Validation Tool

The information contained in this document is provided solely for allowing potential bidders to provide a response to for the services being procured. While you may use the document for internal business purposes as they apply to your response, the information is confidential and must not be shared further.

About COUNTER

COUNTER brings the knowledge community together to agree and adopt the global standard for measuring and reporting content usage. We work with libraries, consortia, publishers, aggregators and technology providers, as well as collaborating with other open infrastructure and standards bodies.

Project background

The original COUNTER Validation Tool was developed as part of our work for Release 5 of the Code of Practice. The original setup had a user interface (https://stats.redi-bw.de/counter-validation-tool/) which was closely coupled with a testing module running on PHP with a Laravel framework. We have subsequently updated the Testing Module that performs the validation to work independently of the user interface, and the updated Validation Tool that uses the Testing Module is available through GitHub (https://github.com/ubfr/counter-r5-validation-tool). It still runs the validation, which is quite memory intensive, in the context of the webserver process. This separation of the Testing Module from the user interface informs our project goals, as outlined below.

Project goals

Assessment Tool

We wish to build a new COUNTER Validation Tool consisting of a user interface, administrative tools, control systems, and COUNTER API (sushi) tests. The Validation Tool makes use of the Testing Module, which will remain separate. The bullets below describe our requirements for the Assessment Tool.

Essential: user interface

- Allow users to register to use the Validation Tool.
- Allow users to upload reports in all supported file formats (JSON, TSV and XLSX). Note that the Testing Module allows for other formats such as ODS and CSV, which are not specified in the Code of Practice. Including the ability to upload additional spreadsheet formats would therefore be appreciated.
- Allow users to obtain all validation results associated with their accounts (Release and Report IDs, and error summaries).
- The interface must meet accessibility standards and should ideally be responsive design.
- It should be possible for the COUNTER team to update text and possibly imagery on the interface without relying on vendor support.
- The interface must not be tied to a specific release of the Code of Practice.

**Essential: administrative tool**
- GDPR-compliant user management.
- GDPR-compliant credential management (currently stored only during a session).
- Automatic deletion of stored reports after a configurable period (currently 7 days).
- Access to the results database for administrative users. This means Release and Report_IDS, error summaries, and key validation information (time, memory used, etc.).
- We would like to see anonymised analytics around use of the Validation Tool.

**Essential: COUNTER API (sushi) tests**
- Provide a UI for returning results on tests against the report provider’s COUNTER API (formerly sushi) server setup.

**Desirable: control systems**
- A queueing mechanism for report validation requests, such that the Validation Tool can control the number of validations running at any given time and run them independent from the webserver process.
- A reasonable size-limit on files submitted for validation, to be determined in discussion with the COUNTER team.

**Essential: open source**
- We require the Validation Tool to be made available under an open source license.
- It must be possible to run the Validation Tool using only open source components.
- We would like a notification mechanism to mitigate the risk of users running older/unpatched versions of the Validation Tool.

**Additional access**
We appreciate that some users may want to use the Testing Module separately from the Assessment Tool, and that users with files exceeding the size limit should be pointed at an option to download and run the Testing Module locally. We therefore have some additional requirements, which could be delivered in a second phase of the project if need be. They are:
- API access for reports within the size limit. We think it would be sensible to require registration for API access as for the UI.
- Bundling the validation module within e.g. Docker to facilitate local hosting. This applies to the Testing Module only, and does not require the Validation Tool itself be downloadable.

**Future extensions**
We plan to continue working on the Testing Module. At present it tests for Release 5 compatibility, but we are extending it to include Release 5.1 and wish to include new tests over time.

Beyond the Testing Module, we may wish to extend the Validation Tool itself in future. For example, allowing users to submit alternative tabular report formats such as ODS or CSV, or creating additional COUNTER API (sushi) tests for exceptions.
If you are able to provide support for either aspect of this work, please include that information in your response.

**Useful information about the current Testing Module**

The Testing Module is implemented in PHP, it can be used in PHP applications as a library (via Composer) or as a command line tool. The Testing Module supports both the validation of COUNTER report files and COUNTER API requests and responses. The results can be returned in JSON format, so that an application using the Testing Module can process and display the results.

