Connecting Illegal Vendors on Darknet Markets: Responsible Authorship Attribution to Link and Connect Online Cybercrimes

About the Presenters:

I completed my bachelor's degree at a European law school and pursued a master's degree in criminology, forensics, and law at Maastricht University. Currently, I am pursuing a Ph.D. at the University of Lausanne. My research focuses on artificial intelligence in the judiciary. Specifically, I am examining various aspects such as risk assessment tools in the justice system, the concept of legal relevance in information retrieval systems, factors influencing the adoption of technology in the judiciary, and methods for testing frameworks within the use of technology in the justice system.

I'm a machine learning enthusiast, in the final year of my Ph.D. at Maastricht University's Law & Tech lab, Netherlands, focusing on research in Natural Language Processing and Computer Vision. My work concentrates on exploring multimodal representational learning techniques for authorship identification tasks, to link and connect illegal vendors by analyzing patterns in writing and photometric styles on online market advertisements.
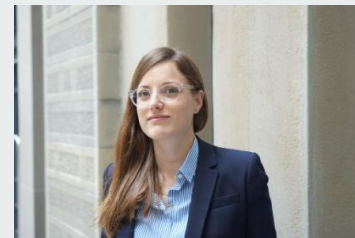
Dr. Aurelia Tamò-Larrieux is an Assistant Professor for Privacy, Security, and Computational Law at University of Lausanne. She has a background in law and economics and is specialized in research at the intersection of law and digital technologies with a particular focus on privacy, data protection, design approaches for privacy-friendly technologies, transparency of automated decision-making, automatically processable regulation, and trust in automation. Aurelia's scientific publications on those subjects are open access and she has presented her research at numerous international conferences.
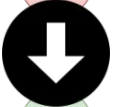
Introduction to Authorship Attribution and Darknet Markets

Responsible guidelines for Authorship Attribution
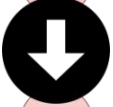
Hands-on Session 1: Getting started, data analysis and pre-processing, and stylometric analysis

Hands-on session 2: Authorship Identification through Statistical algorithms and traditional Neural Networks

