Archival Arrangement and Description in the Cloud
A Preliminary Analysis

Christopher J. Prom
Born Digital in the Cloud: Challenges and Solutions
Archival Science Colloquium, Archivschule Marburg
8th June 2016

Provenance over time

"The principle of provenance is not a new idea, but a theoretical formulation based on experience."

"Archival practices and theory are based on and validated by experience and practice."

--Brichford, Maynard. "The Provenance of Provenance in Germanic Areas." *Provenance, Journal of the Society of Georgia Archivists* 7, no. 2 (1989). http://digitalcommons.kennesaw.edu/provenance/vol7/iss2/5.

InterPARES Trust

Project NA-22

"Arrangement and Description in the Cloud"

Giovanni Michetti
Jennifer Zilm
Darra Hofman
Katherine Timms
Glenn Dingwall
Sarah Romkey
Christopher Prom
Jennifer Douglas
Victoria Lemieux

Luciana Duranti
Corinne Rogers
Joseph T. Tennis
Jim Suderman
Patricia Franks
Richard Pearce-Moses
Andrea Tarnawsky, GRA
Richard Dancy

Cloud Computing Definition

"[A] model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction."

--Peter Mell and Timothy Grance, *The NIST Definition of Cloud Computing* (National Institute of Standards and Technology, September 2011). http://dx.doi.org/10.6028/NIST.SP.800-145

Cloud Characteristics and Models

Essential Characteristics: On-demand self-service, (2) Broad network access, Resource pooling, Rapid elasticity, and Measured service.

Service Models: Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS).

Deployment Models: Private cloud, Community cloud, Public cloud, and Hybrid Cloud.

InterPARES Trust

Roadmap to this talk

- Current standards/practice
- Cloud records from an arrangement and description perspective
- Specific practical problems and alternate models
- Provides baseline for more complete literature review and recommendations (forthcoming)

Core A and D Model

- Three essential concepts
  - describing aggregates
  - levels of control
  - Hierarchies
- Deeply embedded in current descriptive standards and practices.

--Fredric Miller, Arranging and Describing Archives and Manuscripts, Archival Fundamentals Series (Chicago: Society of American Archivists, 1990), p. 19-30. https://catalog.hathitrust.org/api/volumes/oclc/22456382.htm

--Kent Haworth, "Preface to the 1990 Edition," Rules for Archival Description (Ottawa: Bureau of Canadian Archivists, 1990, revised 2008), p. xviii. http://www.cdncouncilarchives.ca/rad/radcomplete_july2008.pdf

