Notes on Vint Cerf discussions on Digital Preservation with CCSDS-DAI

5th April 2016

# Vint Cerf – points made

* He has so far read the first 300 pages of David Giaretta’s book, *Advanced Digital Preservation* (<http://www.springer.com/us/book/9783642168086>).
* He has some knowledge of our ISO work with OAIS Reference Model, Trustworthy Digital Repositories, Audit and Certification of TDRs, etc.
* He acknowledges that emulation is only one tool at best.
* “Digital Vellum” focused primarily on relatively simple electronic documents, software, and games.
  + Mike Kearney has requested Vint Cerf’s slides
  + Digital Vellum presentation: (<https://ecolloq.gsfc.nasa.gov/Current/announce.cerf.html>)
  + Similar presentation at: <https://www.youtube.com/watch?v=STeLOogWqWk>
* Compound digital objects present special problems.
* Interoperability is a key requirement. Mashing up emulated data with current data is problematic. He believes we need to demonstrate multiple trustworthy digital repositories that can interoperate to move the adoption of the standards forward.
* He has no suggestions for a business model for keeping the data usable for as long as it is needed.
* Google is working in collaboration with the Computer History Museum collecting software much like the NIST NSRL.

# Discussions arising from Vint Cerf’s comments – most after he left:

* The OAIS concept of Representation Information provides a general approach to interoperability. However more specific standards such as EAST or DEDSL can be used to provide some of the detailed Representation Information required.
* The users of two repositories with the same Designated Community should be able to understand information they exchange, at least for members of that Designated Community.
* A related but less specific aspect of interoperability would be for us to develop a formal API definition for the OAIS Archival Information Package (AIP)– based on the UML diagrams in OAIS. This would at least help to ensure that repositories can exchange AIPs and extract the data, Representation Information etc.
* The document DAI is working on currently will help in guiding the capture of information needed to exploit (including interoperating with other data) data being created.
* We could try to map DAI’s work and existing books to a more “protocol-like” view. For example, PAIS includes the definition of an information package which can be sent from a Producer to an archive, and the management processes required to organize this exchange.
* Mike Kearney asked if there are any services from DAI which SANA could host. David suggested that it would be possible for SANA to host an authoritative Registry of Representation Information; he has an example which is simply a webserver with a few very small scripts.