NAACCR Plan to Implement XML

November 17, 2017

NAACCR is now firmly committed to changing its Standards Volume II data exchange format from fixed-width to an Extensible Markup Language file format (XML) by 2020. A plan and timeline for this change was approved by the NAACCR Board of Directors at their September 2017 meeting based on recommendations from the NAACCR XML Workgroup and the Standardization and Registry Development Steering Committee (S&RD SC).
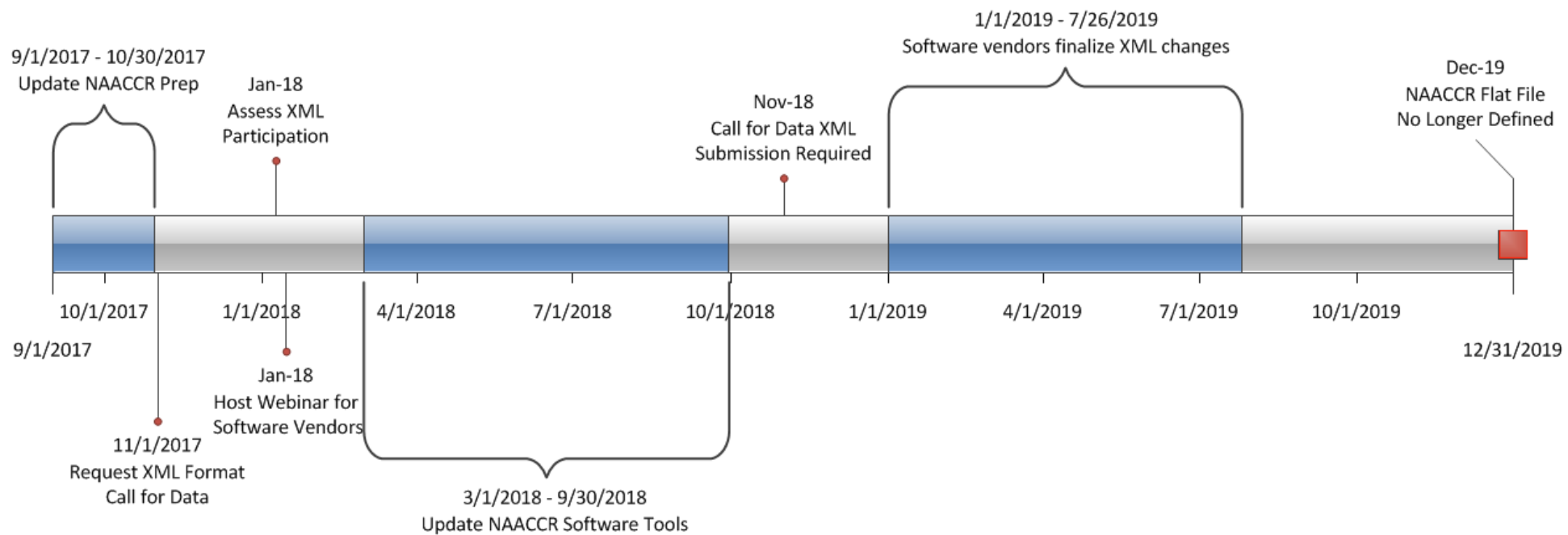
The move to adopt an XML-based data exchange format addresses the rigidity and extensibility limitations of the current fixed-width format as the NAACCR community seeks to provide leadership in an era of personalized medicine, increasing pace of scientific breakthroughs, and rapidly changing data standards. XML will provide greater flexibility in defining and transmitting NAACCR data items, allowing cancer surveillance to keep pace with the rapidly changing needs of the clinical, research, and cancer surveillance communities. XML will also eliminate the loss of data due to restrictions in the fixed-width format for large data items critical to cancer surveillance such as clinical text. In addition, it will enable greater semantic interoperability with healthcare data standards that have already moved to XML.

Transitioning from the fixed-width data exchange format to NAACCR XML will require a coordinated effort throughout the NAACCR Community, including input from many Steering Committees and Workgroups. Initial discussions between the XML WG, S&RD SC, and the Communications Steering Committee have already begun.

The NAACCR organization will be focused on meeting the needs related to this important transition to XML for the various cancer registry, cancer surveillance, and standard setter stakeholders. NAACCR will provide communication, webinars, documentation, other training, tools, and FAQs to support the rollout timeline. The following tasks will be part of that effort:

1. The NAACCR organization should prominently raise awareness of the freely available NAACCR XML software tools provided by SEER and the NPCR program that allow fixed-width NAACCR files to be converted into NAACCR XML.

2. The NAACCR organization updated the NAACCR Prep tool, posted 10/25/2017, to allow input and output of NAACCR XML in time for the 2017 call for data.

3. The NAACCR organization strongly encourages and accepts, but does not require, the 2017 NAACCR Call for Data submission files in the most recent NAACCR XML format (currently v1.3).

4. NAACCR should assess the rate of registry participation in XML format call for data submission. Non-participating registries should be surveyed to assess the barriers registries encountered in submitting XML files. This implementation plan will be modified as needed based on this assessment.

5. The XML Work Group shall provide a NAACCR webinar in early 2018 for all cancer registry software vendors (central registry, hospital registry, commercial and in-house) that will provide a forum to publicize the XML implementation plan and solicit information about vendor readiness to provide XML output capability within this proposed timeline.

6. By November 1, 2018, the NAACCR organization should update its software tools for processing the annual call for data files so that they can natively consume NAACCR XML instead of relying on a conversion to the fixed-width format. The current version of the NAACCR XML data exchange standard is backward compatible with the fixed-width format, allowing conversion both to and from XML using software tools provided free by the SEER and the NPCR programs. However, in order to provide a foundation for the most advanced features of XML, software tools that consume and analyze data exchange files should be updated as soon as possible to handle NAACCR XML natively.

7. The NAACCR organization should require the 2018 and all later NAACCR Call for Data submission files in the most recent NAACCR XML format. The S&RD goal is 80% by 2018.

8. Starting with the 2020 Volume II Data Dictionary changes (typically these would be planned in 2019 and rolled out for 2020 diagnoses and later), NAACCR will no longer support fixed-width formatting. All references to column start, column end, and any other metadata related to the fixed-width format will be removed from the NAACCR Volume II Data Dictionary and other NAACCR documentation where necessary. Once this happens, facility registry software will no longer be able to rely on a NAACCR-defined fixed-width data exchange format. This essentially will end support for the fixed-width format at both the state and national levels.

NAACCR Plan to Implement XML
November 2017