Proposed Project Title: The Design, Deployment and evaluation of Technologies for Refugee Community Resilience by Reem Talhouk

Purpose

Refugee communities face challenges in accessing services [3,12,13,17,25,27], integrating into host communities [9,36,39] and journeying to other countries [34]. They experience limited access to public services, such as healthcare [12,27] and education [13], difficulties in acquiring employment and restrictive economic policies [17,25,26], and difficult journeys to asylum [34]. These aspects render refugee populations vulnerable. This research proposes to build on my previous work with a Syrian refugee community residing in an informal tented settlement (ITS) in Lebanon by investigating the role of technology in building refugee community resilience through designing, developing, deploying and evaluating technologies for this purpose.

Objectives

This research aims to explore the following research questions:

1. What is the definition and characteristics of "community" within the context of protracted refugees?
2. How is technology currently being used by refugee communities to strengthen/support the different aspects of community resilience?
3. How do refugee communities envision using technology to strengthen/support the different aspects of community resilience?
4. How does the involvement of refugee communities in the design, development, deployment and evaluation of community technologies strengthen/support the different aspects of community resilience?
5. What are the appropriate design methods for engaging refugee populations in the development of technological tools aiming at strengthening their resilience?
6. How do community technologies (existing and newly developed) strengthen/support the different aspects of community resilience?
7. How do the current community resilience models lend to refugee contexts and use of technology within refugee contexts?

Motivation

Currently, there are 21.3 million refugees worldwide [37]. Refugees are individuals that have had to forcefully leave their countries due to fear of being persecuted based on race, religion, nationality, political opinion and/or being of a particular social group [35]. Upon leaving their country of origin refugees face challenges in accessing services, integrating into host communities and journeying to other countries. Refugees experience limited access to public services, such as healthcare and education, difficulties in acquiring employment and restrictive economic policies, and difficult journeys to asylum [2]. These aspects render refugee populations vulnerable and impacts their wellbeing in both the initial phase of being a refugee and when they resettle in camps and host communities. Additionally, the difficulties in integrating into host communities and exposure to new laws and cultures have been identified as barriers to community resilience [14]. The definition of community resilience to be used within this PhD proposal is Magis's [20] definition of community resilience: The ''existence, development and engagement of community resources by community members to thrive in an environment characterized by change and unpredictability''. Within refugee contexts, policies and programs that support community resilience through building self-reliance, such as those implemented in Uganda, have been shown to lead to a decrease in refugee reliance on aid [10]. The traditional support for refugees comes in the form of humanitarian aid. However, Berkes & Ross [6] have indicated that there is a need to use psychological and social ecological models that emphasize refugee community resilience to decrease the refugee population's vulnerability. This is especially true since

longitudinal reliance on humanitarian aid marginalizes the natural support systems, such as community experts, family and neighbors [18] that are the foundation of community resilience.

The recent refugee crises differs from previous ones as it has been characterized to be the first refugee crisis in the digital age. Refugees worldwide are actively using technology, mainly smartphones, to (1) access services, (2) support one another and (3) navigating within their new environment [4,7,32]. In line with Digital Civics [24], the increased access of refugees to technology opens up space for us to revisit how refugees access and even create services to decrease their vulnerability and increase their resilience. Indeed, my previous work with Syrian refugees residing in informal tented settlements in rural Lebanon [33] identified the potential for technology to be used in improving access to antenatal healthcare through leveraging currently available services and the peer-to-peer support currently in existence. Additionally, it was highlighted that technology can be used to increase refugee agency within these contexts [33]. A systematic review, I co-authored, on the impact of digital technology on the health of populations affected by humanitarian crises [21] identified that, while there is potential for technology to positively impact the health of refugee communities there are current gaps in the evaluation of these technologies and more importantly that the current technology replicates the paternalistic approach that is characteristic of traditional humanitarian aid. The findings further motivate investigating how technology can be used by refugee communities to build their community resilience and engage with stakeholders within the refugee community and the humanitarian aid system to evaluate the technology.

Refugee Community Resilience & Human Computer Interaction (HCI)

HCI literature involving refugees does not particularly focus on refugee community resilience however aspects of the research do inadvertently aim to strengthen some of the adaptive capacities in figure 1. The majority aim to strengthening the information and communication adaptive capacity [22]. Branoff et al [5] placed near field communication tags around a U.S. city that could be scanned by refugees and would inform them, in their preferred language, where they are in the city and of nearby facilities. Similarly, Ritravan mediates between refugees and formal processes, they have to undergo in the U.S., by placing within the loop of communication asynchronous informal messages between refugees and interpreters [8]. It is important to note that both technology designs aimed to be used by refugees did not involve refugees in the design stage, nothing further than conducting focus groups. This approach has been critiqued to be paternalistic within refugee contexts as they are done for refugees and not always with [19]. Consequently, there is a need to explore participatory and co-design methods when designing technologies with refugees. Other work coming out of Palestine has exemplified the potential for computer clubs to increase community competence and social capital [1,41]. Involving refugee children in computer clubs not only built their capacity in using and making technologies but also had a positive effect on their communication and collaboration skills[1,41]. These skills in turn contribute to active participation within their community which strengthen social capital. Xu et al [40] suggests the use of participatory refugee camp mapping as a way of supporting community building, participation and organizational collaboration. Similarly our team, Talhouk et al [33] highlighted the opportunities for using technology in leveraging the already existing social support practices of Syrian refugees in order to increase access health services.

