysis is carried out to ensure that v3 XML can be used effectively and seamlessly at all stages of the authoring process.*

**Q14 - How important are the following characteristics of the v3 XML format to you?** and **Q15 - How satisfied are you with the following characteristics of the v3 XML format?**

These two questions form a pair for an importance-satisfaction gap analysis, which is used to prioritize areas for improvement. For this technique we calculate the mean importance score for each characteristic and then subtract from that the mean satisfaction score. The higher the difference between the two then the higher the priority for improvement.

| Characteristic | Importance | Satisfaction | Gap |
|---|---|---|---|
| Tools support | 0.91 | 0.23 | 0.68 |
| Documentation | 0.83 | 0.15 | 0.68 |
| Features | 0.54 | 0.26 | 0.28 |

This suggests that our top priorities are to improve the support for v3 XML in our tools and to improve the documentation. The next question goes on to double check this answer.

**Q16 - The majority of the inputs sent to the IETF internet-draft submission tool are still predominantly in plain text and v2 XML. Why do you think submission of v3 XML is still so low? (check all that apply)**

The top three answers to this question are all about knowledge and education, partly confirming the answer from the previous question, though not supporting the perception of the gap regarding tools support.

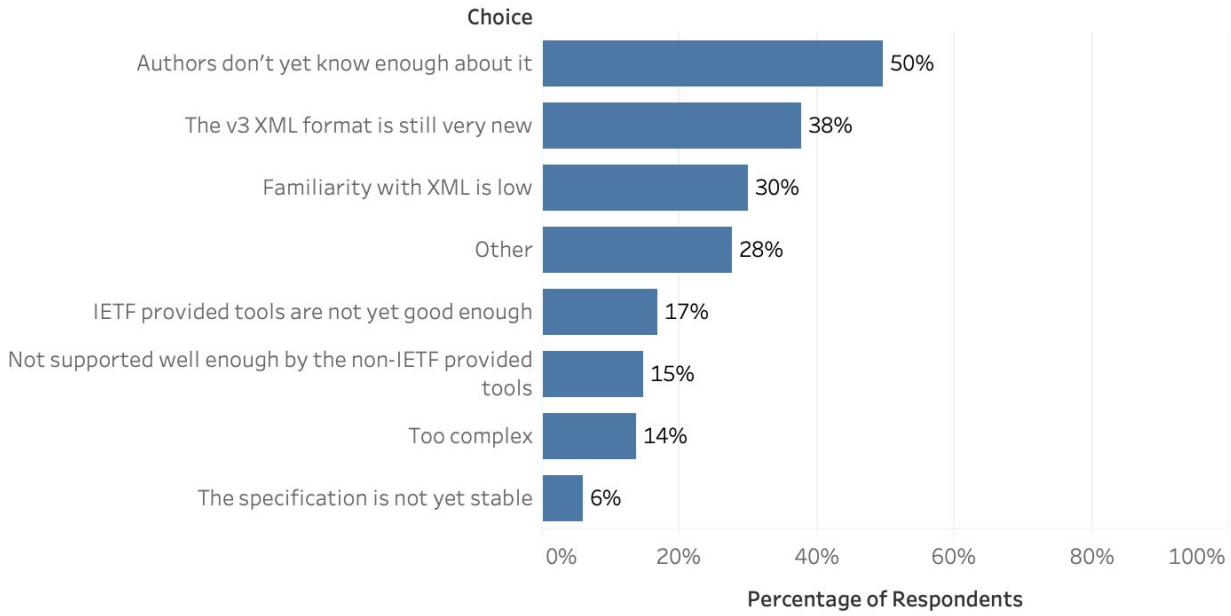


Figure 12: Q16 - individual responses

**Q17 - What more needs to be done to support the rollout of the v3 XML format?**

This was a free text question with many responses. The key themes are:

- Better documentation and tutorials (supported by the evidence).
- XML is too hard, use Markdown instead (this is not supported by the evidence).
- One-stop shop hosted tool that does everything.
- Change the submission process to favour or force v3 XML.

