

Afterwords  
Fiesole 19  
The Evolving Scholarly Environment

21 April 2017

Lilliad at Université de Lille Science et Technologies

M. A. Keller of Stanford

# Plus ça change, plus c'est la même chose

- Fiesole conferences **always** informative & useful; updates & new elements
  - Revealing changing professional practices among librarians & publishers
  - Reminding us that there is a widening spectrum of types – of libraries, librarians, publishers, & interested, if not innocent, bystanders; there is no average...or is there?
  - ~20 years on, so much has changed, but so much of that change is derivative
  - Constants: network effects; data (bigger, more varied); constantly improving ease of analysis, use, presentation; hand wringing; dis-intermediation; dependence on libraries & publishers as service providers, cultural agencies, handmaidens to researchers & teachers; marketplaces;
  - Variables: modes of interaction & access; senses of potential (BIG DATA); increasing engagement across all disciplines with digital possibilities – research & teaching – thus new demands & requirements on librarians and publishers
  - Librarians lack visible concern for overhead costs (time, money) & sustainability; most publishers vitally concerned with money (costs & profit) & efficiency (matching fundamentals to workflow efficiency) & business models (licenses, OA, APC, purchases)

# Unica

- Data.bnf.fr
- Linked data in DNb
- SHARE-Virtual Discovery Environment
- Text & Data Analysis
- AI & Machine Learning – Yewno
- New forms of peer reviews publications from university presses

# Missing elements

- Return on investments
  - For profit publishers constantly calculating
  - For university presses (not for profit, but formal) concern for sustainability (dependence on subsidies & in some cases endowment earnings)
  - For OA operations – is there a sense of total investment & ROI?
  - ROI could be financial, could be reputational, could be “contributions to commonwealth of knowledge”
    - First books; “important books with limited audiences”; supported projects (long term, multi-vols)
  - Where bibliometrics -- New Benchmarks and Methods of Assessment for the 21<sup>st</sup> Century Research Library, project underway
- Digital Humanities & Computational Social Sciences – development, support, publication; consideration of the different contexts for humanities & social sciences vs those of STEM disciplines
- GIS + other geospatial collections & services supporting nearly all disciplines

# Methods

- Top down or bottom up at various levels – national & organizational
  - BUT why not both?
- Network effects – note findings of Flecker & Dempsey analysis of collection overlap among 1<sup>st</sup> five Google Book Project libraries – finding 51% unique holdings across logical collection (NB lack of bureaucratic overhead); multi-nodal, informal results
- Collection development in larger research libraries – beyond formal publications: government & NGO documents; media objects; tech reports, grey literature & ephemera; Websites (why not undertaken by libraries of other categories)
- Usual services, but what of teaching/modeling “information heuristic”, a lifelong skill for students becoming citizens & career professionals
- Data deluges, data refuges, discerning trends, thematic &/or conceptual analysis (LitLab example)
- Pointilistic studies (e.g. Treharne analysis of Salisbury copy of Magna Carta) still and always essential & meaningful
- Preservation – LOCKSS model for copies & preservation applications – avoiding single points of failure (so worries about Internet Archive as single point of failure)

# Analytics

- Circulation statistics could lead to epidemiological results – damage or correct focus? Single language foci or research resources in numerous languages from numerous sources in support of local research programs (e.g. area studies, topical studies)
- On average in 5-10 years, all titles acquired or licenses, circulated internally or externally are circulated
- After 10 years of operations, Stanford Digital Repository (digital library & archive platform based on Fedora, Hydra) supporting 33% of Stanford professors
- Helpful taxonomy of interested/committed parties in scholarly communication
- Ethnographic approach gaining insight on initial career step for new PhDs
- Stanford holds ~1.3M e-books; 2.2M “uses” annually; unknown for 65K e-journals
- In 2015 Stanford authors published 13,000 articles. If APCs average \$1650 per article, required investment to instantiate OA would be \$21,450,000, nearly all of our consolidated library materials allocation

# So...what next?

- Retail services + mass offerings continue
- Innovations in support of DH & CSS via focused support, experienced & communicative specialists, community developments (IIF & Mirador & similar)
- Image recognition and analytics
- Continuing work on ROI algorithmic analysis as a distributable model but will librarians & their masters concerns themselves with ROI empirically derived?
- Overwrought overhead in collaborative approaches among libraries by category & national associations; define projects & methods to 95% level, over determination, little flexibility, *en masse* projects with little or no assessments of ROI
- Too little focus on “just do it” & risk taking