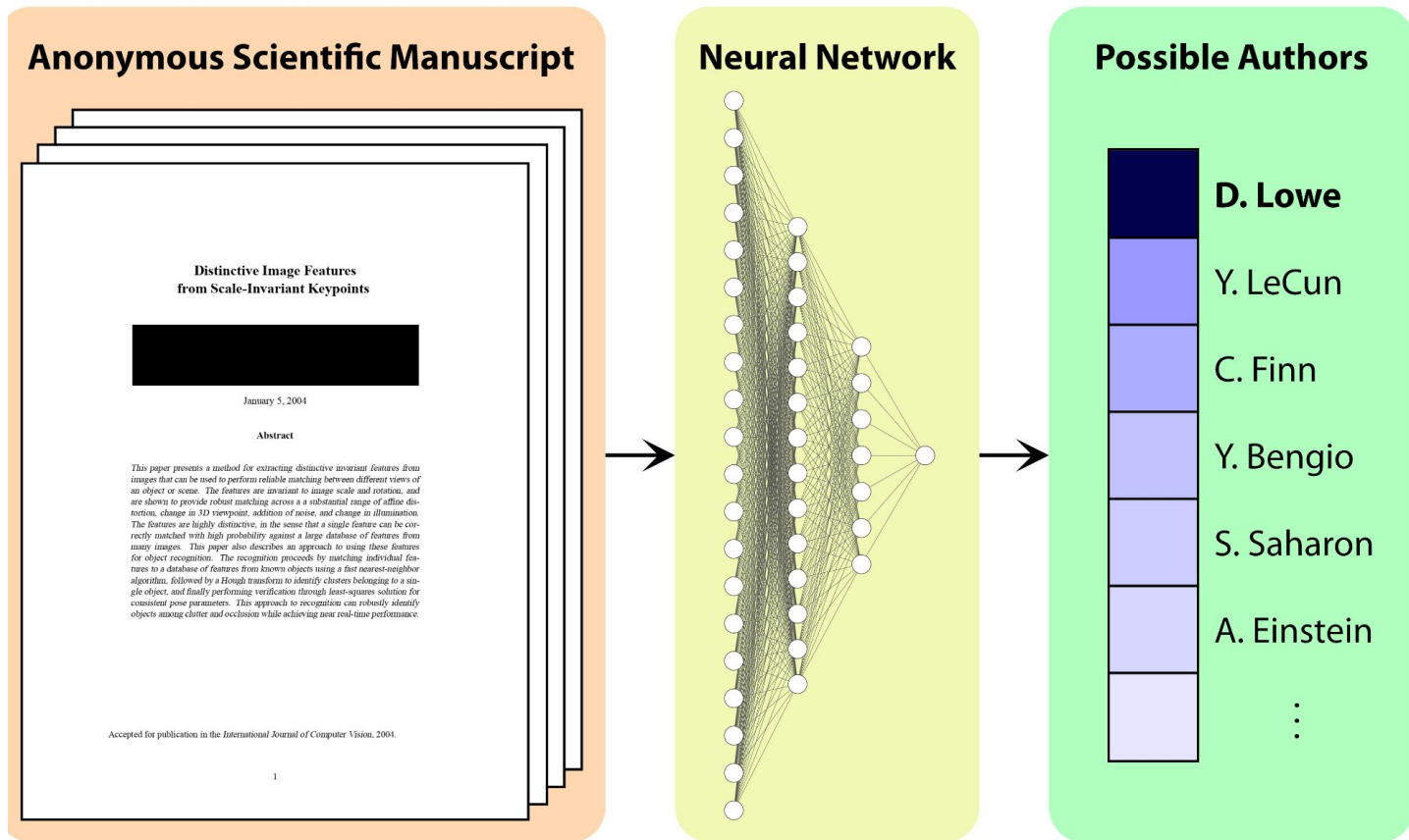**Hands-on session 3: Authorship Identification through Transformers-based models**
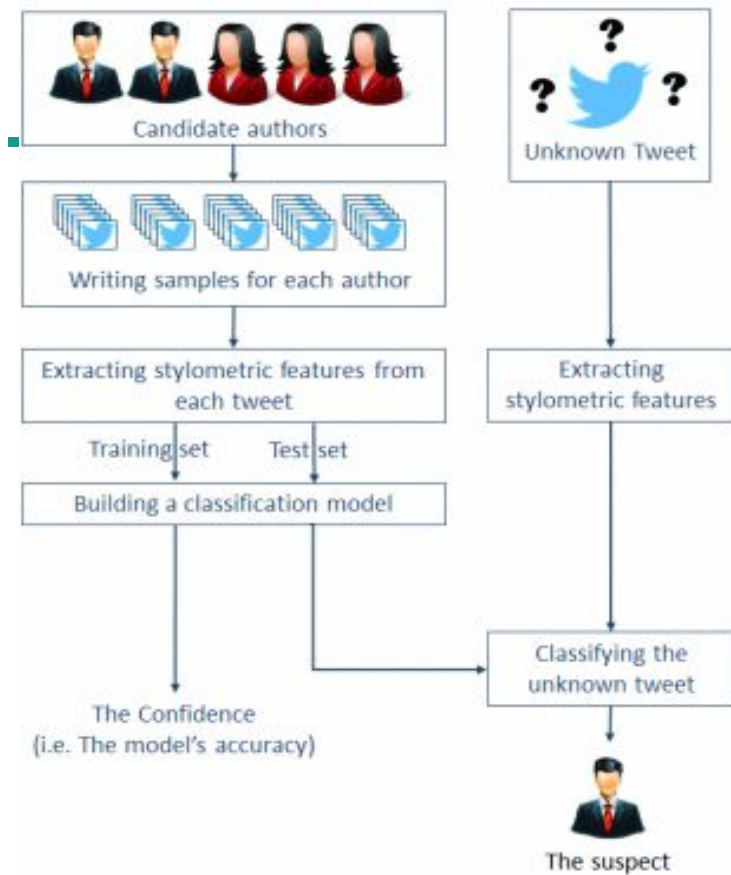
Hands-on session 4: Authorship Verification as retrieval task