**Recommendation 9**: *A significant minority want to use Markdown or similar and a strategic decision is needed on whether or not this should be "officially" supported either in some stages or in the full end-to-end authoring process. (see also Q20)*

**Recommendation 10**: *A 'one-stop shop' tool should be investigated.*

# Part 5 - Future planning of tools overall

**Q18 - How Important are the following characteristics of authoring tools to you?** and **Q19 - How satisfied are you with the following characteristics of authoring tools?**

This is another pair of questions used for importance-satisfaction gap analysis as explained for Q14/Q15. In this analysis some of the gaps are negative indicating that the satisfaction is more than is necessary to meet the importance and therefore more nothing needs to be done on those characteristics.

| Characteristic | Importance | Satisfaction | Gap |
|---|---|---|---|
| Ease of use | 1.55 | 0.46 | 1.09 |
| Overall quality | 1.29 | 0.48 | 0.81 |
| Integration with IETF process | 1.21 | 0.65 | 0.56 |
| Control of output | 0.80 | 0.47 | 0.33 |
| Support for the full range of tags / metadata | 0.61 | 0.36 | 0.25 |
| Integration with version control systems | 0.60 | 0.35 | 0.25 |
| Support of various output formats | 0.52 | 0.56 | -0.04 |
| Choice of different tools | 0.14 | 0.28 | -0.14 |
| Speed at which new features are added | -0.08 | 0.21 | -0.29 |

These results are quite clear - our tools need work to raise them to the standard of ease of use and quality that people expect.

**Recommendation 11**: *Initiate a project to raise the overall ease of use and quality of authoring tools, probably involving a UX professional to create a consistent interaction model for all IETF-supported tools. This is closely linked to recommendation 8 as the end-to-end support of tools is a key element of their ease of use.*

**Q20 - How important is it for you for any new toolchain to support the following authoring formats?**

For this question we need to look at the total figures for ‘very important’ and ‘important’. (Note, this question mistakenly used a different importance scale from the other questions but it is unlikely that had any impact).