During 2023 the Validation Tool with the current iteration of the Testing Module performed validation tests on 19,543 reports. 95% of reports required less than 10MB to run, and 79% of reports returned results in less than 0.1 seconds. On the larger end of the scale, only 11 reports during 2023 required over 1GB and only 104 took more than 1 minute to run.

The memory for the PHP processes currently is limited to 2GB. Uploads are currently limited to 40MB, though we have not placed a size limit on JSON report sizes fetched via the COUNTER API as parsing JSON reports requires significantly less resource than tabular reports. We are happy to consider tightening these limits in the interests of facilitating the additional access requirements above.

**Memory requirements**

<table>
<thead>
<tr>
<th>number of validations in 2023</th>
<th>memory used</th>
</tr>
</thead>
<tbody>
<tr>
<td>16783</td>
<td>less than 4 MB</td>
</tr>
<tr>
<td>1454</td>
<td>4 MB to 8 MB</td>
</tr>
<tr>
<td>676</td>
<td>8 MB to 32 MB</td>
</tr>
<tr>
<td>360</td>
<td>32 MB to 128 MB</td>
</tr>
<tr>
<td>196</td>
<td>128 MB to 512 MB</td>
</tr>
<tr>
<td>63</td>
<td>512 MB to 1 GB</td>
</tr>
<tr>
<td>11</td>
<td>1 GB to 2 GB</td>
</tr>
</tbody>
</table>

**Usage patterns**

<table>
<thead>
<tr>
<th>number of validations in 2023</th>
<th>hours of use in 2023</th>
</tr>
</thead>
<tbody>
<tr>
<td>more than 200</td>
<td>1</td>
</tr>
<tr>
<td>150 to 200</td>
<td>1</td>
</tr>
<tr>
<td>100 to 150</td>
<td>8</td>
</tr>
<tr>
<td>50 to 100</td>
<td>65</td>
</tr>
<tr>
<td>25 to 50</td>
<td>109</td>
</tr>
<tr>
<td>10 to 25</td>
<td>313</td>
</tr>
<tr>
<td>less than 10</td>
<td>1419</td>
</tr>
<tr>
<td>None</td>
<td>6844</td>
</tr>
</tbody>
</table>

Bernd Oberknapp at Freiburg University Library, a crucial member of the COUNTER team, developed and will maintain oversight of the Testing Module, though other members of the community are invited to create issues or pull requests once the new Testing Module is on GitHub. He reports that the new Testing Module will require more memory and more CPU cycles since it is based on objects instead of arrays for processing and storing the data; this has been done to facilitate maintenance and future functionality extensions.
How to respond

Budget and future plans

We have a good understanding of the requirements of this project following our initial build and subsequent activities to develop the Testing Module and Validation Tool, and have set a budget commensurate with our expectations.

We are seeking a partner to support us in the long term through hosting and maintenance of the new Validation Tool, so would like to see annual hosting and maintenance pricing as part of your proposal, but clearly separated from the initial build costs.

Questions

We understand that questions are likely to arise while you plan your response to this RFI. Please send all questions to tasha@countermetrics.org by 17.00 BST on the 6th of June 2024. Any questions received from any potential vendor will be compiled and a package of responses sent to everyone within one week of that date.

Your response

Please provide a response in no more than ten pages, including

- A description of your technical approach to delivering on our project goals.
- A description of your project management approach.
- A summary of the proposed budget, covering both development and the long-term maintenance costs.
- Very short biographies of key personnel.
- Details of at least two referees for whom similar work has been undertaken.

Please send your response to tasha@countermetrics.org by 17.00 BST on the 27th of June 2024. Any responses received after this time will not be considered.

Vendor selection criteria

Proposals will be scored by a selection panel of members of the COUNTER Executive Committee against the following criteria:

- Quality of proposal. The extent to which your response suggests you understand and can meet our project objectives.
- Technological robustness. Use of well-supported technologies and delivery of technical support and advice for our team through the build.
- Long-term support. The level of support you offer as a hosting provider.
- Value for money. Both build and maintenance costs should be commensurate with the nature of our requests.
- Track record. Descriptions of previous similar projects and recommendations from previous clients.

COUNTER reserves the right not to commission the work.