Approach and Methods

In order to answer the above-mentioned research questions several methods will be used under the umbrella of the Action Research approach (AR).

Approach

This study will follow an AR approach [15,31] that has guided my previous work with the community. Previously, focus groups and discussions with the community identified that they would like to have a direct medium of communication with healthcare providers and that there are several barriers to successful communication between refugees and healthcare providers. This process allowed us to design, develop, deploy and evaluate a telephony based technology that configured communication between the women of the refugee community and the healthcare providers to provide health education and address the communication barriers. We were able to observe improvements in trust and understanding between the two stakeholders, an increase in agency amongst the refugees and the interactions of the community members with technology and with each other after the introduction of the technology.

Accordingly, knowledge is co-constructed through engagements with the community members as they identify their primary problems, develop action plans to address them and evaluate the implementation of these plans. In AR's aim of doing "meaningful research" researchers must undergo participatory activities and engagements that allow stakeholders within communities to come to a common definition of their needs/problems and priorities[31]. Furthermore, it calls for stakeholders to jointly formulate solutions to the community needs/problems [31]. Through this co-construction of knowledge, AR enables researchers to "extend their understanding of the experience and perspective of stakeholders" and consider the community as research partners that enrich interventions and evaluations [15,31]. Taking this into account, the research proposed in this study will build on experience and knowledge attained in the field, working with Syrian refugee communities. Furthermore, knowledge will be generated through taking action [15], in this case the designing, development, deployment and evaluation of technology to address community needs.

Community Recruitment

Communities that I have previously worked with in my previous work and communities with which our collaborators are currently working with will be approached to participate in the proposed research. The main community proposed to partake in the research is the current Syrian refugee community that I engaged with when conducting my MRes. The community of 84 refugees reside in Bekaa, a rural area, in Lebanon. This will allow us to build on the current trust between the community and the research team and on the research team's knowledge of the community dynamics and priorities.

Methods

In line with the AR approach previously mentioned, the research suggested in this PhD proposal would follow the stages below. These stages will be followed in the three case studies to be investigated by this PhD.

1) Community Engagement & Design
   a) Procedure: Community engagements in the form of stakeholder meetings, focus groups and interviews will be conducted to understand (1) community needs/priorities, (2) how they fit within community resilience frameworks, and (3) how they are using current resources (including social capital and technology) to address their needs/priorities. The acceptability and feasibility of using technology to capture the afore mentioned community aspects will be explored by integrating technology in to community engagement activities.

      Sample data: Qualitative data in the form of transcribed audio recordings and observational notes Data Analysis: Thematic analysis guided by community resilience literature and frameworks

2) Procedure: Co-design workshops with the community will also be conducted to generate common goals, ideas on how to leverage current resources and how technology can facilitate the solution in a sustainable manner [28,30]. Ideas regarding the possible technologies that would address the issues would be generated. The co-design workshops will be informed by previous literature and data collected by the researcher [23].

   Sample data: Qualitative data in the form of transcribed audio recordings, sketches, story boards created, paper prototypes and observational notes
   Data Analysis: Thematic analysis guided by community resilience literature and frameworks and ICT/ HCI for development literature.

3) Development
   Analysis of data collected from interactions, engagements and workshops would be analyzed. In cases where currently available technologies are applicable, the technologies will be tailored/ scaffolded to meet the needs of the project. If not, technologies would be developed. It is important to note that in accordance with community resilience and AR principles, community members with an interest and skills that would aid in the design and development of the technologies would be encouraged to take a more active role at this stage. They will also be involved in the design of the implementation strategy. Additionally, the community would evaluate the usability and design of the technology while it is being developed. Iterations would be made accordingly.

4) Deployment, Monitoring & Evaluation
   Procedures: Technologies developed would be deployed and used by the community. Interactions while the technology is being used would be observed. Post-deployment evaluations would be conducted through interviews and focus groups with all related stakeholders within the community. Evaluations would include discussions around the technology but also around whether community goals set in the community engagement & design phase were met.

   Sample data: Qualitative data (in the form of transcribed audio recordings, sketches, story boards created, paper prototypes and observational notes) and quantitative data generated from the use of the technology (e.g. data logs).
   Data analysis: Thematic analysis (guided by community resilience literature and frameworks and ICT/ HCI for development literature) and statistical analysis.

