



## **THE ROLE OF AI IN SAFEGUARDING CHILDREN: A CESIUM CASE STUDY**

**Date:** November 25, 2022

*Disclaimer: This briefing note contains the encapsulation of views presented by the speaker and does not exclusively represent the views of the Canadian Association for Security and Intelligence Studies.*

### **KEY EVENTS**

On November 25, 2022, Stephen Anning, Product Manager for CESIUM at Trilateral Research, presented on *The Role of AI in Safeguarding Children: A CESIUM Case Study*. The presentation was followed by a question-and-answer period with questions from the audience and CASIS Vancouver executives. The key points discussed were the future of privacy, CESIUM's ability to audit evidence and provide a user-friendly atmosphere, information management and dataset governance, and mitigating the impact of techno-colonialism.

### **NATURE OF DISCUSSION**

#### **Presentation**

On the role of Artificial Intelligence (AI) in safeguarding children, Mr. Anning defined and described the concept of Ethical AI. He discussed current safeguarding arrangements involving children in the U.K, and used a synthetic creation to demonstrate how the CESIUM platform could be used to augment human judgement.

#### **Question & Answer Period**

Mr. Anning commented on the future of privacy, auditing CESIUM for forensic evidence disclosure, its role in providing a user-friendly atmosphere, best practices in data management and processing, and safeguarding mechanisms that could be implemented to prevent techno-colonialism from overriding cultural norms.

## BACKGROUND

### Presentation

Mr. Anning' discussed the benefits and limitations of Ethical AI in safeguarding children, noting how it is utilised within the context of current safeguarding arrangements in three statutory agencies: the Children's Services in a local authority, the Clinical Commissioning Group, and the National Police Force. Presently, safeguarding arrangements are derived from the Working Together to Safeguard Children Policy, which acts as a government-issued directive for inter-agency collaboration towards the protection of children and promotion of their welfare. It also covers a range of non-statutory partners as well.

Discussing CESIUM, Mr. Anning purported that the above safeguarding arrangement is underpinned by a need for logistical coordination amongst partners with regards to user data. As a result, the purpose behind CESIUM's design is to empower the professionals to share information while facilitating transparency. CESIUM acts as a general-purpose tool, as well as a platform that provides analytical insights, efficiency, and data security; the latter of which can be subject to data siloing and the aim is to connect each database. The intentions behind data siloing differ depending on the agency, and the current dilemma is the balance of individual privacy and the minimum standard of interagency data sharing.

Mr. Anning defined Ethical AI as narrow or weak AI that consisted of augmented intelligence, logic and mathematics, and Machine Learning (ML), distinguished from general or strong AI which is able to exhibit human-level intelligence. Ethical AI augments, rather than replaces human judgement, using outputs derived from ML, logic, and mathematics to generate a rationale for a particular course of action. Ethical AI forms the basis for the platform CESIUM, identifying and assessing vulnerable children that are known to the agencies, but have not been referred for exploitation concerns.

Mr. Anning provided a demonstration of CESIUM, displaying the processes involved in referral to the Multi-Agency Child Exploitation group (MACE), identification of children suspected of being exploited, and the interpretability of the system's outputs. Using an artificial case created through synthetic data, Mr. Anning demonstrated the features of CESIUM, highlighting the ways in which the system can augment and enhance human judgement. First, he used the person search engine to bring up information on the individual, which indicated that they had not been referred to MACE previously. Next, he displayed a list of transcript

statements and used the system to pinpoint indicators flagged through natural language processing, noting the alleviation of the analytical burden for personnel seeking to extract information. He then demonstrated the timeline feature to provide insights into a child's history, creating a more complete picture that could indicate exploitation as a cause of delinquency. Lastly, he used CESIUM to examine association charts, which highlight a child's noteworthy associations, such as an individual with criminal history or severe charges against them.

Mr. Anning stressed that CESIUM is not meant to be a risk assessment tool, and that only a human being should hold this power. CESIUM is intended to act as a searchlight that assists through data synthesis and interpretation, reducing the burden on analysts and enhancing the processes involved in the safeguarding of children.

### **Question & Answer Period**

Mr. Anning provided his thoughts on the future of privacy, stating that current circumstances are characterised by a heightened scrutiny on behalf of professionals cautious of sharing data. He stressed that this over-caution can lead to poor data sharing between organisations, sometimes resulting in cases of child exploitation going unaddressed to deleterious results. He noted that this can be exacerbated by legislation surrounding the protection of privacy.

In terms of auditing CESIUM for evidence disclosure and the possibility of exporting system data into software for link analysis, Mr. Anning stated that CESIUM contains a monitoring and auditing system that keeps detailed logs of pertinent information related to records and users, but that there are no current plans to integrate it elsewhere. Currently, the priority is on a sister-project surrounding the use of graphing technology to heighten visualisation of link analysis. The future applications of this technology could involve the mapping of organised crime with a focus on child exploitation.

Mr. Anning asserted that CESIUM's ease of use and program capabilities is largely a result of regular focus meetings involving designers and end-users, in which a discussion group hosts developers and end-users to discuss requirements and ways to integrate improvements. The platform's design is based on an iterative process that is driven by large amounts of feedback from end users and is meant to be intuitive and user-friendly.

Mr. Anning stated that although CESIUM cannot access the wealth of data stored on an individual's C drive or sharepoint server, the sister project is able to

integrate this data and apply it to organised crime mapping. He suggested that “techno-colonialism” is driven by big data companies who monetize users through their data, and this has led to a centralisation of power around social media companies. Mr. Anning noted that there is now a pushback by proponents of Web3, who hope that decentralisation will create a more equitable online space.

## KEY POINTS OF DISCUSSION

### Presentation

- The purpose of CESIUM is to empower the professionals to share information while facilitating transparency. CESIUM acts as a general-purpose tool, as well as a platform that provides analytical insights, efficiency, and data security
- Ethical AI augments rather than replaces human judgement. Ethical AI consists of augmented intelligence, logic and mathematics, and machine learning, distinguished from general AI which is able to exhibit human-level intelligence.
- Ethical AI forms the basis for CESIUM by assisting analysts in identifying and assessing vulnerable children that are known to the agencies, but have not been referred for exploitation concerns.
- The three primary agencies involved in the MACE arrangement are the police force, Children’s Services, and the Clinical Commissioning Group. At present, each of these agencies safeguard the data under their care through siloing.
- CESIUM is intended to act as a searchlight that assists through data synthesis and interpretation, reducing the burden on analysts and enhancing the processes involved in the safeguarding of children.

### Question & Answer Period

- Current circumstances surrounding privacy are characterised by heightened scrutiny on behalf of professionals cautious of sharing data, which can lead to poor data sharing between organisations.
- CESIUM's Design is based on an iterative process that is driven by large amounts of feedback from end users, leading to an intuitive and user-friendly result.
- It is hoped that Web3, due to its decentralising capabilities, will mitigate negative effects of techno-colonialism and monetisation of individual privacy by large tech companies.



This work is licensed under a Creative Commons Attribution-Non-Commercial-NoDerivatives 4.0 International License.

© (STEPHEN ANNING, 2023)

Published by the Journal of Intelligence, Conflict, and Warfare and Simon Fraser University

Available from: <https://jicw.org/>