

2019-04-01 Technical WG Agenda and Notes

Date

01 Apr 2019

Attendees

- [John Kunze](#)
- [Sheila Morrissey](#)
- Bertrand Caron
- [Mark Phillips](#)
- Adrien Di Mascio
- Greg Janee
- Curtis Mirci
- Tom Creighton
- Roxana Maurer

Goals

- Discuss proposed spec changes

Discussion items

Time	Item	Who	Notes
5m	questions about Experts Day meeting notes?	Curtis, Tom	<ul style="list-style-type: none">• Curtis has reviewed, no issues to report• Trailing ? possible issue, see below
5m	questions about Bertrand's URI compatibility discussion	Bertrand	<ul style="list-style-type: none">• To be discussed on next call, see https://groups.google.com/forum/#!topic/arks-forum/h48XhDk2Hh4• Bertrand: National Library of Finland - some concerns about how IETF will respond to RFC, Bertrand will share link with Google group
5m	NAAN registry URL confirmation	all	Sheila: correct entry, but Portico re-working landing pages for ARKs Mark: Correct entry Greg: (UCSB) will followup Curtis: EZID, works fine Adrien: will check Bertrand: works fine Tom: works correctly, registry entry correct

45m	prioritizing Experts Day spec issues	all	<ol style="list-style-type: none"> 1. <i>Literal character repertoire changes: allow '~', but disallow '#' (which is reserved in URIs fragments and LOD):</i> Greg – are there existing arks with # - if there was no objection, probably safe to assume ok ; to be clear: ok at end of URL, but not part of ARK id itself APPROVED 2. <i>Make the first '/' optional, so that <code>ark:/12345/678</code> is equivalent to <code>ark:12345/678</code>. This would match a near universal practice in other id schemes, and is a commonplace and understandable mistake that currently penalizes ARK users and potential adopters. Seems to have become common practice NOT to include first '/'; do not expect it to affect existing ARKs; TOM: caution - without / in FS environment, will 404; JOHN: when ready, publishing new ARKs, do not have to use initial '/'; TOM after change, must support existing arks (that have '/', but are submitted without; We are saying arks for an object, with or without slash, are equivalent to each other; should be able to resolve (to that same object) whether or not submitted with the '/' ; BERTRAND re ':' Is this a URI compatibility issue? JOHN: per spec should be encoded, so we will have to include this in compatibility question APPROVED</i> 3. <i>Parsers (resolvers) should check for inflections (final punctuation character combinations) before normalization of final structural characters ('/' and ':'), for example, given <code>"ark:/12345/678."</code>, parsers should check if <code>"."</code> is an inflection and only normalize to <code>"ark:/12345/678"</code> if no inflection is matched Resolving code should check for inflections first, strip off if found, then do resolution to stored ARKId when receiving an ARK for storage or for resolution; GREG: concerns; EG ':' at end of ARK that has been cut and pasted: is ':' part of ARK or just punctuation - Roxana: question about suffixes; JOHN: importance of normalizing ARKs before storing (i.e. strip off punctuation at end); GREG stress importance of knowing where an ARKId ends for e.g. sharing purposes - complication is that inflections have meaning, and might make it more difficult to "see" the end JOHN, reserving ? will help clarify, solve many problems, NOT allowing '.' as inflection operator will help clean up many; perhaps we need to make this prohibition more explicit; Tom making further point that proxies are stripping inflection ? already; JOHN possibility is to have N2T convert '?' to standard long-form query string that proxies would have easier time correctly resolving; perhaps we should add long-form equivalent for inflections, and just use ? and ?? as convenience/macro for user; TOM suggests using application specific approaches to returning different representations (different questions returned by inflections) JOHN have spec express way to use content negotiation mechanisms to express equivalences; PENDING: John will work on clarifying wording</i> 4. <i>Make the NAAN more flexible – instead of just 5 digits or 9 digits, allow any "beta-numeric" string (defined to be the same as noid repertoire: bcdghjkmnpqrstvwxyz0-9) with no runs of adjacent letters longer than two, eg. <code>ark/bc8/...</code> but not <code>ark/bcd8/...</code> CURTIS is there a need for this? JOHN benefits of shorter id; future for on-boarding other identifier systems (eg CrossRef DOI prefix mapping could be done); preserves opacity; might remove barrier for new registrants GREG might make ARKs less immediately "grok-able" as ARKS, but not showstopper; BERTRAND questions re: 9 digits – JOHN - none exist now, we can just suppress reference to 9 digits; CURTIS single digit Naan? JOHN yes possible in theory, but just because spec is relaxed doesn't mean NAAN assignment policies will change any time soon APPROVED</i> 5. <i>Update our understanding of what it means for metadata returned by inflections ('?' and '??') in 2018 to be both human- and machine-readable. In 2003, a simple email-header format (eg, ANVL) served both purposes, but now it is common to see a human-readable HTML landing page with machine-readable metadata embedded in it (where it doesn't interfere with the user experience). JOHN Currently simple to embed machine readable version in human-readable response (EG JSON) TOM strong encourage for future of spec, Curtis also APPROVED</i> 6. <i>Max link length for the ARKs : now 128 digit limit change the spec to eliminate the limit, but we will make a recommendation for a minimum of 255 characters supported in implementations. Strongly encourage supporting open ended lengths APPROVED</i>
-----	---	-----	---

Action items

- All: review issues surrounding URI compatibility issues for next meeting (<https://groups.google.com/forum/#!topic/arks-forum/h48XhDk2Hh4>)
- John: clarifications on final punctuation; use of content negotiation; draft changes for the spec
- All: review notes and edit as required