At the end of the three case studies, an evaluation of whether the overall research did strengthen community resilience will be conducted (informed by community resilience frameworks). Results will be continuously presented back to the community and other stakeholders within the humanitarian aid system to inform the decision of whether the technology and program should be sustained.

Ethics

Given that the research conducted is driven by the community. Ethics applications will be submitted for each case study at each stage of the research. Steps one and two of the methods section will have their own ethics form as the study design is different to that of the third step where a technology will be deployed and the technology will be dependent on the results of steps two and three.

Working Title: Artificial Intelligence and Machine Learning Biometrics as Field Intelligence and Criminal Evidence by Luke Chambers

Introduction and Research Significance: Artificial Intelligence (AI) has become intrinsic to our everyday lives over the past two decades and has radically altered many elements of our society - with policing and justice being no exception. We already know that AI and Machine Learning (ML) are capable of identifying people with biometrics such as (but not limited to) Facial Recognition and Gait Analysis - but what is relatively untested is the use of these and similar AI as short-term police intelligence or as evidence against suspects in criminal trials.

AI and ML is far from infallible, and has in fact shown to be subject to many of the same biases as the general population. AI is only as good as the data which feeds it, which means that many of the modern criticisms of our legal system and those who enforce it, such as criticisms of more focus on ethnic minority communities or on those from lower socio-economic backgrounds, will also be present in AI systems. The question is raised on whether AI and ML can be relied upon in an evidential role without violating the suspect's ECHR Article 6 Right to a Fair Trial, or even whether it can be considered a solid basis for the initiation of an investigation of any size which would violate a suspect's ECHR Article 8 Right to Privacy. Both of these Rights are absolutely vital for our democracy and our justice system, so investigation of these issues not only expands society's knowledge in this area but also contributes to our social structure as a whole.

Concise Critical Reflection on the Existing Literature:
There is an abundance of existing literature on the field of AI as well as almost as much on how AI affects and is affected by our laws.

Before considering any of these, it is important to lay one criticism down on the current literature and that it is often far too compartmentalised to achieve its true potential. For example, publications on anti-crime AI systems from the STEM-centric side of academia and practice tend to focus only on the inner workings of the AI systems. Publications from the Law side, on the other hand, tend to focus only on how these systems and their outputs affect and are affected by law. This weakness may seem slight, but the inner workings of the systems have a massive impact on whether or not they can be considered fair. A single slight alteration of how data is processed, or how an output is reached, can change entirely whether or not an AI system could be considered 'fair' or even 'accurate', and so research must balance technical detail with legal acumen and also a good sociological understanding not just of the justice system but those who use it.

I graduated from my BSc (Hons) Computer Science with Artificial Intelligence in 2020, and I am currently studying Space Law(LLM) with modules in AI law which gives me a good overlap between technical knowledge and an understanding of algorithms with how the law works in the cyber domain. Fortunately, there are some reliable and recent sources with a good overlap as well.

Artificial Intelligence, Computational Modelling and Criminal Proceedings by Serena Quattrocolo focuses on how AI can 'affect the investigation and adjudication of crime' with a significant focus on human rights under ECHR and how 'predictive justice' can impact judicial decision-making. It is an Italian work written in English, so does not directly correlate to the law of England and Wales but does correlate well with Europe-centric issues and laws. In the field of counter-terrorism, 'Data Mining Based Crime Analysis Mapping and Intrusion Detection' (Panja et al, 2020) provides a very good algorithmic overview of AI in policing, but suffers from a US-centric focus. The Law Society of England and Wales

published a report in 2019 on algorithm use in the criminal justice system, which provides a good oversight of the landscape as it was 2 years ago. The literature landscape in this area is very active. 'Predictive Policing and Artificial Intelligence' is set for release in February 2021 and looks from previews to provide some value as a source - although these same previews show it may suffer from the lack of technical detail discussed in the opening of this section.

Not all literature is academic, however. Many police forces have published content from proposals to meeting minutes which detail the potential future of AI in criminal justice, as well as debate on ethical and legal concerns. The West Midlands Police and Crime Commissioner Ethics Committee published minutes from a 2019 meeting which detailed concerns around human rights and data protection issues and a lack of privacy impact assessments. Think tanks such as The Royal United Services Institute (RUSI) also publish work in this area, with a 2020 report entitled *'Data Analytics and Algorithms in Policing in England and Wales'* offering insights into current weaknesses in Bias, Discrimination, and a current lack of ethical oversight. Opponents of unrestricted AI use by police are also useful sources of information. Human Rights charity Liberty regularly publishes well-researched articles on this subject area.

