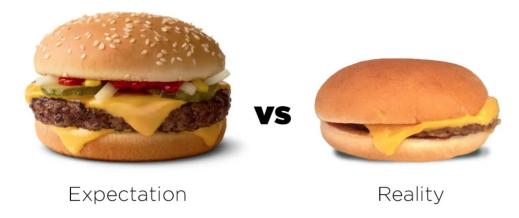# Hybrid Moments: Letters and Email in Conversation

Kathryn Antonelli
Martine A. and Bina Aspen Rothblatt Digital Archivist
American Philosophical Society (APS)

Disclaimer: Work in Progress

- Original question: How can we represent correspondents over time within the constraints of a traditional finding aid?

Expectation vs Reality

- This presentation: Our project, some challenges, implications for how ePADD can be integrated into hybrid collection processing, and our dreams versus our reality in building the finding aid

The Richard Garwin Papers

- Who is Garwin?
  - American physicist, best known for helping develop the H-Bomb
  - Later worked as an advocate for arms control, advised multiple presidential administrations
  - Had a 40 year career at IBM's Watson Laboratory - filed 47 patents
  - Elected to APS membership in 1979

- The Collection
  - Physical component: ~200 linear feet (unprocessed)
    - Incoming and outgoing correspondence, office files, personal papers
  - Digital component: ~8.25GB
    - Email, digital files
    - Digital files include outgoing physical correspondence as SCRIPTXS files and series of early emails from IBM in text file format

Physical Correspondence

- Work so far
  - Physical processing being done by Melanie Rinehart, an APS archivist
  - Correspondence from 1981-2010
  - 72 full bankers boxes -> 220 Hollinger boxes (92 linear feet)
  - Folder level processing - arranged by year, then correspondent
- Digital duplicates?
  - SCRIPTXS copies of outgoing correspondence
  - 1970s onward

RLG:mmr:005+LC (File's ID line if created by X.LET or X.MEM)
(Originator's initials:Staff initials:Filename:Date_plus 4 initials)
...SCRIPTed file follows its name:

 005+LC SCRIPT L1 005+LC SCRIPTXD L1

                    Richard L. Garwin
             IBM Thomas J. Watson Research Center
                      P.O. Box 218
                 Yorktown Heights, NY 10598
                     (914) 945-2555

January 5, 1979

Lawrence Cranberg, Ph.D.
Consulting Physicist
1205 Constant Springs Drive
Austin, Texas 78746

Dear Larry,

Norman Ramsey sent me your letter of 12/03/78 regarding the
APS and neutron therapy. From what little I know about this
problem, I am rather favorably disposed toward organizing
such a study.   Of course, we would have to find a sponsor,
funding, participants, and obtain authorization from the
Council of the American Physical Society.

I   hope that you will be interested in the enclosed
guidelines for APS studies, in the more detailed mechanisms
that POPA has developed in order to intercompare proposals
for studies, and in providing a more formal input of several
pages length to our subcommittee on studies so that we can
begin seriously to consider a POPA study in this field.

My assistant, Bill Gallagher, has provided me with the
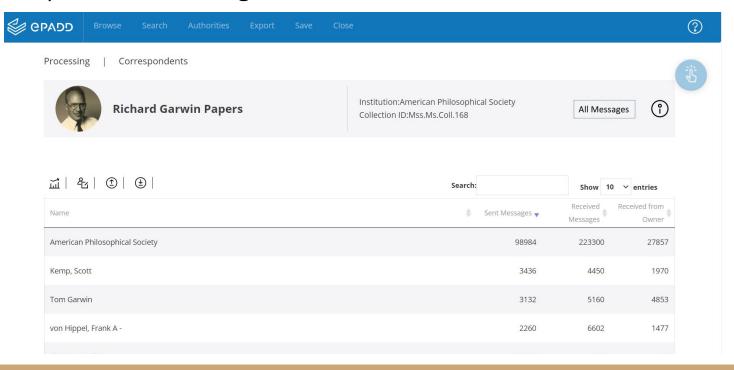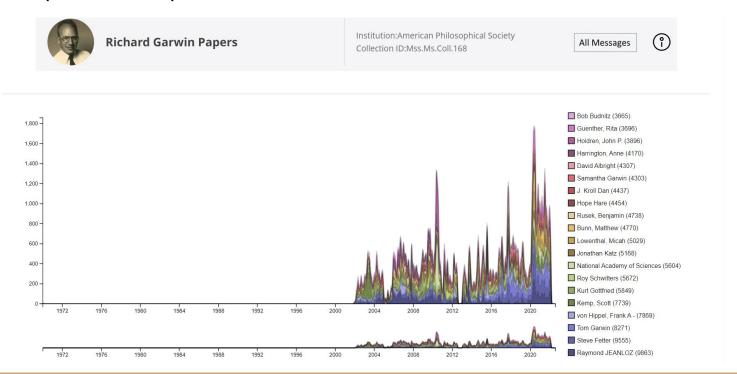enclosures, which will better inform you on the above