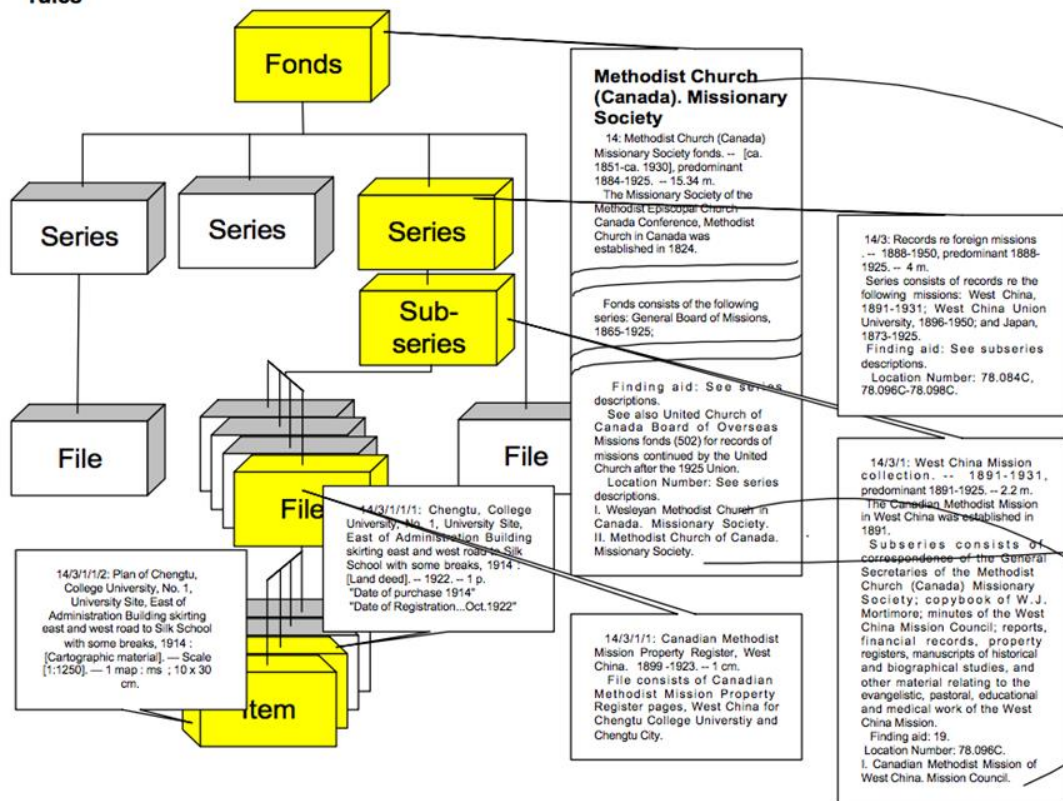InterPARES Trust

Core Int'l Standards

- **ISAD (G)**
- **ISAAR-CPF**
- ISDF
- ISDIAH
- Rules for Archival Description (RAD, Canada)
- Describing Archives: A Content Standard (DACS, USA)
- Encoded Archival Description
- EAC-CPF

InterPARES Trust

RELATIONSHIP BETWEEN DESCRIPTIVE AND AUTHORITY RECORDS
Illustrated with an example

Standardardized archival description(s) according to ISAD(G) or compatible national rules

Archival authority records according to ISAAR(CPF) or compatible national rules
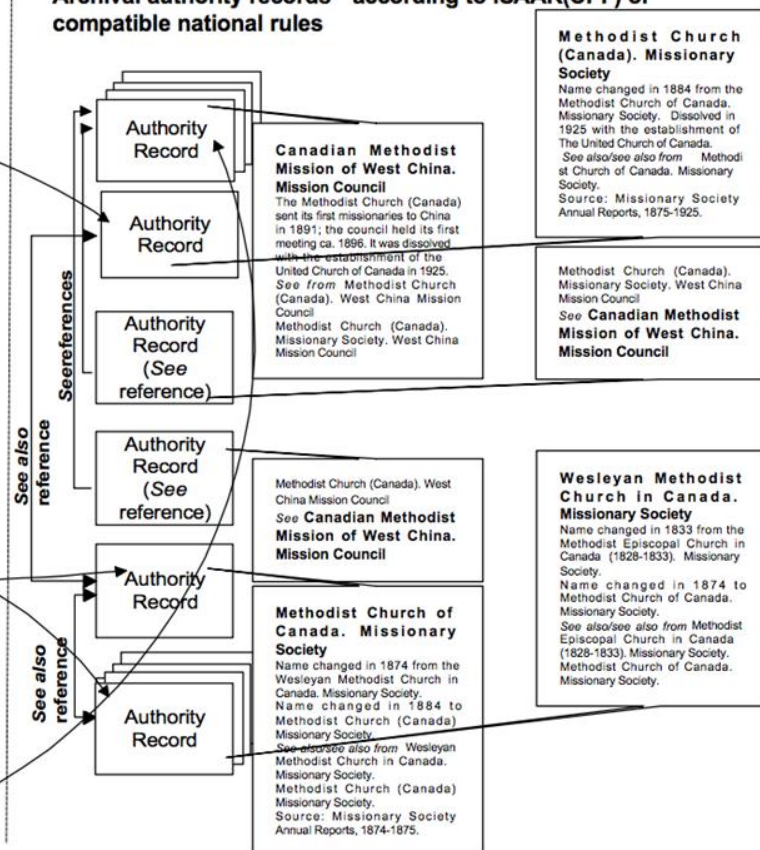
Chart designed and produced by Hugo L.P. Stibbe, c1998

--ISAD(G), page 37.

North American Standards

- ISAD (G)
- ISAAR-CPF
- ISDF
- ISDIAH
- **Rules for Archival Description (RAD): Canada**
- **Describing Archives: A Content Standard (DACS): USA**
- Encoded Archival Description
- EAC-CPF

Data Transmission Standards

- ISAD (G)
- ISAAR-CPF
- ISDF
- ISDIAH
- Rules for Archival Description (RAD): Canada
- Describing Archives: A Content Standard (DACS): USA
- **Encoded Archival Description (EAD)**
- **Encoded Archival Context (EAC-CPF)**

Authority-Based Models in Canada

--Steve Billinton, "The Role of Archival Authority Records in the Finding Aid System of the Archives of Ontario", Journal of Archival Organization, 5:1-2 (2008), 75-93.