Review of Relevant Theoretical Framework:
This project will utilise a Socio-Legal theoretical framework using a primarily empirical focus, though due to the high level of interaction with, and impact on, society it is expected that a Realism framework will also need to be involved - particularly when analysing the human element. An inductive approach will be used to collect and analyse data so that potential frameworks can be developed. A significant literature review and high degree of legal research will form the basis of the legal theory involved around AI in criminal justice. Numerical research methods such as data gathering and statistical analysis means evidence of how widespread AI Biometrics are in law enforcement can be collated, though there will need to be awareness of potential bias from police areas with differing priorities.

Research Questions:
There are a number of questions to be asked and hopefully answered by this research. As a starting point these are:

1. To what extent should Biometric AI be trusted as a predictor of criminal behaviour, an instigator for investigation, or as evidence in trial?

2. Is Biometric AI evidence alone enough to warrant prosecution of a suspected offender?

3. Is there a difference between how Magistrates as lay judiciary and Judges as professional judiciary interpret and rely on Biometric AI evidence?

4. What kind of frameworks could be developed to mitigate the drawbacks and expand the benefits of Biometric AI in policing and criminal justice?

Many of these questions are likely to change in wording during the course of the project, but their focus - on AI's value or, perhaps more importantly, risk as a policing or evidential tool - will remain throughout the project.

Research Objectives:

1. Collate and publish up to date information on the usage, accuracy, statistical validity, and evidential validity of Biometric AI in policing and justice

2. Investigate the socio-legal benefits, risks, and consequences of Biometric AI's increased usage within policing and sentencing

3. Propose possible policing and justice frameworks which would mitigate issues identified in the project

Research Methods:
One of the major research methods will be the use of empirical research to collate the professional opinions of multiple stakeholders in the Biometric AI pipeline. This includes but is not limited to AI developers and researchers, police decision-makers, police officers, national security personnel, criminal lawyers, human rights campaigners, magistrates, and judges. Though the focus would be on the reliability, fairness and legality of the current and future Biometric AI landscape, their opinions would be valuable outside of this as background detail as to the human element of Biometric AI. No AI exists in a vacuum, especially one involved in society and justice, and therefore even the inter-department politics such as human rights groups being against AI for non-technological reasons may still hold research value. My goal would be to access 5 of each type of stakeholder for both a questionnaire and then a semi-flexible 15-minute interview aiming to expand on some of the reasons given as well as potentially unearthing reasoning or factors that I had not considered. By doing this I would accrue both qualitative and quantitative data. Furthermore, a significant literature review of both the criminal law perspective and the STEM perspective would need to be undertaken and used as a base for the research. There is no easy way to tell how many police forces are experimenting with Biometric AI, but a Freedom of Information request to each force should yield some results. Previous experience has shown that most forces respond quickly and in good detail to these requests.

Research Methodology:
For the empirical research, I will reach out to various agencies and individuals who are stakeholders in the subject of this project. This includes members of the justice system such as magistrates, police forces, national security interests, lawyers, and also campaigners and NGOs such as Liberty. Individuals with professional interest in this subject, or who are authorised by their organisation to speak on record on the topic, can then be interviewed about some of the issues identified. This not only helps framework development, but will also highlight issues I or other authors may have missed. The interview could take the form of a survey for statistical purposes, as well as a more open-answer session to help with additional background and research information.

Ethical Considerations:
It is likely that this project will involve analysis of criminal trials both concluded and ongoing, and so it will be important to make sure any ongoing issues are treated carefully in regards to defamation or jeopardizing trials. It is possible I may have to anonymize or change the names of some of the people I interview. I would have to be very careful with how this data is handled.

Potential Contribution to Knowledge:

AI within Law Enforcement is not new, however it is still in its infancy. This seemingly oxymoronic statement results from the exponential nature of AI, in that we are seeing the same technological progress each month or so now that we did each year a mere decade ago. It is not an unfair or hyperbolic statement to say that in a few years time Police, Courts, and Justice will have access to AI powers and abilities that would have been considered science fiction at the start of the century. For such technology to be used fairly, appropriately, and in line with our duties to the public as a Democracy, it is vital that there are sufficient frameworks in place to ensure that each and every individual who is a suspect of such technology receives a fair trial and does not have their right to privacy violated unnecessarily - only when based on solid evidence or reasonable suspicion. This PhD Research will help to ensure that the technologies being developed now and in the future are being used fairly and appropriately as would be expected in a free and fair society. It will highlight issues which may not currently be well explained or well known, and will expand current knowledge in an area of law which is currently very high in public interest, as well as important to our justice system. This research would also function as a foundation for others, both in academia and in society, to make better decisions about how these systems are deployed and regulated - as well as for those looking to expand on my findings.

