2021-08-10 Technical WG Agenda and Notes

Date

10 Aug 2021

Attendees

- John Kunze
- Karen Hanson
- Greg Janée
- Roxana Maurer
- Curtis Mirci
- Mark Phillips
- Dave Vieglais

Goals

New version of the ARK spec

Discussion items

Item	Who	Notes
announcements	John Kunze Dave Vieglais	**Dave:** While working on a legacy project with ARK identifiers, needed a normalization tool. Converted it into a Python library that does normalization, and then into a web tool: https://datadavev.github.io/SNARK. Discovered some ambiguity in the spec. Includes a NAAN lookup for the issuer. The tool is loading a browser version of Python. This is <u>not released</u> - early days but would like feedback. **John:** We received our 800th NAAN registration 2 weeks ago. Looking for assistance with curating the NAAN registry.
calls for papers, submission deadlines, upcoming meetings: Calendar of events		No upcoming meetings/presentations
New spec with changes from last time - diffs here - still work to do	All	**John:** - Now on version 28. Structurally not much changed. - Changed name of section "ARKS Embedded in Language" to "ARKS and Usability" - we have a wider range of inflections and no longer as constrained. - 5.5 added to link from spec to ongoing work at arks.org website. - Broke complex diagram into 2 less complex diagrams and introduced terminology needed for wiki article. Specifically - new concept of "Mapping ARK" to mean a fully qualified ARK. Also, the concept of a "Base Name" with suffixes and prefixes and the " "Check Zone" as NAAN + Base Name. - New table added to clarify how to apply the part names to a NAAN or Shoulder. **Greg:** How will these example paths that end with a NAAN or Shoulder resolve? Should they? **John:** Can add language to say these may or may not resolve, but idea is to provide a way to talk about these different concepts. **Mark:** I find the "mapping ark" concept useful. **Dave:** Does the ark: label have to be immediately after the base URL? Wasn't clear whether it could be further down in a subdirectory - would prefer to have that option. **John:** In that case we would need to parse left to right and look for first instance of "ark:" **Roxana:** Won't use that but can see the use case. Agree that it adds flexibility. Where the label starts, that's where the ARK starts. ***Action Item added*** **Dave:** Is there a way to log and track issues like this? Propose something like GitHub (we don't currently have one) but it requires some commitment. **John:** might be worth doing as we get ready to submit to IETF and open it to commentary.

Team: Agrees this would be a good thing

Mark & Dave volunteer - action item added

John: Other feedback on ARK anatomy section?

Karen: Suggestion to use Base Name consistently now that it is established – document switches between Name, Base, and Base Name.

Team: Generally agrees that this ARK anatomy description is clearer overall.

John: Has added a table to explain the shared NAANs.

Greg: How is uniqueness on 12345, 99999 controlled?

John: similar to other shared NAANs - if you want to resolve these, you can register a shoulder using the form. How to do this is now described later in the spec.

John: Has changed text related to normalization of hex digits

Dave: Suggests the ARK spec could benefit from clear guidance about transmitting ARKs. Conflation of ARK spec with URL spec could create some confusion. Application may expect ARK ID to be a part of a URL path component. Technically, if transmitting ARK as part of the path component, the slashes will be percent encoded to be embedded in the URL.

Greg: References a previous discussion long ago regarding proper URL encoding. The colon, for example, would typically need to be encoded. In practice, no software minds - decided to break the rules, but agrees the spec should have some recommendations.

Dave: Encoding would certainly be less streamlined and most implementations are ok without.

Action Item added

John: Added gentle warning about hyphens showing up in disguise as Unicode - should convert to hyphens first and normalize.

Greg: Spec uses the term "octets" throughout, but also multi-octet Unicode chars. Should be reconciled. New wording may be reaching out of the comfort zone of the spec - recommend not including the paragraph that refers to Unicode.

John: Prefer not to be too rigid. Need to be forgiving so that system does not break frequently. Want to indicate that there are measures that can be taken outside of the spec. May be better to suggest rather than stipulate

Dave: Seem more like something to watch out for than a processing rules of the spec. Could be a slippery slope otherwise - could say if something appears that is not in the character range, it could make the ARK invalid. Has come across this exact hyphen issue in other contexts, but if it is part of the spec it needs to be very prescriptive.

Greg: Agrees re: "slippery slope" references point #6 in normalization steps below, which mentions dash characters in Unicode specifically, if you mention these, do we need to mention others.

John: Need to repair this, but does want to acknowledge real world problems and be tolerant in what is accepted.

Action Item added

John: Added resolver chains section. Has at least one step, may have many steps.

Action items

- John Kunze Dave Vieglais Clarify that using a subfolder after the domain and before the ark: is acceptable
- John Kunze update arks.org to point to latest version of spec
- Mark Phillips Dave Vieglais Mark will research options for managing a spec through GitHub, Dave will come up with ideas/suggestions and work with Mark on a recommendation.
- Dave Vieglais Greg Janée Dave will suggest some text for a section or appendix on best practices for transmitting ARKs in URLs (eg, don't be too strict about %-encoding), Greg will review.
- John Kunze Changes to normalization steps. Point #5 in normalization fix "uppercase" and based on team discussion, point 6 on Unicode should be very loose, something that is good to be aware of but more of an implementation choice. Need to determine where to move it to if it is moved from here
- John Kunze Spec: reduce number of ways to say Base Name (instead of Name or Base)