C51   A Validated, Admissible, Computational Method for Detecting Electronic Authorship

Carole E. Chaski, PhD*, Institute for Linguistic Evidence, Inc., 25100 Trinity Drive, Georgetown, DE 19947

After attending this presentation, attendees will become acquainted with a validated, admissible method for determining authorship of electronic documents including email, blog posts, and chat-room messages.

This presentation will impact the forensic community and/or humanity by demonstrating how the ALIAS method is useful in criminal and civil investigations including cyber crimes, homicide, and terror threats. The ALIAS method allows investigators to obtain essential evidence and, because it is reliable and validated, testimony based on the ALIAS method has been admitted without any restrictions after a *Daubert* hearing and also under the *Frye* standard.

In an increasingly electronic society, the authorship of electronic texts can be key in the investigation of many different types if crimes and civil infractions. A suicide note left on a home computer attempts to cover a homicide. Anonymous emails over the corporate network attempt to ruin a supervisor's reputation. Anonymous "tell-all" letters attempt to scotch a corporate merger. Pedophiles leave an electronic trace in their seduction of children on the web. All of these scenarios and many more demonstrate the importance of a validated, admissible method for determining the authorship of electronic texts. Further, electronic documents significantly differ from handwritten documents, with differences that can disable methods based on handwriting examination (such as forensic stylistics). ALIAS (Automated Linguistic Identification and Assessment System) is computational-linguistics software which analyzes the syntactic patterns, syntactically-classified punctuation and word lengths of texts from which numerical output is statistically analyzed using a leave-one-out, cross-validated discriminant function analysis (Chaski 1997, 2001, 2005). ALIAS has been validated by experiments independent of litigation, at the document level. Using sufficient documents to obtain approximately 2000 words/100 sentences from each of ten authors, ALIAS obtained 95% accuracy; that is, 95% of the documents were consistently classified to the correct author in the ten-author set (Chaski 2005). More recently, ALIAS has been validated independent of litigation, at the sentence level. Using exactly 100 sentences from each of ten authors, and comparing 100 sentences of one author to 100 sentences from another author, ALIAS obtained 85% accuracy; that is, on average over the 10 authors, 85% of each authors' sentences were consistently classified as belonging to the correct author. These results demonstrate the utility of the method even in chat-room scenarios.

The ALIAS method enables investigators to obtain essential evidence against cyber-criminals and other criminals who make use of electronic texts. The method has been used to identify the authors of blog posts, the author of threatening letters and the author of a phony suicide note. The ALIAS method has also been admitted into trial testimony under the Frye standard in Maryland (1998) and also after a *Daubert* hearing in the Federal Court of the District of Columbia (2001).

Computer Forensics, Authorship Identification, Questioned Documents