

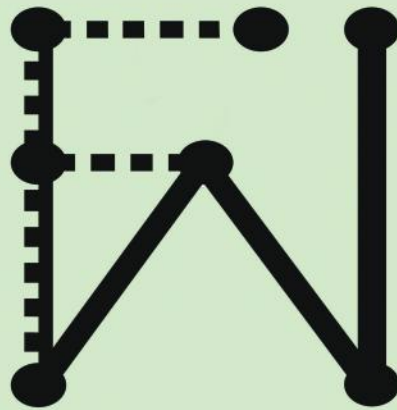


Edinburgh Napier  
UNIVERSITY

Centre for Mind,  
Creativity & Environment  
Research (CEDAR)

# False Webs Network Seminar

**Algorithms, Social Media,  
and Fake News**



**3rd Seminar  
27 March  
10:30-15:30**

**Edinburgh Napier  
Merchiston Campus  
10 Colinton Rd,  
Edinburgh EH10 5DT**

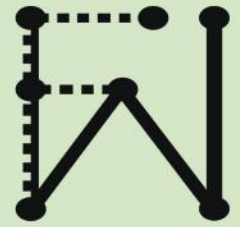
**Made possible  
through funding by:**

**RSE**

*The Royal Society  
of Edinburgh*

KNOWLEDGE MADE USEFUL

# Algorithms, Social Media, and Fake News Seminar



**10:30 - Introduction (Marina Wimmer & Ted Webster)**

## **Speakers Slot 1 - Room G23**

- |   |          |  |
|---|----------|--|
| <b>10:45 - Michael Marshall</b><br><b>Merseyside Skeptics</b>                   | <b>-</b> | <b>How social media pulls people into rabbit holes, and how to get them back out</b>                       |
| <b>11:05 - Ciara Greene</b><br><b>UC Dublin</b>                                 | <b>-</b> | <b>A novel intervention to reduce cognitive biases in the formation of false memories for fake news</b>    |
| <b>11:25 - Charlotte Buecken</b><br><b>KU Leuven</b>                            | <b>-</b> | <b>Can We Reverse False Autobiographical Memories? Testing a Source Sensitization Memory Training</b>      |
| <b>11:45 - Caroline Are</b><br><b>Northumbria University</b>                    | <b>-</b> | <b>Flagging, De-Platforming and Appeals: Content Moderation of Nuanced Content on Instagram and TikTok</b> |
| <b>12:05 - Katarina Alexander &amp; Md Zia Ullah</b><br><b>Edinburgh Napier</b> | <b>-</b> | <b>Fake News Detection and Warning Labels</b>  |

**12:20-2:00 - Complimentary lunch at local restaurant**

## **Speakers Slot 2 - Room H11**

- |  |          |  |
|--|----------|--|
| <b>2:00 - Pip Brown</b><br><b>University of Warwick</b>      | <b>-</b> | <b>Inconstant Vigilance: Adolescent Selective Trust in Online Information</b>            |
| <b>2:20 - Yvonne Skipper</b><br><b>University of Glasgow</b> | <b>-</b> | <b>Project Real. Co-creating interventions to help young people spot misinformation.</b> |
| <b>2:40 - Yaniv Hanoch</b><br><b>Coventry University</b>     | <b>-</b> | <b>Can you spot the real one</b>   |
| <b>3:00 - John Collomosse</b><br><b>University of Surrey</b> | <b>-</b> | <b>Building Safe and Fair Generative AI with Content Provenance</b>                      |

**3:20-3:40 - Coffee Break**

**3:40 - Policy Paper and next seminar discussion**

**5:00 - End of Seminar - Informal pub social**

**6:30 - Complimentary Seminar Dinner at SEN Restaurant**

# Public Engagement Event

On March 26th, the day before the seminar, a public engagement event will take place, organised and hosted by Alice Howarth in collaboration with the Glasgow and Edinburgh Skeptics. Michael Marshall, Yvonne Skipper, and David Robertson will give talks followed by audience Q & A.

This event will take place at 7:30 pm at  
The Three Sisters pub (139 Cowgate, Edinburgh EH1 1JS)

## Seminar Information

This seminar takes place at Edinburgh Napier University  
Merchiston Campus, in room G23 (Morning Session) and H11 (Afternoon Session)

Address: 10 Colinton Rd, Edinburgh EH10 5DT

Starting from 10:30 AM (GMT; Coffee available from 10:00)  
and ending at 5:00 PM

15-minute talks are related to the topic of the seminar:  
“Algorithms, Social Media, and Fake News”  
and include research presentations and topic discussions  
given by researchers and science communicators

Talks will conclude at 3:00, after which will be a discussion about the policy  
paper, and some planning for the next seminar. The seminar will end at 5.

For those staying there, the campus is a 10-15 minute  
walk along Bruntsfield place from the Black Ivy Hotel

There are two  
nearby bus stops:



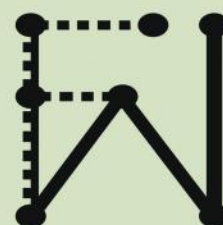
Napier University

Bus 45



Holy Corner

Buses 11, 15, 16, 23,  
36, 45, 101/102



Edinburgh Napier  
UNIVERSITY

Centre for Mind,  
Creativity & Environment  
Research (CEDAR)



# Talk Details

**10:45 – Michael Marshall - Merseyside Skeptics**

**Title: How social media pulls people into rabbit holes, and how to get them back out**

**Description:** When COVID-19 brought the world to a standstill, online groups formed around the idea that the pandemic was a hoax concocted by the governments of the world to instil fear into their people, as a pretext for introducing new, permanent totalitarian laws. These groups spilled out onto the street, in an effort to recruit members of the public to join encrypted messaging channels to learn more about what was really going on. Michael Marshall, full-time skeptical investigator and activist, joined these groups, spending months undercover in the messaging app Telegram, to report on the various conspiracy theories spread there, and how groups used the Covid crisis to radicalise vaccine hesitant members of the public into a dangerous ecosystem of misinformation and extremism.

