CMIS Unified Search

Design Discussion - 11 February, 2009

Attendees:
> Al Brown, IBM
> Paul Goetz, SAP
> Ethan Gur-esh, Microsoft
> Laurent Hasson, IBM
> Gregory Melahn, IBM
> John Newton, Alfresco

Topics:

1. Consensus on the proposed dates for proposing this change

   02/13 (First draft 0.1 posted to Oasis for subgroup comment) - *DONE*
   02/13 - 02/16 (Collect comments from subgroup - email)
   02/17 (Review meeting of subgroup - phone)
   02/20 (Second draft 0.2 posted to Oasis for subgroup comment)
   02/20-02/23 (Collect comments from TC - email)
   02/24 (Review meeting of subgroup - phone)
   02/27 (Third draft 0.3 posted to Oasis TC)
   by 03/15 (vote by TC)

2. Need to ensure that this service, which is really an observation pattern, can be extensible in the future to satisfy other use cases than search.   For example, if this were to be applied to audit as well as search, then it would need to also answer who changed the item, why it was changed (e.g. comments), what was changed (specific properties that were changed). Another use case to which this could be applied is replication. Perhaps this could be done by extending the schema in 2.0 to include change details such as that by extending *cmisChangedObjectType.*

3. A lot of discussion on the need for a way for a service consumer to resume where the consumer left off. One proposal is for a token to be generated by this service that could later be returned to the service as a way for the consumer to tell the provider where to begin to return results. This token would be opaque, verifiable and stable (for some repository-determined specified period of time).   Perhaps the token could be also be comparable for before/after comparisons but consensus was that this is not needed. If the consumer did not provide the token then the provider would answer the 'first' maxItems. The list could be derived by the producer from a repository query or it could be derived from a journal or transaction log. If the supplied token was not valid, because the repository could not produce the next page of items from it, then the producer would answer an exception and the consumer would need to start without a token (i.e. from the beginning).

4. The scope of the token was discussed - it could be at an item level or at a page level (probably the latter is sufficient).

5. An alternative idea is just to generate a 'tail' of the most recent changes and let the consumer navigate through the result until it discovered something it cared about (something it had not seen yet). This would be easier for the repository to produce but harder for the search consumer to process.

6. Another alternative would be to allow the consumer to specify a time period, and the provider would answer the first maxItems in that time period.

Action Items:

1. Greg to update the document today with a specific proposal - email discussion on that proposal would follow

2. Next meeting, tentatively scheduled for Tuesday 02/17/2009. Greg will send out an invite.