Email

- This is the first ePADD collection for the APS!
- Pre-weeding statistics
  - November 2001 - present
  - 320,137 emails, 54,811 correspondents
    - Wait...54,811??
- Need for multiple migrations may have impacted metadata
  - Garwin converted NSF (LotusNotes) -> PST
    - Done via MailsDaddy software before donation
  - Emailchemy license for ePADD converted PST ->MBOX
- Another factor: text files of emails c. 1990-2001
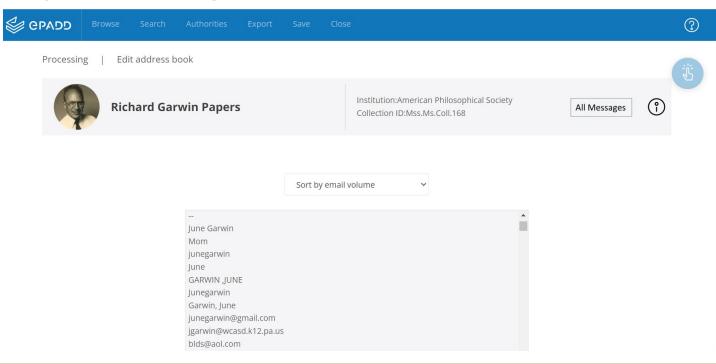 - Garwin migrated from IBM computer to text files

ePADD Capabilities: Correspondents

- Correspondent browsing view

ePADD Capabilities: Correspondents

- Correspondent report view

ePADD Capabilities: Correspondents

- Correspondent editing view

Our Correspondent Issues

	A	B	C	D	E	F	G	H	I	J	K	L
1	reference	names	email_addresses									
47	.nia.nih.gov	.nia.nih.gov	kom@grc									
48	.org	.org	a.gulick@deepcove;x-ray-list@lists.mitre									
49	.stanford.edu	.stanford.edu	martin@slac									
50	.WASH MIL-5-2 Dudde, Alexander	.WASH MIL-5-2 Dudde, Alexander	mil-5-2@wash.auswaertiges-amt.de									
51	/	/										
52	/a	/a										
53	/CN=Lotus	/CN=Lotus	mailfax@ccfax									
54	:	:										
55	;blankenship	;blankenship										
56	;dykhuizen	;dykhuizen										
57	;hurst	;hurst										
58	;Susan	;Susan	pitbulldurhams@aol.com									
59	?? ???? Suzuki	?? ???? Suzuki;??????? Suzuki;éˆ´æœ¨	atsuto.suzuki@kek.jp									
60	???	???	cchuang@trmc.aec.gov.tw									
61	??KuanKuan	??KuanKuan	minimermaid0308@gmail.com									
62	--[ML]	--[ML]	c1mariap@us.ibm.com;presletter@lists.brandeis.edu									
63	--[ML]	--[ML]	jeffrey@armscontrolwonk.org;fxbcenter_info@hsphsun3.harvard.edu									

	A	B	C	D	E	F
1	reference	names	email_addresses			
246	Ã‰ditions Odile Jacob	Ã‰ditions Odile Jacob	noreplay@odilejacob.fr			
247	Ã‰ditions Odile Jacob	Ã‰ditions Odile Jacob;Odile Jacob;Ed	info@odilejacob.fr;o.jacob@odilejacob.fr			
248	å´	å´	wubinglin1940@163.com			
249	å´æˆæ±Ÿ	å´æˆæ±Ÿ	xiangshan80@163.com			
250	å´æ˜¥åŽ	å´æ˜¥åŽ	wuchuenhua1945@163.com			
251	å´çŽ‰æ°	å´çŽ‰æ°	wuyujie2013@cgnpc.com.cn			

ePADD Capabilities: Correspondents

- Disadvantage of first collection: Would correspondents have been cleaner if the emails were originally PST/MBOX? You tell me!
- Email/digital versus physical correspondents
  - Cleaning mostly done, but need to test what that does to ePADD contents on reingest...

ePADD Capabilities: Other

- Entity analysis
  - People, places, corporate names
  - Can be kept in mind for subject headings
- Authority control
  - Can be used as a quick way to find "correspondents of note"
- Lexicon search
  - Can be used to find PII or certain topics for series
  - Get creative by making your own lexicon

Workflow Possibilities

- Should we process email first?
    - ePADD can help draw out subjects, important correspondents
    - ePADD can find terms that you're interested in, sometimes including PII
- Don't overlook the potential issues
    - Email may be different time span - focus of career might change
    - Personal versus professional correspondents - the move to texting
        - This deserves more discussion!

Finding Aid: Expectations

- Need to work within constraints of existing style sheet

PARENT NAME
     CHILD NAME   **YEAR**   **BOX/FOLDER**
so...
PARENT NAME
     CHILD NAME   **YEAR**   **BOX/FOLDER**
     CHILD NAME   **[null]**   **ePADD (link to search)**
     CHILD NAME   **[null]**   **SCRIPTXS file (link to DigLib)**
     CHILD NAME   **[null]**   **IBM file (link to DigLib)**

- Way too labor intensive! Over 17k SCRIPTXS and IBM files
  - Plus, file names logical but not clear/obvious

Finding Aid: Reality

- Note with ePADD link in Scope and Content and at top of Correspondence Series
- Separate link - downloadable file with list of correspondents that appear in email and physical correspondence
  - Have correspondent output from ePADD
  - Can search digital files for names from physical list
  - This at least will give researchers a heads up and easier/faster search method
  - Balance between staff time and access
- Similar solution (downloadable guide) for SCRIPTXS outgoing correspondence and text files of pre-2001 emails
- Hoping to still do some fun correspondent analysis for future Garwin digital project!

Thank you!

kantonelli@amphilsoc.org