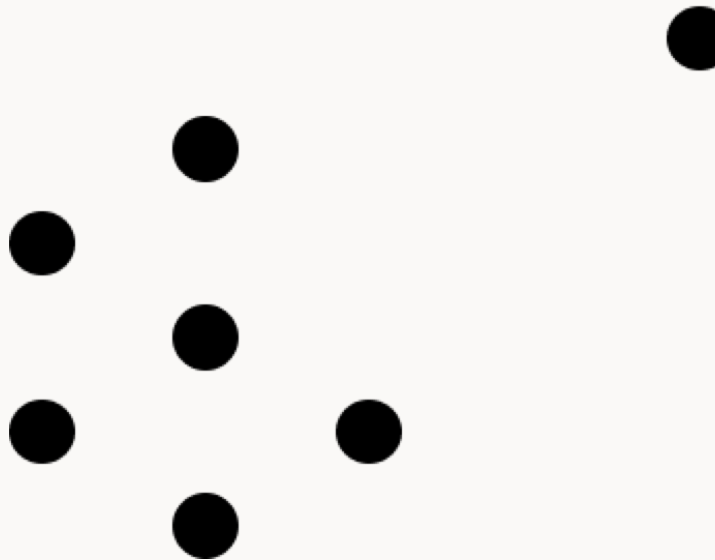


::: Primer

Primer AI: Deep Learning Text Forensics

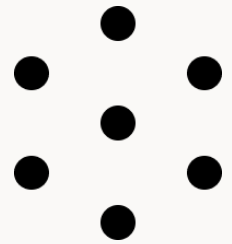
IARPA HIATUS Program Lightning Talk

01/14/2022





Agenda



- _ Introduction to Primer
- _ Primer's Experience with Authorship Detection
- _ Human Authorship/Privacy
 - Stylometry
 - Steganography



INTRODUCTION TO PRIMER

Primer: Leader in Machine Learning for Defense and Intelligence

Primer is headquartered in San Francisco, California and focuses exclusively on the world's best natural language processing and generation solutions

Our technology is used by some of the world's largest government agencies, financial institutions, and Fortune 50 companies

We are dedicated to delivering game changing machine learning software solutions to the world's most demanding customers

- 60+ machine learning engineers focused on NLP
- 120+ total engineers
- 20% of company with IC/DOD experience



2015

FOUNDED



215

EMPLOYEES



50+

CUSTOMERS



6

OFFICE LOCATIONS

SELECT CUSTOMERS & PARTNERS



Walmart*



IQT
IN·Q·TEL



Primer's Experience with Authorship Detection

Primer has built a next-generation NLP platform to detect and analyze influence operations

01 Dispute Detection and Summarization

02 Claims Extraction

03 Claims Partitioning and Attribution

04 Bot Detection

01

Claims Discovered in Dispute

- China has denied allegations of abuse and genocide, blaming "anti-China forces" for "some distorted coverage" in Western media.
Australia urged to act after human rights report reveals 'horrifying' case of Uighur jailed in Xinjiang
Original Source: Special Broadcasting Service | Apr 19, 2021
- China has denied the United Nations unfettered access to the region to investigate.
Group urges UN to probe China against humanity
Original Source: The Independent | Apr 19, 2021

02

Claims Discovered in Dispute

- On Wednesday, China described the US assertion as "the biggest lie of all that rides roughshod over international law" and suggested the report was based on testimonies from "trained" actors and actresses.
US, China fight over claims of 'genocide' in Xinjiang
Original Source: Daily Nation (Kenya) | Apr 02, 2021

