# Linguistic Anomaly Detection at Scale:
## Unveiling the Elsagate Phenomenon and Ensuring Online Child Safety

Panagiotis Soustas | Dr. Matthew Edwards | Dr. Claudia Peersman
Panagiotis.Soustas@bristol.ac.uk

## Introduction

The ability to identify abnormal language patterns at scale is crucial for ensuring a secure and engaging online environment.

## The Phenomenon

Elsagate refers to a phenomenon on online video-sharing platforms, particularly YouTube, where certain children's videos contain inappropriate or disturbing content.

## Methodology

A diverse dataset of 28 million YouTube comments will be analyzed, with advanced natural language processing techniques.

This research aims to develop and evaluate a scalable approach for detecting linguistic anomalies in extensive datasets, with a particular focus on its adaptability and generalization strength.

## Examples of Comments



- "I really do not like this video this is weird"
- "Hihoupjljbffuihyhuuy"
- "Password iPhone"
- "*WHAT AM I WATCHING???* -GOD HELP-"
- "I have officially entered the weird side of Youtube"
- "What did I click on??"
- "Suck! Elsagate!"

- "why does someone want to kill dora that's weird.....actualy its not because I want to kill her to so yeah"
- "Would you say you've drank less alcohol than that guy smoked crack, or that your bigger alcoholics than he is a crank head jw. I have my guess"
- "You're right about the ghost remedy...my grandmother used to do it. With urine, sometimes period blood and nails. She would put a concealed jar near the front door to keep the bad spirits away lol"



| VIEW | ENCODE DECODE | VIEW |
|------|---------------|------|
| Ciphertext ▾ | ROT13 ▾ | Plaintext ▾ |
| bg]bhfbbc[ a`]gb`efh | VARIANT<br>○ ROT5 (0-9)<br>○ ROT13 (A-Z, a-z)<br>○ ROT18 (0-9, A-Z, a-z)<br>◉ ROT47 (!-~)<br>→ Decoded 20 chars in 2ms | 38.397334, 21.831679 |