APPLICATION FOR ETHICAL APPROVAL

In order for research to result in benefit and minimise risk of harm, it must be conducted ethically.

The University follows the OECD Frascati manual definition of **research activity**: "creative work undertaken on a systematic basis in order to increase the stock of knowledge, including knowledge of man, culture and society, and the use of this stock of knowledge to devise new applications". As such this covers activities undertaken by members of staff, postgraduate research students, and both taught postgraduate and undergraduate students working on dissertations/projects.

The individual undertaking the research activity is known as the "principal researcher".

This form must be completed and approved prior to undertaking any research activity.

SECTION A: About You (Principal Researcher)

1	Full Name:	Habibullah Shaikh
2	Student Number:	2503799
3	Email address:	2503799@student.uwtsd.ac.uk
4	Programme of Study:	MSc Cybersecurity and Digital Forensics
5	Director of Studies/Supervisor:	Suresh Shanmugasundaram

SECTION B: Internal and External Ethical Guidance Materials

	Please list the core ethical guidance documents that have been referred to during the completion of this form (including any discipline-specific codes of research ethics, location-specific dodes of research ethics, and also any specific ethical guidance relating to the proposed methodology). Please tick to confirm that your research proposal adheres to these codes and guidelines. You may add rows to this table if needed.	
1	**UWTSD Research Ethics & Integrity Code of Practice**	☒
2	**UWTSD Research Data Management Policy**	☒
3	UWTSD Data Protection Policy (GDPR/DPA 2018 compliance)	☒

SECTION C: Details of Research Activity

1	Indicative title:	AI-Powered Threat Detection in SIEM		
2	Proposed start date:	June 2026	Proposed end date:	September 2026
	Introduction to the Research (maximum 300 words in each section) **Ensure that you write for a <u>Non-Specialist Audience</u> when outlining your response to the three points below:** • *Purpose of Research Activity* • *Proposed Research Question* • *Aims of Research Activity* • *Objectives of Research Activity* Demonstrate, briefly, how **<u>Existing Research</u>** has informed the proposed activity and explain			

	• *What the research activity will add to the body of knowledge* • *How it addresses an area of importance.*
3	**Purpose of Research Activity** This research investigates whether machine learning can improve the detection of brute force cyber-attacks within Security Information and Event Management (SIEM) systems compared to traditional rule-based detection methods. Rule-based SIEM systems currently generate very high numbers of false alerts, wasting security analysts' time and causing real attacks to be missed. This project designs, builds, and tests an AI-based detection system using the open-source Wazuh SIEM platform, evaluates six machine learning algorithms, and compares performance against a conventional rule-based baseline using standard, non-personal, publicly available network traffic data
4	**Research Question** Can a machine learning model, trained on publicly available network traffic data, detect brute force attacks within a SIEM environment with higher precision and fewer false positives than conventional rule-based detection methods?
5	**Aims of Research Activity** To design, develop, and evaluate an AI-powered brute force attack detection system integrated with the Wazuh open-source SIEM platform, achieving measurably higher detection accuracy and fewer false positives than conventional rule-based detection.
6	**Objectives of Research Activity** 1) To review existing literature on machine learning-based threat detection in SIEM systems. 2) To develop and compare multiple machine learning classifiers trained on publicly available network traffic datasets and independently captured network traffic, integrating the best performing model with the Wazuh SIEM platform. 3) To evaluate detection performance using standard classification metrics and critically assess the results against a rule-based baseline.
	Proposed data collection methods (maximum 600 words) Provide a brief summary of all the methods that **may** be used in the research activity to collect data, making it clear what specific techniques may be used. If methods other than those listed in this section are deemed appropriate later, additional ethical approval for those methods will be needed. You do not need to justify the methods here, but should instead describe how you intend to collect the data necessary for you to complete your project.
7	This should describe how you intend to collect data. It should not include a discussion of the theoretical basis for your data collection methods. Please note, that if you intend to collect any audio/video recordings of interviews with participants then these will be are classified as Personal Data under GDPR/DPA2018. If you intend to use these then note this in section H. This research does not involve any human participants or personal data. All data used is either (a) the publicly available CICIDS 2017 network intrusion detection dataset, produced by the Canadian Institute for Cybersecurity specifically for academic research and containing no personally identifiable information, or (b) network traffic independently captured by the researcher using Wireshark on their own personal machine in an isolated local test environment, involving only the researcher's own simulated traffic and no third-party data. No live organisational networks, production systems, or third-party users are involved at any stage.

