

REPHRAIN MAP: Progress Update



Agenda



- 01 Aim of the MAP
- 02 Our Approach
- 03 Current State
- 04 Demo
- 05 New features and challenges
- 06 Future plans

Leading Team



Kopo Marvin Ramokapane
marvin.ramokapane@bristol.ac.uk



Emily Godwin
emily.godwin@bristol.ac.uk



Alicia Cork
ac974@bath.ac.uk



Partha Das Chowdhury
partha.daschowdhury@bristol.ac.uk



Andrés Domínguez Hernández
andres.dominguez@bristol.ac.uk



Emily Johnstone
emily.johnstone@bristol.ac.uk

REPHRAIN MAP



AIM

to establish a baseline of current state-of-the-art.

Features

- a living resource (updated regularly)
- inspired by the Mitre ATT&CK framework 11 for technical cyber attacks
- (key distinction) the REPHRAIN Map is socio-technical
- Allow ability to drill down into advances that mitigate against particular harms.
- a barometer to evaluate the Centre's progress with regards to the baseline

Our Approach

Phase 1: Scoping sessions

- Various scoping sessions with academia, industry, partners, and organisations
- Identified five key components

Phase 2: Design

- Drafted visual designs
- Harm centric instead of project centric

Phase 3: Populating the MAP

Data Collection

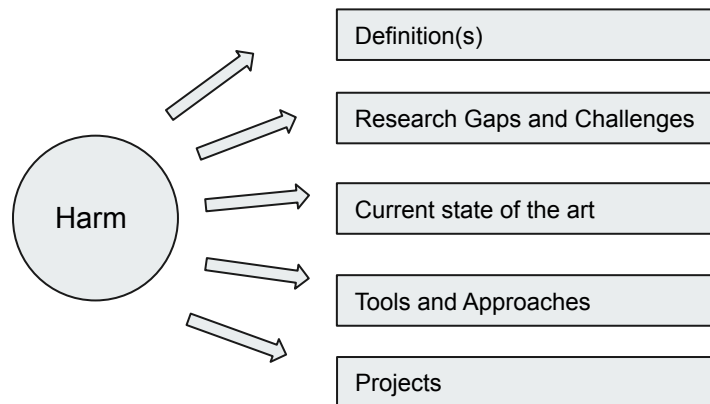
- Online forms
- Workshops
- One-on-One meetings
- Online searches
- Emails

