

HIDDEN REALITY HACKATHON - HACK PACK

Contents

Trigger Warning	2
Core Hackathon Team	2
Logistics	2
Agenda	3
Snowflake Technology & Skills	3
Intro to Hackathon	4
Data Providers & Data Sources	5
Themes & Problem Statements	6
Submission	9
Partnerships & Collaboration	9
Code of Conduct	10
Contact Information	10

TRIGGER WARNING

This Hackathon contains references to themes of harassment including sexual assault and violence and to other violence against women and girls, including domestic violence and abuse. You are welcome to leave this session at any time.

If you are affected by the issues raised, you can access direction to recommended support services and advice through these [lines of support](#).



CORE HACKATHON TEAM

- **Sarah Kavanagh** - Director of Research (Women in Data)
- **Michelle Conway** - Hackathon Co-Lead (Women in Data)
- **Molly Bevan** - Hackathon Co-Lead (Impactic Volunteers)
- **Galiya Warriar** - Senior Sales Engineer (Snowflake)
- **Harini Gopalakrishnan** - Field CTO Lifesciences (Snowflake)
- **Alex McCarthy** - Senior Solutions Architect (Snowflake)

LOGISTICS

- **Date:** Thursday, 7th March 2024
- **Time:** 9:50-17:45 (agenda below)
- **Venue:** InterContinental London – The O2, 1 Waterview Drive, Greenwich Peninsula, London SE10 0TW
 - **Room:** Peninsula Ballroom (below main auditorium)
- **Arrival instructions:**
 - Registration will open at 9:50 - please do not arrive before then
 - Please bring photo ID
- **Team Formation**
 - Teams will either be:
 - From 1 organisation (organisation to select their team ahead of the day)
 - Made up of individuals (allocated on the day)
 - Each team will have 1 Team Lead

- Teams should contain 6-10 data professionals, ideally with a mix of the following experience:
 - Project Manager
 - Data Analyst
 - Data Scientist
 - Data Engineer
 - Data SME
- **Please bring your own tech (e.g. laptops, chargers)**
 - This can be your work or personal laptop (as long as you are allowed to use your work laptop for this purpose)
 - You will be using the venue's WiFi on the day
- **Lunch & refreshments will be provided**
- **Dress code:**
 - We are asking you to wear something red. This doesn't need to be a full outfit (unless you want it to be). It could be as simple as a headband, bag, shoes or bracelet. Red signifies confidence, ambition and drive, and is the colour of strength and vitality. Please note, this is optional.

Note: This event is running at the same time as the Flagship event. Committing to the Hackathon means you will not be attending the Flagship, although there will be some live-streamed sessions, as well as opportunity to visit the exhibition stands throughout the Hackathon.

AGENDA

- 9:50-10:20 - Registration
- 10:30-11:00 - Hack kick-off
- 11:00-11:15 - Bright Data Spotlight
- 11:15-13:00 - Hacking
- 13:00-14:00 - Working lunch (+ Crime Data Mapping - Imy Younge from Capgemini)
- 14:00-16:30 - Hacking
- 16:30-17:30 - Live-streamed main stage sessions
- 17:30-17:45 - Hack close
- 17:45-19:00 - Drinks & networking

SNOWFLAKE TECHNOLOGY & SKILLS

- Snowflake

- Please watch the Pre-Event Familiarity Session recordings ahead of the day:
 - [Familiarity Session \(Part 1\) - for all](#)
 - [Familiarity Session \(Part 2\) - for Team Leads only](#)
 - Team Leads should watch both parts
 - Part 2 covers how to access the data and create an environment to start executing code in
- **Snowflake [Team Lead Guide](#)**
 - **Team leads should set up their Snowflake environment for their team ahead of the day (one Snowflake account per team)**
- You will not need to download any software - this can all be done in your browser
- We will open up access to the survey data and give instructions for how to add the data “share” to your account. You will not need to import data - you will all be accessing the same data stored in the WiD account
- Skillsets
 - Everyone is welcome - from beginner to expert!
 - This will be a safe space for you to meet new people, leverage your skills, and ultimately make a huge difference to women’s safety using some incredible datasets!
- Coaches
 - There will be coaches on the day to support you and guide your team

INTRO TO HACKATHON

Survey Context

Violence against Women and Girls (VAWG) is an endemic issue. Most incidents go unreported, creating a disturbing gap between reported and unreported gender-based violence.

Throughout September and October 2023 Women in Data® fielded a nationwide survey called ‘**The Hidden Reality**’ - the research was the first in the UK to quantify and understand the societal, personal and economic impact of such abuse.

Hackathon Purpose

In Women in Data’s largest ever hackathon, our community will explore the findings, uncover new insights and deliver recommendations

Delivered as a comprehensive report to help legislators, organisations and survivor groups make data-led decisions that can affect change.

Teams will combine the 'Hidden Reality' survey results with other rich data provided by CARTO, Cybersyn, Bright Data, using a Snowflake environment.