Side A
United States
Side B
China

03

04

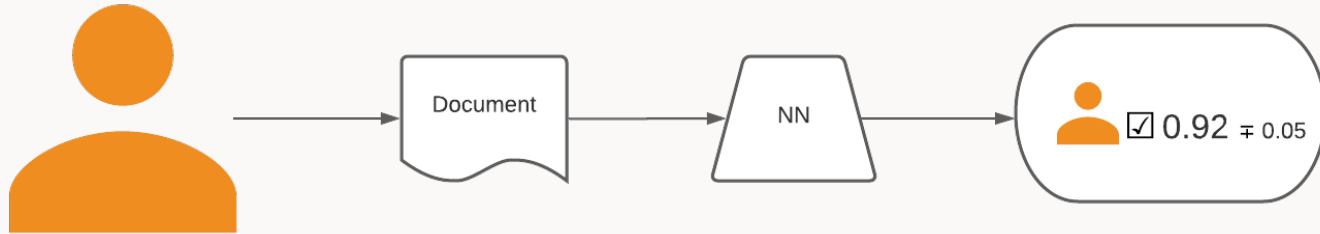
- // Synthetic Text Detection
- // Native Language NLP for Russian and Chinese
- // Event Detection and Summarization
- // Information Content Tracing

- // Dispute Identification and Clustering
- // Social Network Analysis
- // Claims Identification, Clustering, and Summarization
- // Topic Modeling

Primer's GPT-2/GPT-Neo classification models detect synthetic text with high performance