Research Plan/Timetable:

Year 1: Conduct and write literature review and secure research participants

Year 2: Undertake research with participants via questionnaires and interviews

Year 3: Complete data analysis, edit previously written research, compile and distribute findings

Bibliography Note

West Midlands Police and Crime
https://www.westmidlands-pcc.gov.uk/wp-content/uploads/2019/05/Ethics-Committee-03042019-MINUTES-.pdf?x39505

Artificial Intelligence, Computational Modelling and Criminal Proceedings
https://link.springer.com/book/10.1007%2F978-3-030-52470-8#authorsandaffiliationsbook

RUSI
https://rusi.org/sites/default/files/rusi_pub_165_2020_01_algorithmic_policing_babuta_final_web_copy.pdf

'Predictive Policing and Artificial Intelligence'
https://books.google.co.uk/books?id=J-MSEAAAQBAJ&dq=%22Artificial+Intelligence%22++criminal+law&lr=&source=gbs_navlinks_s

Data Mining Crime Analysis -
https://search.proquest.com/docview/2472178831?pq-origsite=gscholar&fromopenview=true

PhD Proposal – The University as a Platform: Co-creating Community Infrastructure (RDF23/CIS/DOW)

Applicant – David Clark

Proposed Title - Exploring the role of the university as a broker in citizen-led encounters with AI-driven surveillance technologies

Overview

The increasing use of AI-powered tools has brought forth new opportunities in various aspects of modern society. However, it has also raised concerns from different voices, predicting potential ramifications for entire professions, civil liberties and ways of life. For instance, a study by Buolamwini and Gebru (2018) has highlighted the ethical implications of AI automation, specifically addressing fairness and potential discrimination based on race, nationality or gender. As such, those most adversely affected belong to marginalised communities often living in identified areas of multiple deprivation and facing material threats to their standard of living such as the worst effects of poverty. Additionally, Fontes et al (2022) conducted research on the deployment of Facial Recognition Technology (FRT) and assessed its impacts in terms of function, consent and societal implications. These studies underscore the importance of keeping AI systems in the best interests of the public and promoting awareness and education among marginalised groups to ensure their inclusion in the ongoing discussions surrounding AI ethics. It is not clear, however, on whom this responsibility falls and with commercial interests enjoying increasing successes with their AI-driven tools (e.g. OpenAI's ChatGPT LLM[1]) it seems unlikely they can be relied upon to prioritise the inclusion of marginal voices over the expansion of profit margins. Rather, there is an urgency to explore the potential for other organisations already working with communities, for example, HEIs, to broker the investigation and understanding of AI tools in collaboration with citizen partners.

Objectives

The objectives of this PhD will be explored by carrying out a case study approach. The case studies are outlined below:

Case Study One – *Researching the problem with communities*

How the implementation of AI in camera surveillance impacts the perception of surveillance in communities experiencing multiple deprivations. This case study investigates how an individual's level of digital literacy affects their attitudes and experiences with AI surveillance technologies.

Case Study Two – *Creating encounters with AI community surveillance*

Designing a platform to create citizen encounters with AI surveillance technologies, building awareness of AI-driven technology for citizen community safety. This study will allow participants to contribute to the design, for example through the use of Design Fiction (Blythe et al., 2016), and interact with the platform educating and informing them of interactions with surveillance technology and exploring data implications.

Case Study Three – *Supporting communities to research AI encounters*

Creating and evaluating a trustworthy process which empowers communities to create new encounters with AI-driven surveillance technologies for themselves. In this case study a framework will be developed that evolves the role of AI research expertise in such processes and transforms the relationship between citizens and the university (Olivier & Wright, 2015). This framework will be evaluated to explore its efficacy in comprising one facet of the university

[1] https://openai.com/

as a platform, specifically supporting citizen-led research into the future of AI tools that impact citizen safety.

Rationale

AI encroachment, especially in the realm of surveillance, raises significant concerns that will impact society as a whole. Marginalised groups, often excluded from the benefits of AI, face increased exposure to these technologies due to limited access and knowledge. Biases within AI classification systems, such as predicting criminality based on facial features (Wu and Zhang, 2016) or detecting sexuality from face images (Wang and Kosinski, 2018), perpetuate patterns of racial and gender bias, deepening inequalities within marginalised communities. Urgent revaluation of these systems is needed (West et al., 2019) to address these issues raise awareness around the potential impact of AI systems and ensure those impacted are aware, educated and equitably represented in their design (Hayes, 2011).

Conducting research on the role of the university as a broker in trustworthy AI is crucial due to its potential impact on different groups, especially when considering marginalised communities that are often voiceless in such processes. This investigation is timely, since modern, civic universities are increasingly tasked with finding new ways to work with the communities in which they are situated (Goddard et al., 2016) and are measured against the UKRI's Knowledge Exchange Framework[2] to assess their success in this regard. This speaks the universities research growth and the research work done at a regional level. Moreover, universities can serve as brokers, fostering collaboration, dialogue and co-creation between communities, local government and academic researchers (Dow et al., 2018). Concurrently, recent research in HCI demonstrates how citizen engagement with their data can impact trust, inclusivity and individual agency (Bowyer et al., 2022). This represents an upskilling of local government and community members that is crucial in enabling marginalised groups to understand and utilise their data effectively. By adopting an inclusive design method and engaging in action research, universities can empower marginalised communities to shape the AI research agenda and develop technologies that better serve their needs and help them understand the technologies better. The collaborative effort allows for the equitable inclusion of diverse perspectives, whilst ensuring AI systems are designed ethically, equitably and with community understanding that both empowers and educates.

