Cedric Neumann

received his PhD in Forensic Science from the University of Lausanne, Switzerland, host of the oldest and most prestigious Forensic Institute in the World (est. 1909). In the 2000s, Cedric led the Statistic and Interpretation Research Group of the Forensic Science Service (Home Office, UK). During that time, Cedric and his team developed various statistical models for the interpretation of partial and mixed DNA profiles, latent prints, and shoe impressions. Cedric also led a team of contractors tasked with building and deploying the Digital Ink Library for the United States Secret Service (U.S.S.S., U.S.A).

During the 2010s, Cedric was an assistant/associate Professor of Statistics at the Pennsylvania State University and at the South Dakota State University, where he led a research team on foundational issues related to the use of statistics in forensic science. At that time, Cedric contributed to U.S. government-sponsored groups and committees, such as the Scientific Working Group on Friction Ridge Analysis, Study and Technology (SWGFAST) and the Organisation of Scientific Area Committees for Forensic Science (OSAC). In 2021, Cedric joined the Battelle Memorial Institute as a lead data scientist (Data Scientist IV).

Cedric uses Bayesian statistics, machine learning (ML) models and large language model (LLM)-based artificial intelligence (AI) to support projects related to military health, development of advanced materials for the U.S. Department of Defense (DoD) and the U.S. Department of Energy (DoE), and to other areas of national interest. As a result of his work, Cedric was recognized with the 2009 European Emerging Forensic Scientist Award, a Certificate of Appreciation from the U.S. Department of Homeland Security in 2009, the 2016 Sherwood and Elizabeth Berg Young Faculty Award and the CAPITAL Card Services Scholarship in Machine Learning and Artificial Intelligence.

Cedric Neumann, PhD

Advances in Statistical Interpretation of Non-DNA Forensic Evidence

The increasing use of forensic DNA evidence in the 1990s and 2000s has brought to light the possibility to quantify the weight of forensic evidence, and to report probabilistic or statistical values as results of forensic analyses. It is often forgotten that, in the early XX century, the founding father of criminalistics were already concerned with providing statistical support for forensic evidence, starting with anthropometry and fingerprint evidence. During the 1960s and 1970s, early statistical models were introduced for trace evidence. These models defined how forensic DNA evidence came to be evaluated and reported. Finally, the past 25 years have seen countless models developed and proposed to quantify the weight of trace, fingerprint, toolmarks, footwear, document and other types of evidence. Yet, in most jurisdictions, inferences from non-DNA evidence are still based on expert judgment and made with very little statistical support. The purpose of this talk is not to claim that these inferences are incorrect; rather, we will explore why non-DNA forensic evidence types have yet to widely adopt statistical approaches to support conclusions.

Despite progresses, developing statistical models for non-DNA evidence presents several unique challenges. Unlike DNA, which has a well-defined structure and sequence, non-DNA evidence can be hard to describe mathematically. Non-DNA forensic evidence can be complex and multidimensional, involving multiple variables and intricate relationships. Capturing this complexity in statistical models requires sophisticated techniques and large datasets. However, obtaining sufficient data for model training and validation can be challenging, especially for rare or unique types of evidence. Non-DNA evidence has also a much wider range of variability. For example, fingerprints can be smudged, distorted, be patent or latent, positive or negative, and only show a few ridges or more than 100 features, including sweat pores and scars. This inherent variability makes it difficult to develop models that can account for all possible variations and ensure accurate interpretations.

Finally, forensic evidence should be presented in a clear and understandable manner. Complex statistical models may be difficult for decision-makers to comprehend, potentially affecting the weight and credibility of the evidence. Balancing the need for advanced statistical methods with the requirement for clarity and simplicity in legal contexts is an ongoing challenge (even for DNA evidence).

This talk will briefly review some historical models and the arguments in favor of the need for more quantifiable support when reporting forensic evidence. We will also discuss the issues related with developing statistical models for non-DNA evidence. We will survey some of the recent developments in different fields, such as trace evidence and fingerprints. And finally, we will explore how statistical information can be used to support expert judgements, without necessarily be used to replace them and directly report forensic evidence.

There will be very little to no mathematics in this talk. And you are welcome to bring your questions and comments.