

# Technically Assisted Review: Fact, Fiction or Jam Tomorrow?

Tim Gollins

Head of Digital Archiving  
The National Records of Scotland

# Technically Assisted Review (TAR)

- TAR, Is it real?
- TAR, What is it ?
- TAR, What is it good for?
- TAR, A real example
- Other published work
- My View
- Why does this matter? - Panel

# TAR, is it real?

- Yes !
- In e-Discovery it works
  - For reviewing for responsiveness to the “Matter”
  - For reviewing for “Legal Privilege”
- Significant academic research
- Significant commercial investment
- Economic driver - Money to be made!

# TAR, What is it?

## Breaking it down

- Technical
  - Tools and Techniques developed by Information Retrieval (IR) community (largely Machine Learning)
  - System learns from examples what “good” looks like (Lots of different approaches)
- Assisted
  - HELPS humans be more efficient - does not replace.
- Review
  - conceptually “looks” at everything.

# TAR, what is it good for?

- Not Absolutely Nothing!
  - Say it again! Huh !
- BUT the detailed task matters
- e-discovery is NOT
  - e-Archiving selection
  - e-Archiving appraisal
  - e-Archiving sensitivity review!
- Not all dogs can herd sheep!

# TAR

## A real example of use with Email

- TAR used to assist in the release of Email of Governor Tim Kaine (Library of Virginia)
- Total Collection 1.3 Million emails
  - Being (or has been?) released in batches
- Reviewing for
  - Privacy
  - Legal Privilege
- See Grossman and Cormack – University of Waterloo for all complex details

# Other Published Work not e-mail specific

- The National Archives Research Report (February 2016)
  - Largely positive, qualitative, results of applying e-discovery tools on sample records
- Papers on “Sensitivity Review” by Graham MacDonald (Glasgow University)
  - Largely positive, quantitative, results of fundamental algorithmic research

# My View

- Fact – YES
- Fiction – YES
  - Significant danger of oversimplification of tasks
  - Significant danger of hype
- Jam Tomorrow – YES
  - No generic solution (outside e-discovery)
  - Small amount of core research being done
  - Specifics of “e-mail in archives” needs much work
  - Little or no work on “Archival” User Interfaces!

# Why does this matter for e-mail Preservation?

Panel?