--Hudson's Bay Company Archives, "Arrangement and Description of Records in HBCA: a Brief History". Retrieved June 1, 2016 from https://www.gov.mb.ca/chc/archives/hbca/holdings/arrangement.html;

--Davies, Dan. Thinking outside the box: re-imagining archival description with the 'series' system. MA Thesis, University of Manitoba, 2004.

Personal Archives User Studies

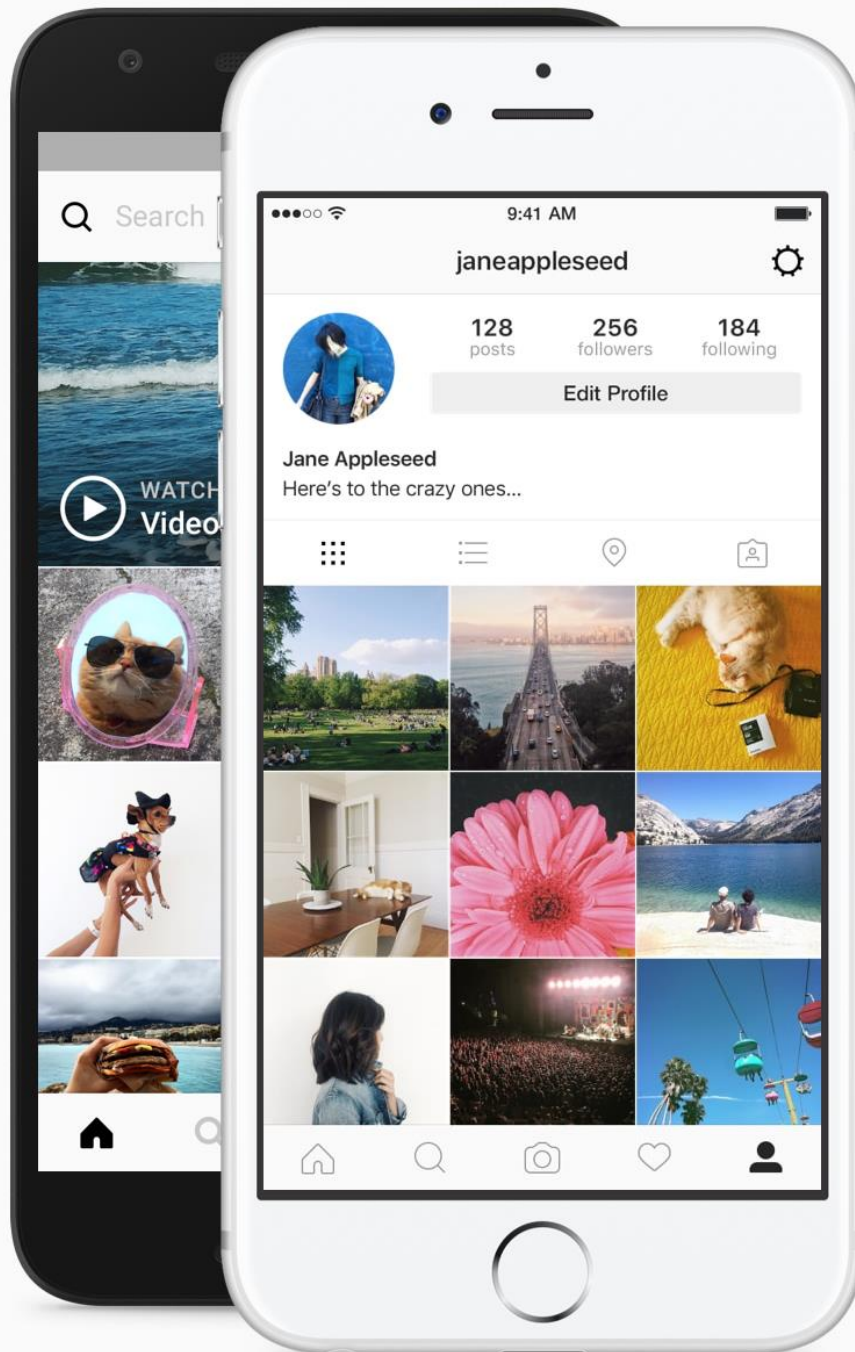--Bushey, Jessica. The archival trustworthiness of digital photographs in social media platforms.. University of British Columbia, PhD dissertation, 2016. https://open.library.ubc.ca/cIRcle/collections/24/items/1.0300440