DATA PROVIDERS & DATA SOURCES

1) Hidden Realities Survey (access via Snowflake)

You can click through the survey as many times as you would like, your responses will not be recorded

- [Survey - Familiarity Link](#)



- Survey dataset available on the day in the WiD Snowflake account

2) CARTO

CARTO provide geospatial analytical toolbox and also different datasets:

- Crime data for UK region with latitude and longitude: can be used to identify areas of interest close to which crime occurs .
 - Available on the day via WiD Snowflake account
- [Demographics data](#)
 - It is called spatial feature data set and has Demographics data about population, age, gender, and more like weather and temperature into an easy to use dataset. It's called that because it adds many *features* that are geo in nature and in a really nice format for machine learning
- [Points of interest data](#)
- Open street maps: can be used to identify locations like police stations, hospitals to analyse if there are any trends on crime occurrence near a landmark
- Analytics tool box: you can create cool geospatial applications like the examples here <https://academy.carto.com/>
- [CARTO Quick start guide](#)
 - Shows how to use the maps for rendering)

3) Cybersyn

All data sets will be available on the day (as part of the private listing). It has data sets on the following topics:

- Arrests
- UK Prison population
- Sexual offences prevalence and victim characteristics
- Violence against girl and women

4) Bright Data

Access will be available via WiD Snowflake account

- **Social Media Posts and Discussion Structured Datasets**

Platforms: Instagram and Tiktok

Data Collection type:

1. Keyword/Hashtag search
2. Comments

Collection Geo-Location: UK

Datasets:

1. Andrew Tate

Applied keywords (individually):

#andrewtate #cobratate #hustleruniversity #datingtips #datingapp #dates
#onlinedating #dailydoseofredpill #onlinedatingfail #matches #datingadvice
#gerardkeyspodcast #redpilltruth #redpill #datenight #redpilltraining
#gerardkeys #hustlersuniversity #datejust #dailyredpilldose #trump #tristantate
#tateism #motivation #successmindset #motivationalquotes #moneymindset
#inspiration #tate #god #success #mindset #grind #strongwill #masculinity
#strong #demigod #king #money #wealth #talismans #lifestyle #gops
#churchoftate #democrat #talismantate

Snapshot Names:

Tiktok - '#AndrewTate Posts Search by Keyword'

Tiktok - '#AndrewTate Post's Comments'

Instagram - #AndrewTate Posts Search by Keyword + Comments

2. MeToo

Applied keywords (individually):

#metoo #feminism #feminist #timesup #metoomovement #women #equality
#love #survivor #sexualassault #womenempowerment #believesurvivors
#blacklivesmatter #domesticviolence #abuse #rapeculture #mentalhealth
#womensrights #consent #healing #intersectionalfeminism
#womensupportingwomen #humanrights #girlpower #trauma #selflove
#sexualabuse #nomeansno #bhfy #prochoice #metoo #feminism #feminist
#timesup #metoomovement #women #equality #love #survivor #sexualassault
#womenempowerment #believesurvivors #blacklivesmatter #domesticviolence
#abuse #rapeculture #mentalhealth #womensrights #consent #healing

#intersectionalfeminism #womensupportingwomen #humanrights #girlpower
#trauma #selflove #sexualabuse #nomeansno #bhfp #prochoice

Snapshot Names:

Tiktok - '#MeToo Posts Search by Keyword'

Tiktok - '#MeToo Post's Comments'

Instagram - '#MeToo Posts Search by Keyword + Comments'

Mumsnet.com 'Leave The Bastard (LTB)' Discussions Custom Dataset:

A custom dataset collection of the keyword search 'LTB', collecting all topics and replies resulted when applying the 'past year' filter.

The delivered dataset includes the following key data points:

- keyword_topic: Post topic
- post_title: title of original post
- post_date: Date of the original post
- comment_date_time: the date of the comment/reply to the original post
- comment_content: The comment content

You are not limited to the data sources provided - if you are aware of other open data sources, you are welcome to use those too. There is plenty in the Snowflake marketplace or you can easily load in others.

THEMES & PROBLEM STATEMENTS

It is up to each team which theme and problem statement they would like to focus on. For each problem statement, there are some suggested focus areas, but you are not limited to these.

Theme 1: Perpetrators & Crime

Problem Statements (& some suggested focus areas)

1) Understanding Perpetrator Characteristics: While data exists on VAWG crimes, comprehensive information on perpetrator profiles remains limited.

- Identify common characteristics and motivations among VAWG perpetrators
- Explore potential links between these characteristics and specific types of VAWG crimes

- Highlight potential gaps in data collection and reporting on perpetrator profiles

2) Geographic Variations in VAWG: VAWG prevalence and its characteristics can vary significantly across regions

- Analyse and visualise existing data to identify geographical trends in VAWG prevalence and types
- Investigate potential factors contributing to these variations
- Develop recommendations for tailored VAWG prevention and intervention strategies based on regional differences.

3) Improving VAWG Reporting and Data Availability: The vast under-reporting of VAWG crimes hinders accurate assessments of the problem and effective response strategies. For example, according to research ([Guardian](#) and referenced [ONS data](#)), in 2018 there were just 1128 rape convictions, estimated to be 1% of offences.)

