

## **Grammatical Induction and Recognition of the Documentary Form of Records**

William Underwood, Sheila Isbell, Matthew Underwood  
Georgia Tech Research Institute  
250 14<sup>th</sup> St. N.W. Atlanta, Georgia 30332-0832  
[william.underwood@gtri.gatech.edu](mailto:william.underwood@gtri.gatech.edu)  
[sheila.isbell@gtri.gatech.edu](mailto:sheila.isbell@gtri.gatech.edu)  
[matthew.underwood@gtri.gatech.edu](mailto:matthew.underwood@gtri.gatech.edu)

### **Abstract**

This paper presents digital curators with a more precise understanding of the concept of documentary form, and how documentary form can be automatically learned from a sample of records of a particular document type. The ability to automatically recognize documentary form enables item description. Item description enables file unit description and this enables automatic series description. This technology can reduce the effort required of an appraisal archivist to assess the value of record series containing a large number of e-records of different documentary forms. It can also provide archivists with earlier intellectual control of accessioned e-record series by providing preliminary scope and content notes for these series. Item descriptions provide additional ways for indexing and searching collections of records.