AAC meetings – Tuesday 10-4-16

* rules for matching strings to the data
* need to know what effort we need to plan over the next 7 months
* what will this look like in the end in terms of posting – how to handle root domains etc.
* what is the target model? Which fields are we aiming to have eventually
* Matt Lincoln – LOD pessimist – tricky to use
  + He’s conflicted
  + Assortativity/E-1 Index
  + British Museum released an LOD representation in a SPARQL endpoint
  + Also pulled JSON data from Rijksmuseum –
  + SPARQL query – not many good interfaces that help people compost queries
  + Need documentation of the data mapping to understand what is going on with the data underneath
  + Failure in finding “nationality” for Van Dyck (same for JSON and SPARQL)
  + How can we make clear to our users the state of our data – where do we have missing data, etc. Need to hint to our users…
  + He thought LOD was addressable, but LOD is just data on someone’s computer
  + Live SPARQL endpoints are finicky and hard to keep running
  + British Museum offered a bulk download – so he could build a SPARQL instance on his own computer for these queries
  + Need an indication of a “time out” when a query must get cut off – this was a silent failure – then you don’t know if you got everything from the DB
  + When it comes time to disseminate, he needs links to be able to do the same queries again. But since these are live db, you can never get the same results again.
  + Licensing and archiving various versions of data might be need to be supported
  + We need to start thinking of data like a biological organism, rather than books off a shelf
  + Replicability – give exact data and workflow to get the same results reported in an article for a journal. (important in science, etc.)
  + Reproducability – different purpose, to see if you get the same results – being able to see what changes is useful
  + “Using the same types of data will give the same types of results”
  + What is the shelf life of the data mapping?
  + How do you build the infrastructures for this?
  + LOD enables complex data expression – if the producer values that use case
  + Library of Congress – collections as data – keynote was wonderful
  + Our collections data produce possibilities for researchers –
  + By creating these db, we enable new generation of scholars to be wildly experimental in the research they do
  + We are creating the conditions for research possibilities
  + An enabling technology
  + Embed the researcher in our models for work – for funding as well
* Thomas Moran used Prang watercolors – (Amon Carter has that info)
* David Newbury
  + Created a utility to make it easy for us to post comments, questions, etc.
* **Github tutotial for AAC**
  + Notes from Craig – some changes are OK – if they have no impact on ISI work -- as long as we don’t change name, number or order of fields, check with Kate first, and then proceed cautiously