- Using the data, calculate the updated figures
- Analyse barriers to VAWG reporting and suggest evidence-based strategies to enhance reporting rates
- Design an optimal reporting system for VAWG across private, public, and online spaces

4) Shifting Public Discourse on VAWG: Movements like #MeToo have sparked critical conversations, but the public mood and support for VAWG action can fluctuate.

- Analyse current public sentiment towards VAWG and profile who is supportive and actively against the movement
- Identify emerging trends in online conversations, including potential backlash elements and counter-narratives
- Profile who Andrew Tate is appealing to and his key ideology

Suggested Data Sources:

- Hidden Realities Survey dataset
- CARTO datasets
- Bright Data datasets
- Cybersyn datasets

Theme 2: Victims

Problem Statements (& some suggested focus areas)

1) Vulnerability and Risk Assessment: VAWG impacts women differently based on their individual circumstances.

- Develop a framework to identify and assess the risk of VAWG for different groups of women
- Prioritise interventions based on identified risk factors and specific needs of vulnerable groups
- Create a visual tool (e.g., interactive dashboard, stats generator) to present and explore the risk assessment framework

2) Supporting Victims in Leaving Abusive Relationships: Online communities offer valuable support to women seeking to leave abusive relationships i.e. 'Leave the Bastard - LTB' (e.g. copy their financial details, go to Women's Aid)

- Extract and analyse information from online resources to identify common strategies and resources shared by and recommended to victims
- Develop a comprehensive resource guide or knowledge base summarising information and support options for victims leaving abusive relationships
- Explore potential digital tools that can further facilitate access to resources, anonymised support networks, and safety planning for victims

Suggested Data Sources:

- Hidden Realities Survey
- Bright Data datasets
- Desk research

Theme 3: Workplace

Problem Statements (& some suggested focus areas)

1) Addressing the Impact of VAWG in the Workplace: Women from diverse backgrounds can face different forms of VAWG within workplaces

- Analyse how various forms of VAWG manifest differently for diverse working women based on factors like industry, age, position, geography, ethnicity
- Prioritise and propose interventions tailored to address specific needs and vulnerabilities of different working women groups
- Identify best practices and develop a toolkit for leaders and organizations to create a safer and more inclusive workplace environment.

2) Problem Statement 2: Modeling the Economic Impact of VAWG: Women may choose to avoid work due to fear or experience of VAWG, leading to individual and societal consequences.

- Develop a model to estimate the economic impact of VAWG on individual victims, including lost wages, career prospects, and long-term financial security
- Analyse the broader economic impact of VAWG on various sectors and the overall economy

Suggested Data Sources:

- Hidden Realities Survey dataset
- CARTO datasets
- Desk research e.g. economic data

SUBMISSION

Please complete the [submission form](#) at the end of the Hackathon. This will feed into our version 2.0 of the Hidden Realities report. Please be as clear and thorough as possible.

PARTNERSHIPS & COLLABORATION

This Hackathon would not have been possible without our incredible partners:

Snowflake

For Snowflake, the data warehouse built for any cloud, culture is everything and partnering with Women in Data is another route to ensure we are doing all that we can to cultivate a culture of equality in Snowflake and the industry.



Impactic Volunteers

Impactic is a non-profit social enterprise that runs skills-based volunteering events for corporates. Partnering with Women in Data unlocks the fantastic skills of this community to make a huge difference to the charity sector. Impactic's female founder, [Molly Bevan](#), is the Co-Lead for this Hackathon.



A special thanks to our partners who have provided us with invaluable data sources for the Hackathon: [CARTO](#), [Cybersyn](#), and [Bright Data](#).



[Palantir](#) - Immersive AI Experience

Note: This is a side stream and will not interfere with the Hackathon. This is available for one team member per hack team.



Please join us for an immersive, hands on keyboards, dive into the use of Large Language Models, building an understanding of where they can and cannot be trusted, and how you can harness their power safely. In this one hour session you will learn some fundamentals of employing LLMs and have a go at using them to solve challenges on Palantir's Artificial Intelligence Platform. Expect to leave the session with ideas for how AI could help you in your own role.

[Palentir sign up form](#) (to be completed on the day following instructions)

Brand Partners

This event also brings together many of Women in Data's incredible brand partners, with their teams using their skills for good:

- BAE
- Boots
- Capgemini
- Lexis Nexis
- Dstl
- Esure
- Haleon
- Jman Group
- J.P.Morgan
- Kingfisher
- Knight Frank
- Kubrick
- LBG



Impactic
Volunteers



- The Visionaries
- Police NPCC
- M&G
- MOD
- NovoNordisk
- Ofcom
- Moneysupermarket
- QBE
- Sainsbury's
- Schroders
- Vanquis

CODE OF CONDUCT

Please review our full [Code of Conduct](#).

CONTACT INFORMATION

If you have any questions, please reach out to our Women in Data Hackathon Co-Lead, Molly Bevan at molly@womenindata.co.uk

For up to the minute news, be sure to follow our social channels:

[LinkedIn](#), [Twitter](#), [Instagram](#), [Facebook](#)