The outputs of this research will contribute new knowledge to top tier, international venues in Human-Computer Interaction, such as the CHI Conference, whilst providing insight impacting the ethical design of AI-driven tools more broadly. Importantly, this project seeks to provide a practical framework that affords new citizen-led research opportunities that are produced collaboratively between communities and Higher Education Institutions.

Methodology

This research will take a case study approach which allows for significant interaction with research participants providing an in-depth picture (Bloomberg & Volpe, 2019). Additionally providing an opportunity to investigate complex phenomena with a high level of detail and depth (Yin, R.K 2009). To ensure for a comprehensive analysis of a rich and diverse data set involving community partners, the project will employ a mixed methods approach, integrating quantitative and qualitative data collection and analysis techniques as both combined provide considerable value and depth (Bryman, 2006). Furthermore, the research will adopt an Action Research (AR) framework to emphasise the collaboration between research, community members and stakeholders promoting mutual learning and shared decision-making. AR use in Human-Computer Interaction (HCI) research offers methodological approaches when conducting collaborative research (Hayes, 2011). Finally, participatory design principles will be incorporated into the research methodology, as described by Björgvinsson et al. (2010). Emphasising the active involvement of community members in the research.

[2] https://kef.ac.uk/

The methodologies aim to generate a comprehensive understanding of the role of the university as a broker in trustworthy AI.

Skills Development

- AI tools discussion and evaluation.
- Mixed methods research methods.
- Academic Writing.
- Understanding of ethical and responsible practices.

Personal Motivation

My personal motivations for this project come from my passion around the potential AI brings and its ability to positively impact society. However, I firmly believe that this technology must be developed and used ethically and fairly, working for people rather than against them. I enjoy collaborating with others who share this passion and are committed to leveraging this technology for good. My experience and passion for this can be seen in my undergraduate project looking at attitudes towards the use of AI in surveillance, titled 'AI Surveillance: Big Brother's Nerdy Cousin? Investigating Public Attitudes on the Use of AI in Surveillance Technology'. This opportunity to contribute to the Ph.D. position would allow me to further explore these important issues whilst evaluating the university as a platform, working with local organisations to leverage the universities position locally.

Timeline

A provisional timeline for the project, including the three case studies and the writing up and publication of the findings of these can be found below. Subject to adjustments based on the specific requirements of each task and the progress made during the research process. To be in line with conference submissions.

Timeframe		Anticipated Activities
Year One	Months 1-6	- Finding relevant literature and preparing for literature review. - Identify local groups (charities, organisations, councils) to work with and start to build relationships.
	Months 7-8	- Composing literature review. - Maintain relationship with local groups. - Design fact finding questionnaires and interviews.
	Months 9-12	- Carry out fact finding study. - Maintain relationship with local groups. - Analysis of fact-finding data - Compose related thesis chapters
Year Two	Months 1-6	- Commence case study 1 (problem research) - Design and carry out community-based research activity - Write-up case study findings for conference submission.
	Months 6-12	- Commence case study 2 (event/interaction type) - Write-up case study findings for conference submission. - Compose related thesis chapters
Year Three	Months 1-9	- Commence case study 3 (framework design/evaluation) - Build design framework and evaluate for AI encounters in the future and understandable/explainable AI (XAI). - Analyse and write-up for thesis.
	Months 10-12	- Compose Introduction & Conclusion chapters - Final thesis assembly - Viva Preparation &Delivery

References

Björgvinsson, E., Ehn, P. and Hillgren, P.A., 2010, November. Participatory design and" democratizing innovation". In Proceedings of the 11th Biennial participatory design conference (pp. 41-50).

Bloomberg, L.D. and Volpe, M., 2018. Completing your qualitative dissertation: A road map from beginning to end.

Blythe, M., Andersen, K., Clarke, R. and Wright, P., 2016, May. Anti-solutionist strategies: Seriously silly design fiction. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (pp. 4968-4978).

Bowyer, A., Holt, J., Go Jefferies, J., Wilson, R., Kirk, D. and David Smeddinck, J., 2022, April. Human-GDPR Interaction: Practical Experiences of Accessing Personal Data. In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (pp. 1-19).

Bryman, A., 2006. Integrating quantitative and qualitative research: how is it done?. Qualitative research, 6(1), pp.97-113

Bryman, A., 2016. Social research methods. Oxford university press.

Buolamwini, J., Gebru, T. (2018) 'Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification', Proceedings of Machine Learning Research, 81, p1-15.