**11:05 – Ciara Greene – University College Dublin**  
**Collaborator: Gillian Murphy**

**Title: A novel intervention to reduce cognitive biases in the formation of false memories for fake news**

**Description:** Our work over the last few years has shown 1) that people can readily form false memories for fake news stories, and 2) that those false memories are more likely if the news item aligns with the individual's preexisting beliefs or ideology. In this talk I will summarise some of that work, and present evidence regarding a new intervention we have developed to combat this effect.

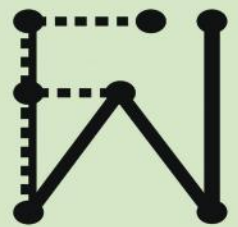
**11:25 – Charlotte Buecken – KU Leuven**

**Title: Can We Reverse False Autobiographical Memories?**  
**Testing a Source Sensitization Memory Training**

**Description:** Much misinformation regarding mental illness is spread on social media platforms such as TikTok. This misinformation can generate false self-diagnoses and could even lead to false autobiographical memories, for example when faced with suggestions on events that could have triggered the self-diagnosed mental illness. In the current study we examine if false memory rates of autobiographical events can be reduced using a source monitoring training aimed at differentiating between memories from different sources.

**11:45 - Carolina Are - Northumbria University**

**Title: Flagging, De-Platforming and Appeals: Content Moderation of Nuanced Content on Instagram and TikTok**



**Description:** In this talk I will briefly introduce the struggles that creators of nuanced content face on mainstream social media platforms such as Instagram and TikTok, sharing the main findings from my postdoc project at the Centre for Digital Citizens. Showing how platform policies can assemble with algorithmic and human moderation to facilitate discrimination and mistaken enforcement, I will provide examples from the experiences of sex workers, sex educators, activists and journalists to explain how platform infrastructure has been gamed in newsworthy health, labour, political and self-expression scenarios.

**12:05 - Katarina Alexander & Md Zia Ullah**

**Title: Fake News Detection and Warning Labels**

**- 12:30-1:45 Lunch at nearby restaurant -**

**Speakers Slot 2**

**2:00 - Pip Brown - Warwick University**

**Title: Inconstant Vigilance: Adolescent Selective Trust in Online Information**

**Description:** The development of epistemic vigilance towards online information is crucial for adolescents in the context of widespread online 'information pollution'. My doctoral research focuses on trust decisions between online sources. While pre-adolescents have demonstrated selective mistrust of webpages with typographical errors, this is not the case for webpages with semantic errors. We use a selective trust task to investigate whether this pattern of results changes through adolescence. In this task, participants read two pairs of sources about different scientific topics, each pair containing a webpage with either semantic or typographical errors. When asked a novel factual question, which source participants draw the answer from indicates the presence or absence of selective trust in the source. I will discuss findings from our first experiment and their implications for our understanding of adolescent epistemic vigilance. Additionally, in view of the fact that dialogic collaboration helps buttress adolescent reasoning skills, I will also present some initial findings from our second experiment, in which we investigate collaborative selective trust judgements through the use of online dyads.

**2:20 – Yvonne Skipper – University of Glasgow**

**Title: Project Real. Co-creating interventions to help young people spot misinformation.**

**Description:** Misinformation is one of the greatest challenges of our time (WEF, 2024), therefore, equipping young people with the skills to navigate misinformation is essential. However, most interventions are designed for, rather than with, young people. Project Real develops co-created, research-based interventions which help young people develop their skills in spotting misinformation. This talk will describe key interventions including lesson plans, comics and simple games—developed in collaboration with Education Scotland, Police Scotland and influencers. It will also discuss the co-creative process and explore how collaborating with young people can enhance digital literacy and resilience in an increasingly complex online world.

**2:40 – Yaniv Hanoach – Coventry University**

**Title: Can you Spot the Real One**

**Description:** In this presentation, I will discuss research focusing on individuals' ability to identify ticket fraud and an intervention to improve their abilities.

**3:00 – John Collomosse – University of Surrey**

**Title: Building Safe and Fair Generative AI with Content Provenance**

**Description:** Provenance facts, such as who made an image and how, can provide valuable context for users to make trust decisions about visual content. Emerging standards and provenance enhancing tools promise to play an important role in fighting fake news and the spread of misinformation. In this talk we contrast three provenance enhancing technologies: metadata, fingerprinting and watermarking, and discuss how we can build upon the complementary strengths of these three pillars to provide robust trust signals to support stories told by real and generative images. Beyond authenticity, we describe how provenance can also underpin new models for value creation in the age of Generative AI. In doing so we address other risks arising with generative AI such as ensuring training consent, and the proper attribution of credit to creatives who contribute their work to train generative models. We show that provenance may be combined with distributed ledger technology (DLT) to develop novel solutions for recognizing and rewarding creative endeavour in the age of generative AI.