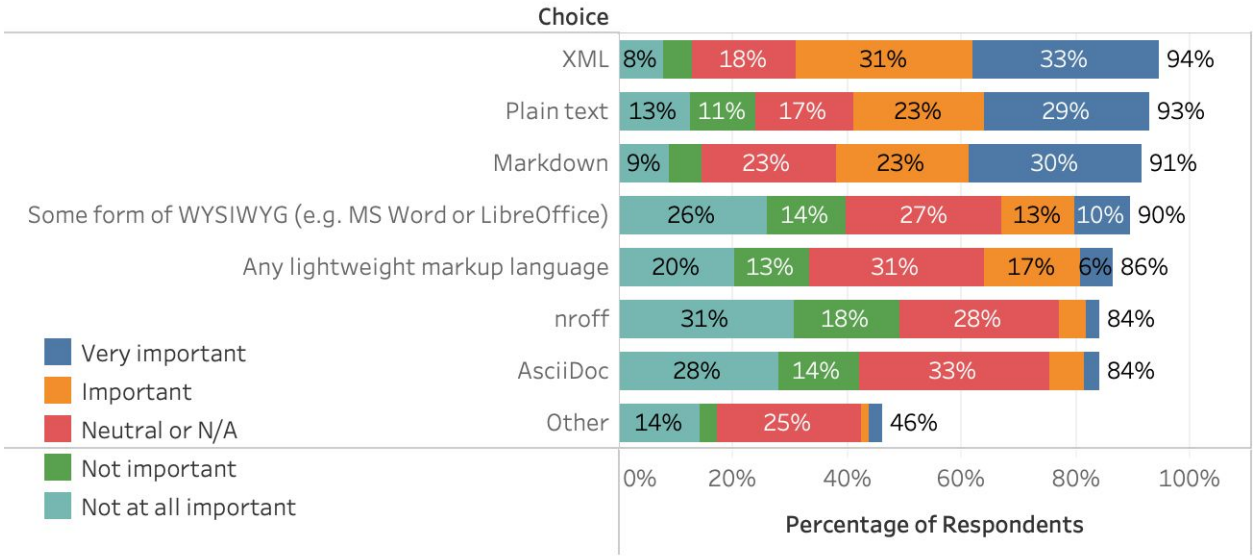


Figure 13: Q20 - individual responses

The top three are XML (64%), Markdown (53%) and Plain text (52%). Given that this question asked about Plain text separately from all simple markup languages it seems that people are specifically identifying support for un-marked up plain text.

Recommendation 10 above already covers investigating additional support for Markdown.

**Q21 - Do you author standards in another Standards Development Organisation?**

This question was in two parts, the yes/no answer and then the text to identify the other SDOs, with 34% of those who answered this question (148 out of 432) saying yes. A simple analysis of the text responses gives the following top answers:

| **SDO** | ITU | IEEE | W3C | ETSI | 3GPP | ISO |
|---|---|---|---|---|---|---|
| **Count** | 31 | 27 | 23 | 19 | 18 | 12 |

Several more SDOs were mentioned 1-5 times.

**Q22 - How does your experience of authoring tools in other Standards Development Organisations compare to IETF authoring tools?**

There is a suggestion that the wording of this question was ambiguous and therefore some people may have answered “better” to mean the other SDOs are better than the IETF, while others answered that the IETF is better than other SDOs.

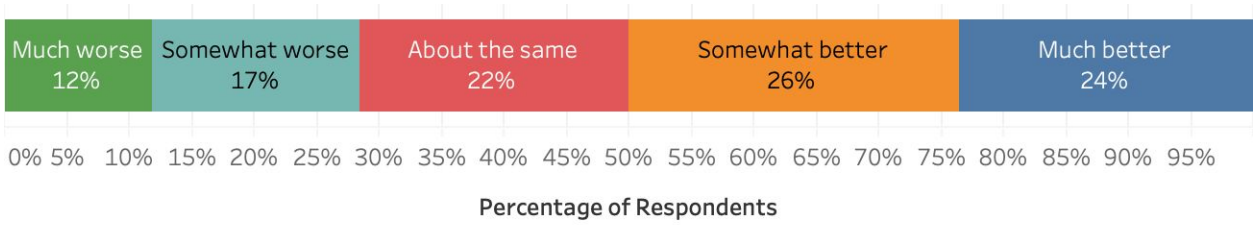


Figure 14: Q22 - individual responses

With that warning in mind, the figures do not look good for the IETF with 50% rating other SDOs as better and only 29% rating the IETF better than other SDOs.

**Recommendation 12**: *A comprehensive study is needed of other SDOs and their authoring tools and processes to see what the IETF can learn and/or adopt to improve its own processes and tools.*

**Q23 - Is there anything else you would like to say about I-D authoring tools and formats?**

This was a free text question with most of the answers following similar themes to the answers to questions above. Some new themes to emerge were:

- Many documents are authored by a team who need to share the source, track changes across the team, discuss changes etc, and the tools need to recognise and support this.
- Various recommendations about what tools should be officially supported.
- The IETF should have a recommended toolchain.
- Better communication needed about the processes and decisions of the RPC.
- The IETF should look at completely different authoring processes (e.g. ReSpec).
- The IETF is generally making the process of document production too complex and so restricting those that can participate to those with the necessary technical skills.
- The IETF should join with other SDOs to create a shared toolset, or the IETF should adopt the toolset of another SDO.

Most of these will fall out of the various recommendations above, with one notable exception addressed in the recommendation below:

**Recommendation 13**: *A conversation needs to start on whether our tools should be adapted to integrate with one or more collaboration platforms or whether collaboration is achieved by third party services and tools that work side-by-side the IETF tools.*