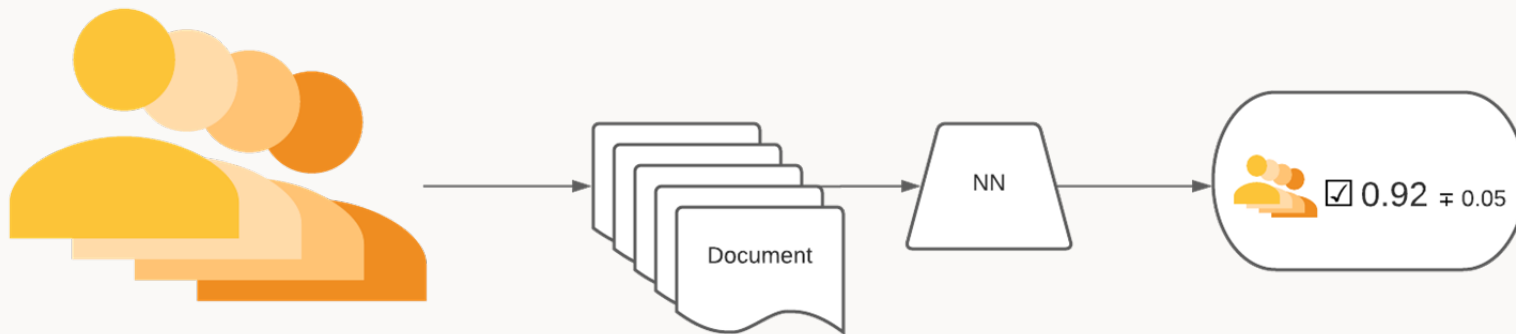
Updating stylometry with deep learning

Problem 1: Individual authorship



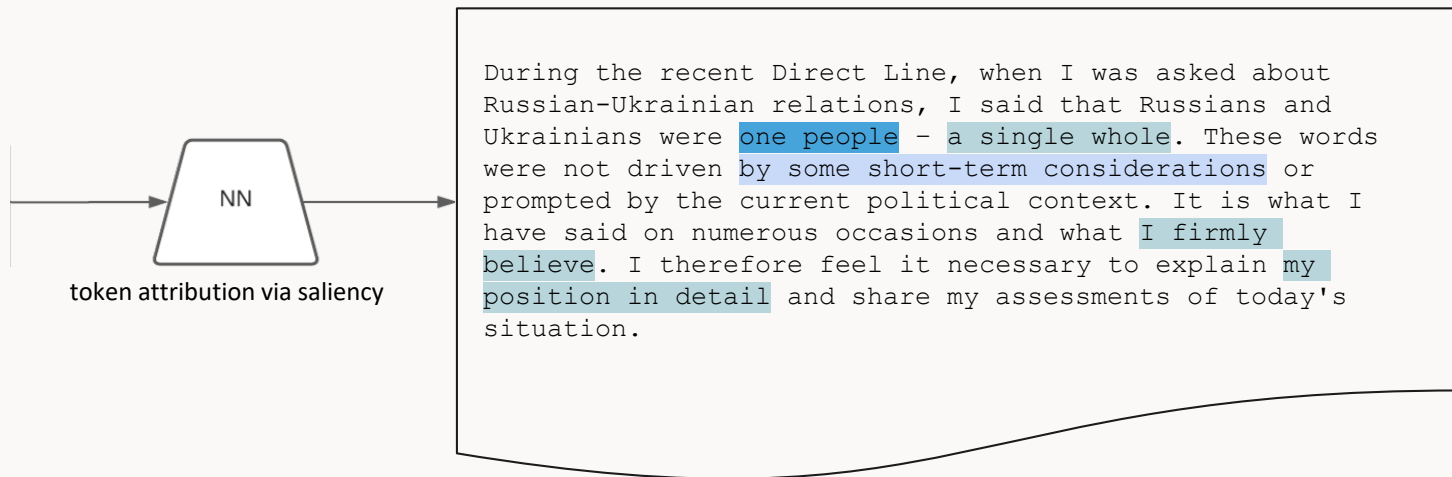
Updating stylometry with deep learning

Problem 2: Group authorship



Updating stylometry with deep learning

Problem 3: Explainability



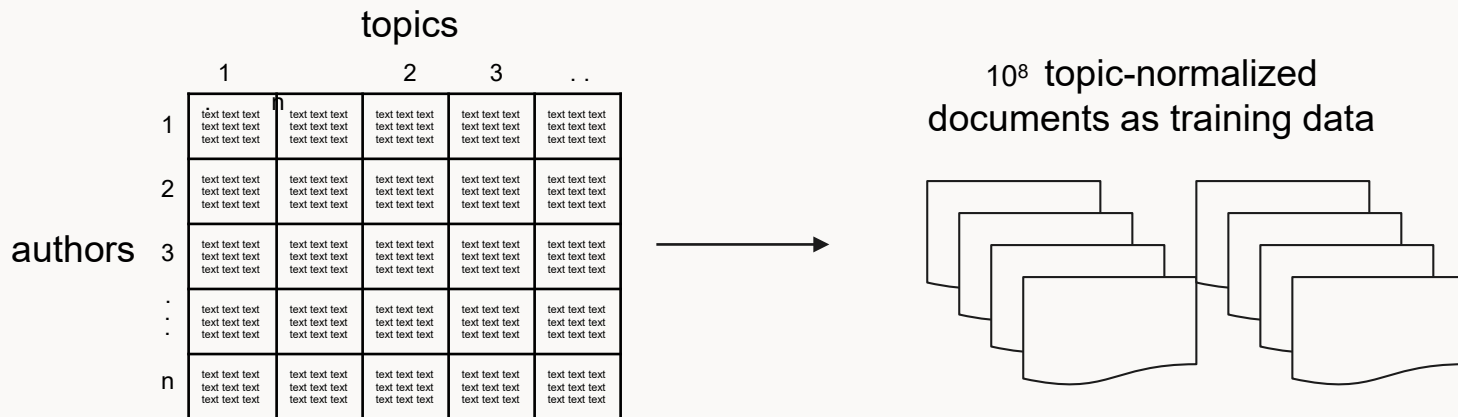
Updating stylometry with deep learning

Problem 4: Discovery



Updating stylometry with deep learning

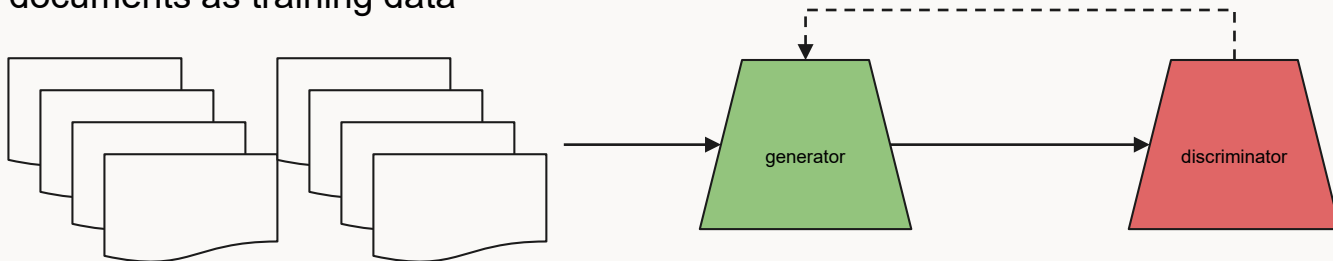
The real problem: Training a language model to focus only on style!



