DESI VII Keynote – June 12, 2017

<u>Selective Digital Amnesia</u>

Maura R. Grossman & Gordon V. Cormack
David R. Cheriton School of Computer Science
University of Waterloo

When private information is comingled with information to which others are entitled, privacy rights and access rights collide. Both can be fulfilled only to the extent that private and non-private information can be accurately segregated.

Imagine a hypothetical scenario in which unlimited time and resources (both human and technological) were available. Would it be possible to separate the information perfectly, such that neither party's rights were impinged, or is perfection an illusion?

In practice, the goal is to do the best possible job of separating the information, using reasonable time and resources, where what is considered reasonable is informed by what is possible. Arguably, if a new method can do a better job using fewer resources than the accepted practice, it should be considered reasonable. And once the new method is established, using the previously accepted practice may cease to be reasonable.

We illustrate these points by comparing technology-assisted review (TAR) to an exhaustive manual review conducted by the Virginia Senior State Archivist to identify restricted records among 400,000 email messages from the gubernatorial administration of Tim Kaine.