Data Curation

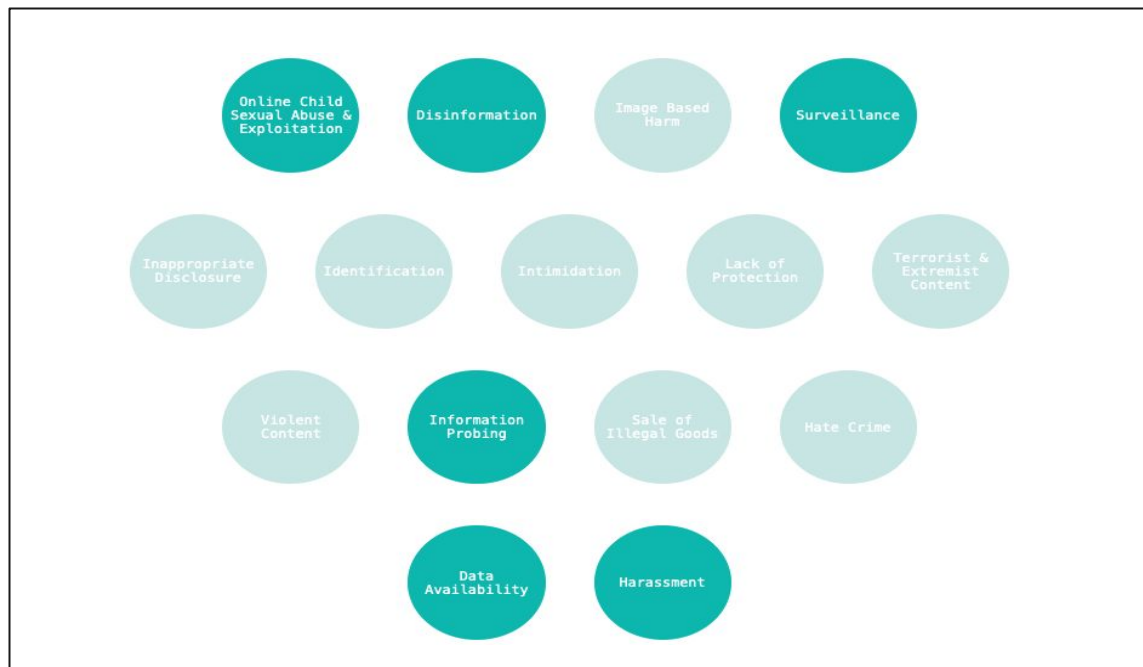
Update Map

Current State

- **Six (6) online harms**
 - Disinformation
 - Surveillance
 - Online Child Sexual Abuse and Exploitation
 - Information Probing
 - Human trafficking
 - Inappropriate/non-consensual disclosure
- **Five (5) Components**
 - Definition
 - Definition of online harm
 - Research Gaps and Challenges
 - List of identified gaps in literature
 - Existing Challenges addressing the harm
 - Current State of the art
 - Existing literature on each harm
 - Tools and Approaches
 - Existing tools and REPHRAIN Initiatives
 - Projects
 - REPHRAIN Projects aiming to address the harm



Demo!



Upcoming Changes

Landing page

- We will add a MAP description and rationale

Categories

- Explanations of the headings:
 - Research and practical challenges,
 - State of the art,
 - Bills, laws and regulations

New harms and updated terminology

- Human trafficking / modern day slavery
- Information probing and phishing
- Cyber bullying and harassment

Frontend design

Demo!

Go Back to Harms



Definition(s)

Online child sexual abuse and exploitation (OCSEA) involves the use of information and communication technology as a means to sexually abuse or sexually exploit children, or both concurrently (Interagency Working Group, 2016, p. 23 and 28).

The United Nations Economic and Social Commission for Asia and the Pacific (UN ESCAP, 1999) defines **child sexual abuse** as "contacts or interactions between a child and an older or more knowledgeable child or adult (stranger, sibling or person in a position of authority such as a parent or caretaker) when the child is being used as an object for the older child's or adult's sexual needs. These contacts or interactions are carried out against the child using force, trickery, bribes, threats or pressure."

Child sexual exploitation involves child sexual abuse or other sexualised acts using children that involves an exchange of some kind (e.g., affection, food, drugs, and shelter) (UNODC, 2015).

Demo!

Go Back to Harms



State of the Art

Academic Publications

Liberatore, M., Erdely, R., Kerle, T., Levine, B. N., & Shields, C. (2010). Forensic investigation of peer-to-peer file sharing networks. *Digital Investigation*, 7, S95–S103. ([DOI](#))

Peersman, C., Schulze, C., Rashid, A., Brennan, M., & Fischer, C. (2016). iCOP: Live forensics to reveal previously unknown criminal media on P2P networks. *Digital Investigation*, 18, 50–64. ([DOI](#))

Sanchez, L., Grajeda, C., Baggili, I., & Hall, C. (2019). A Practitioner Survey Exploring the Value of Forensic Tools, AI, Filtering, & Safer Presentation for Investigating Child Sexual Abuse Material (CSAM). *Digital Investigation*, 29, S124–S1 ([DOI](#))

Phippen, A., & Brennan, M. (2019). *Child Protection and Safeguarding Technologies: Appropriate or Excessive 'Solutions' to Social Problems?* Routledge. ([DOI](#))

Policy Documents

[Draft Online Safety Bill](#)

White Papers & Reports

Information is being curated

Demo!

[Go Back to Harms](#)



Challenges

Research Gaps & Challenges

- Access to data to train AI models in a secure and privacy preserving way.
- Overfitting of AI models based on bias in training data.
- Sharing data in a secure and privacy preserving way.

Demo!

[Go Back to Harms](#)

Tools & Approaches



REPHRAIN Initiatives

[AUTAPP](#)

A multidisciplinary approach will be used, combining the development of intelligent technologies to tackle online harms with a qualitative approach that will incorporate a series of semi-structured interviews with relevant stakeholders combined with an open-source literature review.

[SURVEY](#)

A better understanding of the policy process around this complex issue would provide valuable insights and learning opportunities and, critically, it would provide clear indications of areas of common approach that could lead to international policy collaborations or alignment..

Other Resources and Approaches

[The 4Cs: Classifying Online Risk to Children](#)

Demo!