Dow, Andy, Rob Comber, and John Vines. "Between grassroots and the hierarchy: Lessons learned from the design of a public services directory." In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, pp. 1-13. 2018.

Fontes, A.C. et al. (2022) "AI-powered public surveillance systems: Why we (might) need them and how we want them," SSRN Electronic Journal [Preprint]. Available at: https://doi.org/10.2139/ssrn.4112042.

Goddard, J., Hazelkorn, E. and Vallance, P. eds., 2016. The civic university: The policy and leadership challenges. Edward Elgar Publishing.

Hayes, G.R., 2011. The relationship of action research to human-computer interaction. ACM Transactions on Computer-Human Interaction (TOCHI), 18(3), pp.1-20.

Olivier, P., and Wright, P., 2015. Digital Civics: Taking a Local Turn (Interactions)

Wang, Y. and Kosinski, M., 2018. Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. Journal of personality and social psychology, 114(2), p.246.

West, S.M., Whittaker, M. and Crawford, K., 2019. Discriminating systems. AI Now, pp.1-33.

Wu, X. and Zhang, X., 2016. Automated inference on criminality using face images. arXiv preprint arXiv:1611.04135, pp.4038-4052.

Yin, R.K., 2009. Case study research: Design and methods (Vol. 5). Sage

PhD Proposal	
Advert Reference	RDF21/EE/CIS/WARNERMark
Applicant	Lauren Scott

Exploring the relationship between family and health misinformation (HCI/CSCW).

Overview

Health misinformation has existed for many years and impacts a large proportion of the world's population on a day-to-day basis. A famous example of which is the now-retracted study and subsequent statement by Andrew Wakefield (Wakefield *et al.*, 1998; Wakefield, 2004; The Editors of The Lancet, 2010), ignited by the misinformation of his MMR-Autism study. The media coverage of Wakefield's "badly designed study" (Walker, 1998), has led to an anti-vaccination movement which still resonates today. With the spread of the Covid-19 virus and the race to create a vaccine, there have been a number of *anti-vaxxer*s insisting that vaccinations are not safe (ITV News, 2020). However, misinformation does not only surround vaccinations. It has been seen to impact public's opinions of the 2015 Zika virus (Bode and Vraga, 2018), oral contraceptive use in Egypt (DeClerque *et al.*, 1986) and has seen the upsurge of fake cancer "treatments" spread via twitter (Ghenai and Mejova, 2018), to give a few examples. Overall, articles misinforming members of the public of health facts, have seen a significant increase from 2016 onwards (Wang *et al.*, 2019).

This study proposes to both investigate the methods through which this misinformation is spread and to determine the effects that this misinformation can have on family dynamics and relationships.

Objectives

- Explore how misinformation is flagged, understood, evaluated, shared and discussed within family networks across different cultural and/or socio-economic groups,
- Identify strategies and behaviours used by family members across different cultural and/or socio-economic groups when challenging shared misinformation,
- Determine the effects that misinformation can have on family relationships and dynamics.

Rationale

Although associated with social media, health misinformation and "fake news" has been seen to spread between family members over WhatsApp (Gragnani, 2018). Although interventions have been investigated for use on social media (Walter *et al.*, 2020), further research is needed into how misinformation spreads across other medium and how misinformation can impact families of different cultural and socio-economic backgrounds. As health misinformation impacts individuals across the world, it is important to understand how misinformation impacts multiple cultures, and backgrounds.

This investigation is key to limiting the spread of misinformation between family networks as it is clear that more understanding is needed for the development of interventions for messaging platforms, and for the education of individuals on how misinformation is presented and spread. This education will be essential to limiting misinformation in future generations.

Methodology

The proposed methodology for this study includes three studies, across multiple nationalities. The first stage will be a questionnaire produced in multiple languages. This questionnaire will be used as an initial data gathering exercise to explore widespread opinions on misinformation, its presence in their family network, whether participants can identify misinformation, and what effects that can have on family relationships.

This will be used to inform the second study through the design of more specific questions and case studies. The second study will involve focus groups (held virtually due to geographical distance) and will investigate each participants' experiences with misinformation.

The results of these studies will be used to produce educational resources on misinformation, suitable for both e-safety lessons in education and for social media

providers to use as advertisements. Additionally, they will be used to further specify sources of misinformation, languages used and how family.

The impact of these resources will be explored in relation to family networks and the impact that they have on younger generations' spotting misinformation will be included. This is due to the rise of misinformation, and although schools focus on e-safety, *fake news* has only begun to be included in teaching, with limited resources. By educating individuals on misinformation, it may serve to limit the spread

Skill Development

- Research Methods
- Academic Writing
- Software Development
- User Research
- Computer Support Cooperative Work methods

Challenges

The primary challenge related to this project is that of identifying individuals who have been affected by misinformation. Although those who have misinformed family members will be aware that the misinformation has occurred, finding participants that have been misinformed will be more difficult as they may not be aware that this misinformation has occurred.