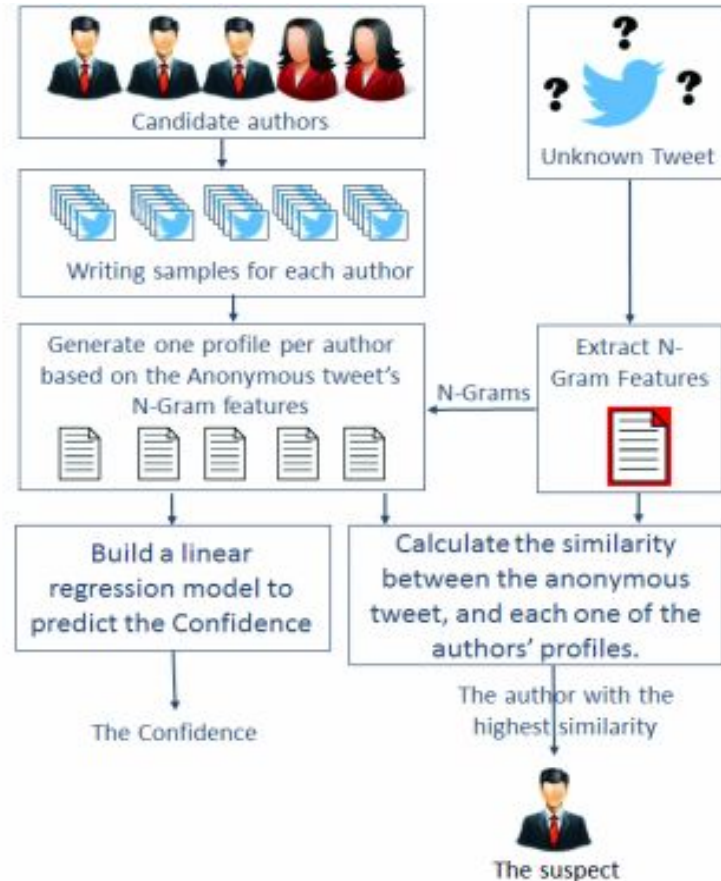**Limitations, Future Work, and Conclusion**

What is Authorship Attribution (AA)? How does it apply to NLP?

Authorship Identification vs Authorship Verification

(a) Instance-based authorship attribution.

(b) Profile-based authorship attribution.

Dark Web and Darknet Markets

Dark Web Content

A study by *Gareth Owen of Portsmouth University* discovered that content on the Dark Web was dominated by:

Illegal pornography, black markets, hacking groups and botnet operations (those commonly associated with spam, fraud and malicious attacks).

Hidden Services on the Dark Web (January 2015)

- Drugs
- Fraud / Counterfeit
- Black Market Sites
- Bitcoin
- Whistleblowing
- Hacking
- Pornography
- Blogs
- Abuse
- Other / Not Indexed

SHUT IT DOWN!
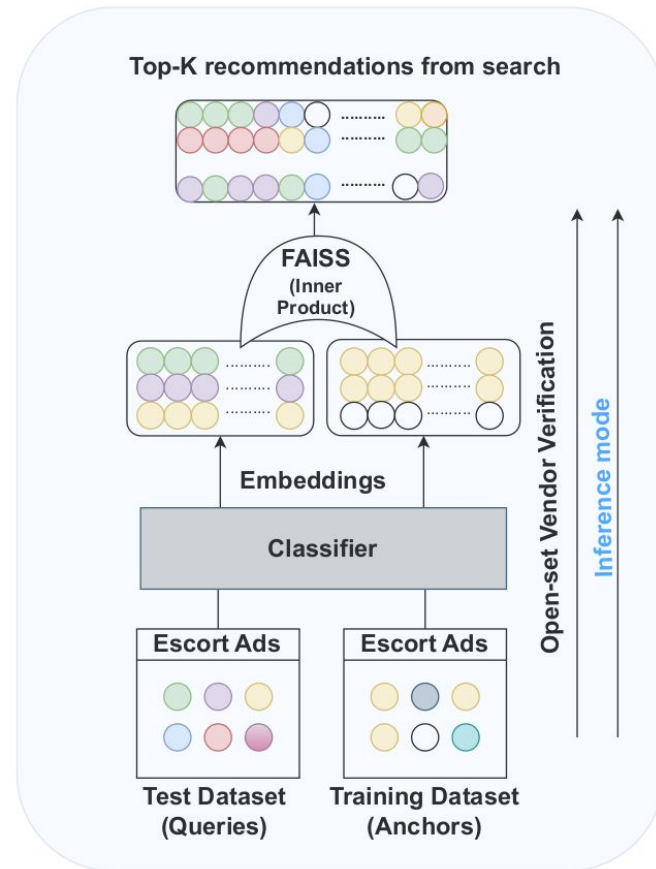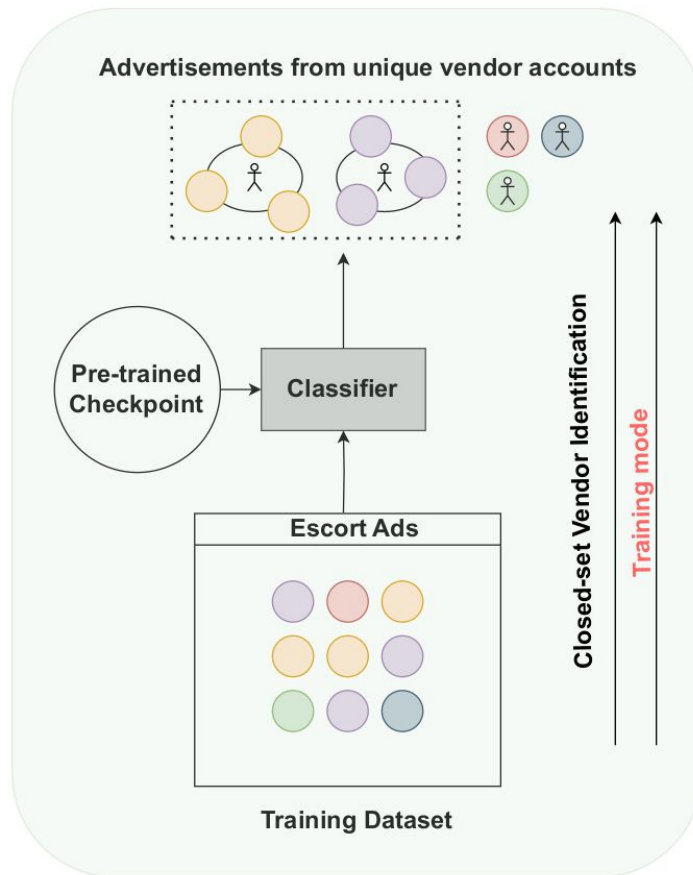