[Go Back to Harms](#)

Projects

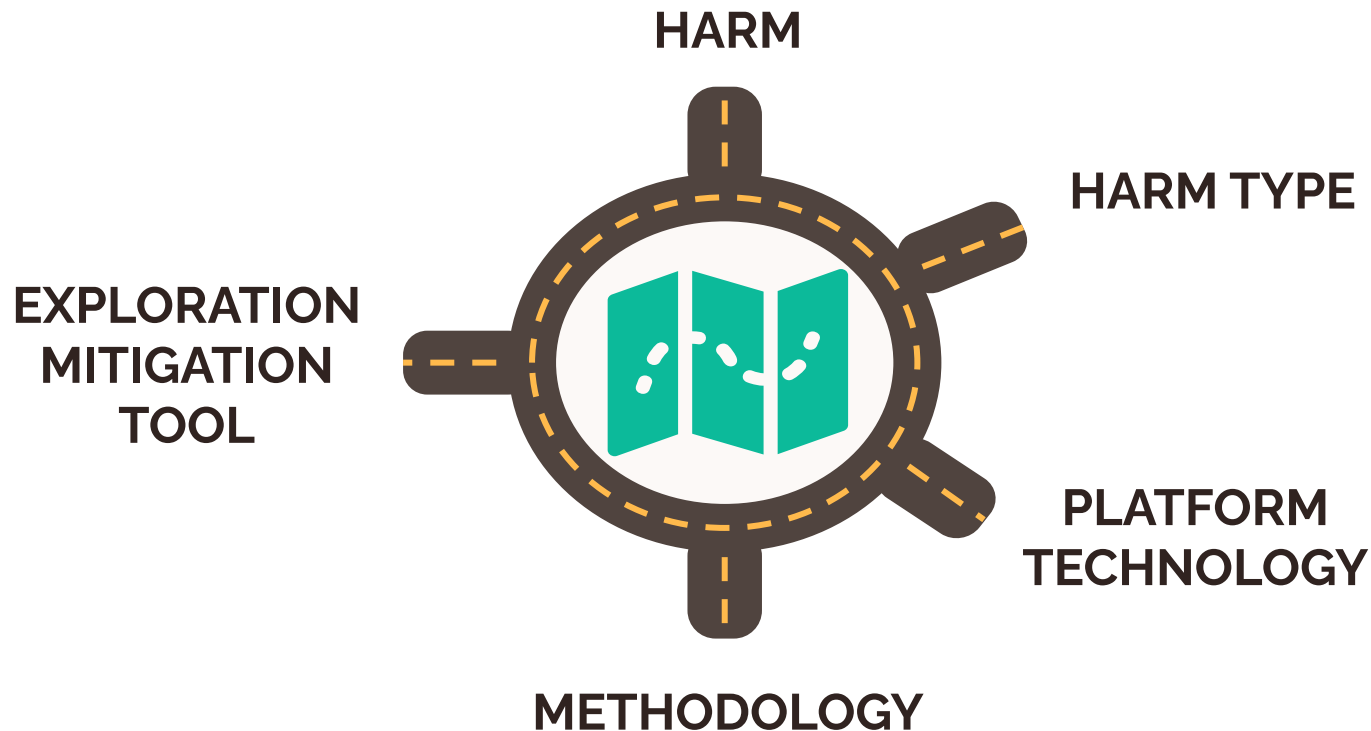


Projects	
AUTAPP	Automated Detection of Online Harms for Social Media Applications
SURVEY	Global Survey of Policy Approaches to Protecting Citizens Online

Mapping literature from REPHRAIN researchers!

Emily Johnstone and Emily Godwin

Our coding scheme



Challenges



1

EVOLVING LANDSCAPE

We anticipate new harms will be added to the map as research and technology develops.

2

DUAL-USE DILEMMA

Some tools mitigate against one harm but potentially lead to another.

3

OVERLAPS

Importance of using an iterative process to ensure categories are clear and distinct.

Future Plans



- Seamless approach to updating the MAP
 - Usability of updating and maintaining the MAP
- REPHRAIN researchers profiles to the MAP
 - Link Profiles to projects
- Timeline
 - Expected outputs
 - Evolution of the MAP
- Podcast from experts
 - Audio and Video
- UX design



All Hands Meeting Friday workshop

We want to test:

- New ways of visualising and categorising harms
- A workflow to update and maintain the map



Thank you!