Creating and Accessing Email Archives

Anton Leuski and Douglas W. Oard

Email management
- Implicit context
- Task: Defining structure
  - hierarchies
- Task: Using structure
  - arrival (classification)
  - workflow management
  - retrospective

Email archives
- Context reconstruction
- Task: Redaction
  - Isolate sensitive content, links, patterns
- Users:
  - Archivists

Email collections:
1. NSC email collection from “White House E-Mail”, edited by Tom Blanton
2. Ben Shneiderman’s email from UMD library

Facts (in 2000, in the world):
- 600 billion messages
- 570 million e-mailboxes

Task: Access
- Users:
  - Historians
  - Social scientists
  - Teachers and students
  - What about you?...

Technologies
- Segmentation
- Search
text, links, and patterns
- Topic and link detection
- Topic tracking
- Traffic analysis
- Social network analysis
- Data mining
- Visualization
- Summarization
- Privacy

Technology example: Visualization and browsing
eArchivarius