WHOA CHILL OUT BRO

SETBACK

- Anonymity on the Dark Web
- Limited human resources
- Too much data to analyze
- Vendors creating aliases and appearing across different markets
  - Makes it difficult to estimate the scope and size of a market

Assistance

Investigation

Data
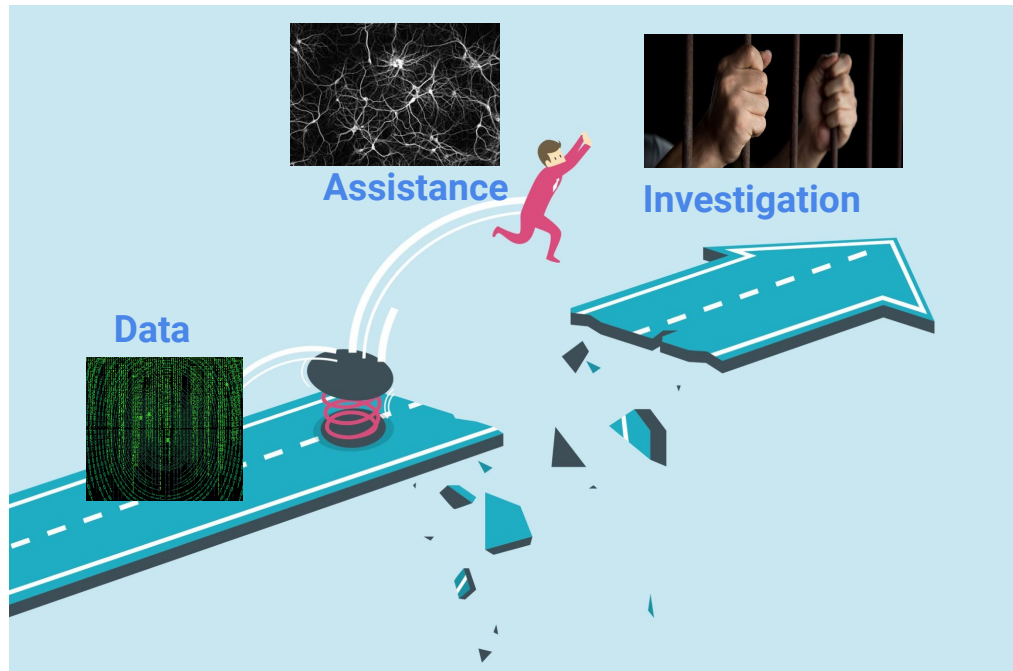
Vendor Migrants vs Vendor Aliases

VendorLink: An NLP approach for Identifying & Linking Vendor Migrants & Potential Aliases on Darknet Markets

IDTraffickers: An Authorship Attribution Dataset to link and connect Potential Human-Trafficking Operations on Text Escort Advertisements