--Toby Burrows "Personal Electronic Archives: Collecting the Digital Me," *OCLC Systems & Services:* 2, no. 2 (April 2006): 85–88. doi:10.1108/10650750610663932

--Clifford Lynch, "The Future of Personal Digital Archiving: Defining the Research Agendas," in Donald T. Hawkins [Ed.], Personal Archiving: Preserving our Digital Heritage, Medford, New Jersey: Information Today, 2013), p. 259-277. https://perma.cc/R8X4-ES47

--William Odom, Abi Sellen, Richard Harper, and Eno Thereska, "Lost in Translation: Understanding the Possession of Digital Things in the Cloud," *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (ACM Press, 2012), 781-790. doi:10.1145/2207676.2207789.

Cloud Record Capture Tools

Social Feed Manager

Helping researchers and archivists build social media collections

View on GitHub

Correspondents (114)

Get TAGS

Locations (345)

Image attachments (4)

Document attachments (92)

Int'l Standard Framework

- **ISAD (G)**
- **ISAAR-CPF**
- **ISDF**
- **ISDIAH**

Provides Separation of Concerns at **Conceptual** Level

- Rules for Archival Description (RAD): Canada
- Describing Archives: A Content Standard (DACS): USA
- Encoded Archival Description
- EAC-CPF

InterPARES Trust

Records Management: Multi-entity models

--ISO 23081-1:2006, Information and documentation -- Records management processes -- Metadata for records -- Part 1: Principles
--ISO/TR 26122:2008 - Information and Documentation -- Work Process Analysis for Records.

InterPARES Trust

Alternative strategies

Robert B. Allen, "Using Information Visualization to Support Access to Archival Records." *Journal of Archival Organization* 3, no. 1 (October 12, 2005): 37–49. doi:10.1300/J201v03n01_04;

Victoria L. Lemieux, "Toward a 'Third Order' Archival Interface: Research Notes on Some Theoretical and Practical Implications of Visual Explorations in the Canadian Context of Financial Electronic Records," Archivaria 78 (Fall 2014), http://journals.sfu.ca/archivar/index.php/archivaria/article/view/13493;

Greg Bak,"Continuous Classification: Capturing Dynamic Relationships among Information Resources." Archival Science 12, no. 3 (September 2012): 287–318. doi:10.1007/s10502-012-9171-8

Jefferson Bailey, "Disrespect Des Fonds: Rethinking Arrangement and Description in Born-Digital Archives," Archive Journal: 3 (2013). https://perma.cc/P4SX-F3HR.