The secondary challenge specific to this project is the language barrier, and access to participants from different socio-economic backgrounds. When conducting a study investigating a number of different nationalities and cultures, identifying the presence of a language barrier will be important. Additionally, when conducting these focus groups virtually, it is important that this does not exclude the possibility for participants from less affluent socio-economic backgrounds. As one of the objectives is to explore these issues across different socio-economic backgrounds, it is important to ensure that all backgrounds can take part.

Personal Motivations

My motivations for this project come from two main areas: my interest in HCI research, and my experiences of friends and family believing misinformation during the Covid-19 pandemic.

I became interested in HCI during my summer industry placement in Agile Development. Since then, I have undertaken two university modules on the topic, obtaining a 100% score in the MSc-level module and have prepared a CHI paper related to our group's findings (the status of this paper is still unknown as it was late-breaking work). This was with the support of the NORTH Lab research group, based at the university, whose virtual talks I attended during the first lockdown. As HCI is an area I hope to work in one day, this project seemed very suitable.

Regarding misinformation, over the course of the Covid-19 pandemic I have observed educated individuals turn to social media and conspiracy theories, believing misinformation above the facts and then trying to spread that throughout family and friends. As part of my teaching placement, I have also been teaching secondary school students about the importance of identifying *fake news* and the identification of facts.

Timeline

- First year
    - Months 1-6: finding relevant literature and preparing for literature review
    - Months 7 & 8: composing literature review
    - Months 9-12: study design and advertisement. Participant groups identified.
- Second year
    - Months 13 & 14: preliminary study undertaken across multiple nationalities.
    - Month 15: data analysis and evaluation of preliminary study
    - Months 16 & 17: secondary study undertaken with (virtual) focus groups from multiple nationalities.
    - Month 18: data analysis and evaluation of secondary study.
    - Months 19-22: composing experiment sections(s) of thesis
    - Months 23 & 24: Ascertain conclusions from results and prepare for conference submission(s)
- Third year
    - Months 24-26: Design of educational materials to highlight misinformation methods and the effects that they can have.
    - Months 27 & 28: online study investigating the impact of these materials.

- Months 29-end: Final thesis assembly, viva preparation, viva delivery.

References

Bode, L. and Vraga, E. K. (2018) 'See Something, Say Something: Correction of Global Health Misinformation on Social Media', *Health Communication*, 33(9), pp. 1131–1140. doi: 10.1080/10410236.2017.1331312.

DeClerque, J. *et al.* (1986) 'Rumor, Misinformation and Oral Contraceptive Use in Egypt', *Social science & medicine (1982)*, 23, pp. 83–92. doi: 10.1016/0277-9536(86)90327-8.

Ghenai, A. and Mejova, Y. (2018) 'Fake Cures: User-centric Modeling of Health Misinformation in Social Media', *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW), p. 58:1-58:20. doi: 10.1145/3274327.

Gragnani, J. (2018) 'Pesquisa inédita identifica grupos de família como principal vetor de notícias falsas no WhatsApp', *BBC News Brasil*. Available at: https://www.bbc.com/portuguese/brasil-43797257 (Accessed: 3 January 2021).

ITV News (2020) *Revealed: Anti-vaxxers' plans to undermine trust in Covid-19 vaccine*, *ITV News*. Available at: https://www.itv.com/news/2020-12-22/revealed-more-than-5-million-anti-vax-followers-in-uk-as-leaders-spread-distrust-about-covid-vaccine (Accessed: 3 January 2021).

The Editors of The Lancet (2010) 'Retraction—Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children', *The Lancet*, 375(9713), p. 445. doi: 10.1016/S0140-6736(10)60175-4.

Wakefield, A. (2004) 'A statement by Dr Andrew Wakefield', *The Lancet*, 363(9411), pp. 823–824. doi: 10.1016/S0140-6736(04)15710-3.

Wakefield, A. J. *et al.* (1998) 'RETRACTED: Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children', *The Lancet*, 351(9103), pp. 637–641. doi: 10.1016/S0140-6736(97)11096-0.

Walker, D. R. (1998) 'Autism, inflammatory bowel disease, and MMR vaccine', *The Lancet*, 351(9112), p. 1355. doi: 10.1016/S0140-6736(98)26018-1.

Walter, N. *et al.* (2020) 'Evaluating the Impact of Attempts to Correct Health Misinformation on Social Media: A Meta-Analysis', *Health Communication*, pp. 1–9. doi: 10.1080/10410236.2020.1794553.

Wang, Y. *et al.* (2019) 'Systematic Literature Review on the Spread of Health-related Misinformation on Social Media', *Social Science & Medicine*, 240, p. 112552. doi: 10.1016/j.socscimed.2019.112552.