Updating stylometry with deep learning

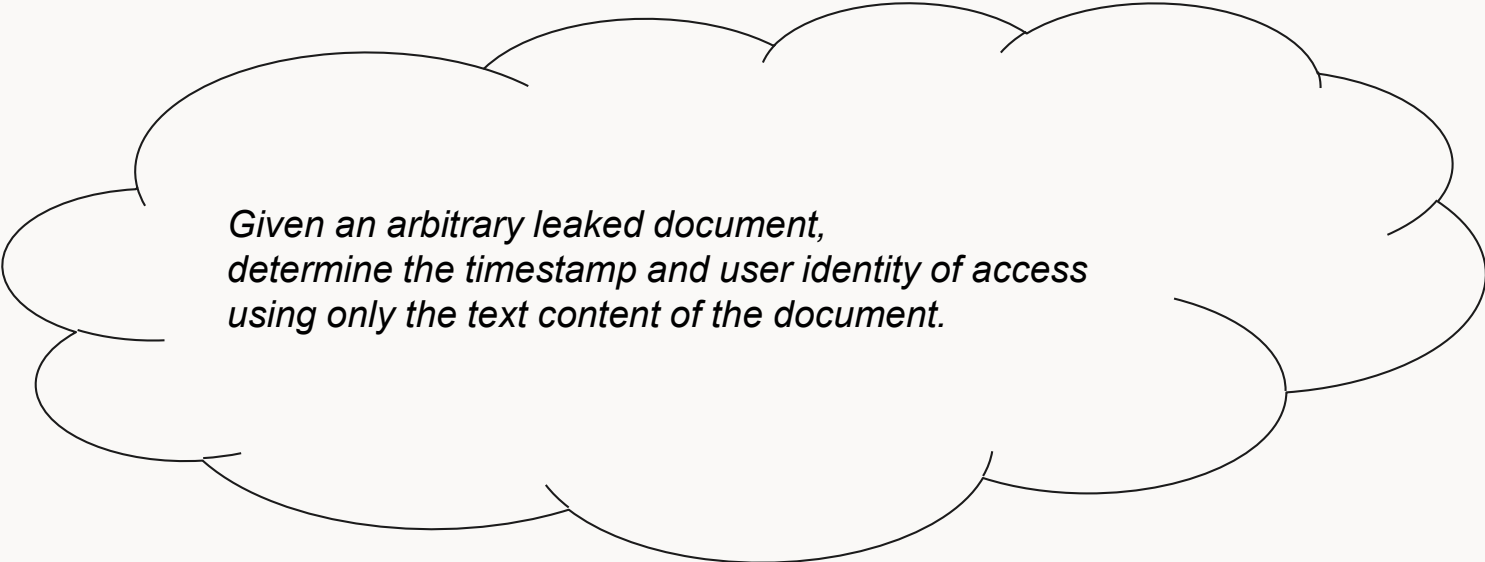
The real problem #2: Style generation

10^8 topic-normalized
documents as training data



Updating stylometry with deep learning

The problem you didn't realize you had: Provenance control



*Given an arbitrary leaked document,
determine the timestamp and user identity of access
using only the text content of the document.*

⋮ Primer

Thank You

John Bohannon

Bohannon@primer.ai

WEB: <https://primer.ai>

BLOG: <https://blog.primer.ai/>

TWITTER: [@primer_ai](https://twitter.com/primer_ai)