Sarah Higgins, Christopher Hilton and Lyn Dafis, "Archives context and discovery: rethinking arrangement and description for the digital age", Girona 2014: Arxius i Industries Culturals http://www.girona.cat/web/ica2014/ponents/textos/id174.pdf.
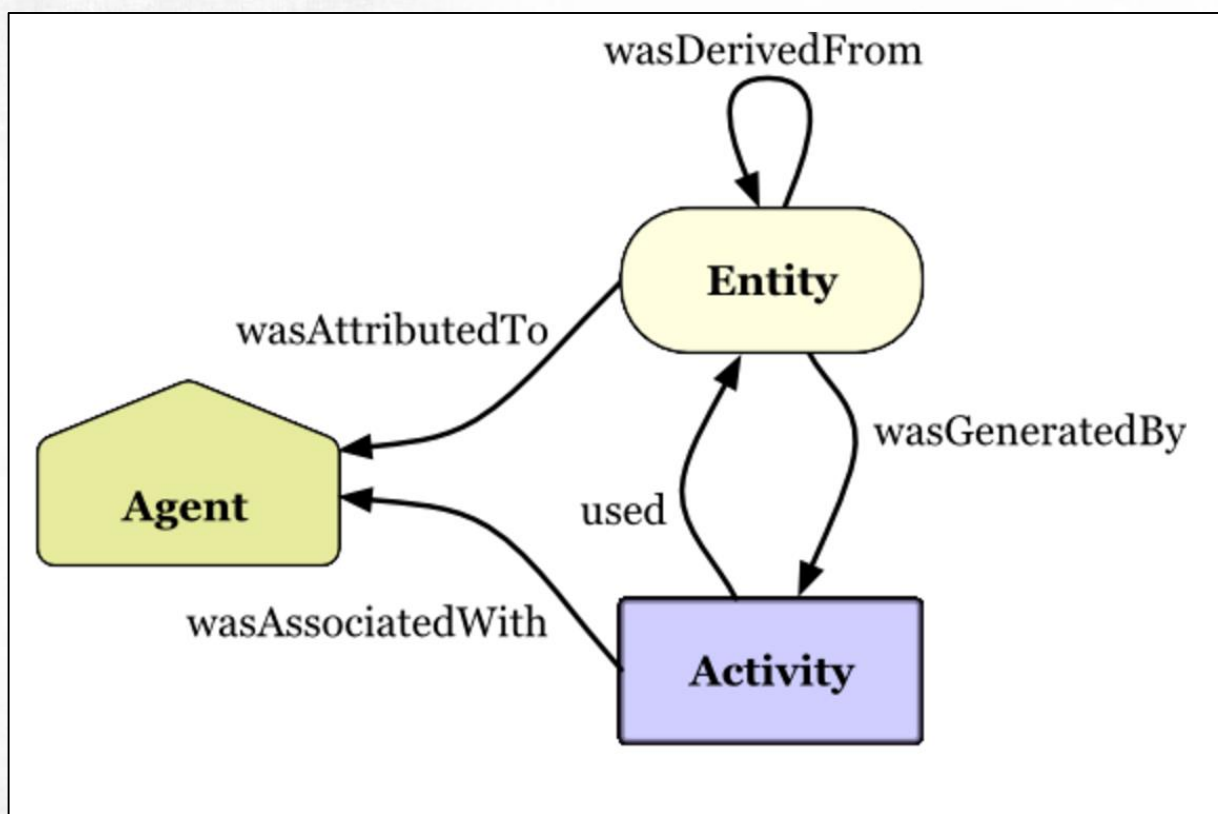
Linked Open Data

-- Tufts University, LiAM: Linked Archival Metadata Guidebook, http://sites.tufts.edu/liam/

-- William Wueppelmann, "Out of the Trenches : A Linked Open Data Project," (2012). http://www.canadiana.ca/en/pcdhn-lod

-- Claire Sibille de Grimoüard and Alice Motte,"Archives and Linked Data: Are Our Tools Ready to 'complete the Picture'? Two Case Studies." In *Girona 2014: Archives and Cultural Heritage Industries.* http://www.girona.cat/web/ica2014/ponents/textos/id9.pdf;

--"Linking Lives | Using Linked Data to Create Biographical Resources." (2012) http://linkinglives.archiveshub.ac.uk/ .

--ALOD, Archival Linked Open Data, http://alod.ch/

--Karen F. Gracy, "Archival Description and Linked Data: A Preliminary Study of Opportunities and Implementation Challenges," *Archival Science* 15, no. 3 (September 2015): 239–94. doi:10.1007/s10502-014-9216-2

WC3 Prov-O Ontology

"Provenance is information about entities, activities, and people involved in producing a piece of data or thing, which can be used to form assessments about its quality, reliability or trustworthiness"

InterPARES Trust

Developing Models

- EAD3:
  - experimental <relations> element
  - Available at any component level
  - https://www.loc.gov/ead/EAD3taglib
- Experts Group on Archival Description
  - http://www.ica.org/en/about-egad

Final Citations

--Geoffrey Yeo, "Bringing Things Together: Aggregate Records in a Digital Age," Archivaria; Archivaria 74 (Fall 2012), 2012. http://journals.sfu.ca/archivar/index.php/archivaria/article/view/13407.

--Elings, Mary L. "Using NLP to Support Dynamic Arrangement, Description, and Discovery of Born Digital Collections: The ArchExtract Experiment." bloggERS!, May 24, 2016. https://perma.cc/R5YK-Z7ZF.

InterPARES Trust

Archival Arrangement and Description in the Cloud
A Preliminary Analysis

Christopher J. Prom
Born Digitals in the Cloud: Challenges and Solutions
Archival Science Colloquium, Archivschule Marburg
8th June 2016