SECTION D: Scope of Research Activity

	Will the research activity include:	**YES**	**NO**
1	Use of a questionnaire or similar research instrument?	☐	☒
2	Use of interviews?	☐	☒
3	Use of focus groups?	☐	☒
4	Use of participant diaries?	☐	☒
5	Use of video or audio recording?	☐	☒
6	Use of computer-generated log files?	☐	☒
7	Participant observation with their knowledge?	☐	☒
8	Participant observation without their knowledge?	☐	☒

9	Access to personal or confidential information without the participants' specific consent?	☐	☒
10	Administration of any questions, test stimuli, presentation that may be experienced as physically, mentally or emotionally harmful / offensive?	☐	☒
11	Performance of any acts which may cause embarrassment or affect self-esteem?	☐	☒
12	Investigation of participants involved in illegal activities?	☐	☒
13	Use of procedures that involve deception?	☐	☒
14	Administration of any substance, agent or placebo?	☐	☒
15	Working with live vertebrate animals?	☐	☒
16	Procedures that may have a negative impact on the environment?	☐	☒
17	Other primary data collection methods. Please indicate the type of data collection method(s) below.	☐	☒
	Details of any other primary data collection method: (this box should expand as you type)		

If you have ticked NO to every question then the research activity is (ethically) low risk and you may skip section E and continue to section F.

If YES to any question, then no research activity should be undertaken until full ethical approval has been obtained.

SECTION E: Intended Participants

	Who are the intended participants:	**YES**	**NO**
1	Students or staff at the University?	☐	☐
2	Adults (over the age of 18 and competent to give consent)?	☐	☐
3	Vulnerable adults?	☐	☐
4	Children and Young People under the age of 18? (Consent from Parent, Carer or Guardian will be required)	☐	☐
5	Prisoners?	☐	☐
6	Young offenders?	☐	☐
7	Those who could be considered to have a particularly dependent relationship with the investigator or a gatekeeper?	☐	☐
8	People engaged in illegal activities?	☐	☐
9	Others. Please indicate the participants below, and specifically any group who may be unable to give consent.	☐	☐
	Details of any other participant groups: Complete this only if your participants cannot give consent. This includes animals Not applicable — this research does not involve any human participants (see Section D).		

	Participant numbers and source Provide an estimate of the expected number of participants. How will you identify participants and how will they be recruited?	
10	How many participants are expected?	Ballpark figures are fine, but make sure that you explain how you will identify and contact your participants. N/A — no human participants are involved in this research

11	Who will the participants be?	N/A
12	How will you identify the participants?	N/A

	Information for participants:	**YES**	**NO**	**N/A**
13	Will you describe the main research procedures to participants in advance, so that they are informed about what to expect?	☐	☐	☐
14	Will you tell participants that their participation is voluntary?	☐	☐	☐
15	Will you obtain written consent for participation?	☐	☐	☐
16	Will you explain to participants that refusal to participate in the research will not affect their treatment or education (if relevant)?	☐	☐	☐
17	If the research is observational, will you ask participants for their consent to being observed?	☐	☐	☐
18	Will you tell participants that they may withdraw from the research at any time and for any reason?	☐	☐	☐
19	With questionnaires, will you give participants the option of omitting questions they do not want to answer?	☐	☐	☐
20	Will you tell participants that their data will be treated with full confidentiality and that, if published, it will not be identifiable as theirs?	☐	☐	☐
21	Will you debrief participants at the end of their participation, in a way appropriate to the type of research undertaken?	☐	☐	☐
22	If NO to any of above questions, please give an explanation			
	You should be able to tick YES for all of these questions. If not, then explain why not in this box. Not applicable — no human participants are involved in this research (see Section D).			

	Information for participants:	**YES**	**NO**	**N/A**
24	Will participants be paid?	☐	☐	☐
25	Is specialist electrical or other equipment to be used with participants?	☐	☐	☐
26	Are there any financial or other interests to the investigator or University arising from this study?	☐	☐	☐
27	Will the research activity involve deliberately misleading participants in any way, or the partial or full concealment of the specific study aims?	☐	☐	☐
28	If YES to any question, please provide full details			
	You should be able to tick NO for most of these questions. For any cases that you have ticked YES then provide details in this box. If you are using cameras/voice recorders to record interviews then please state that in this box. Not applicable — no human participants are involved in this research.			

SECTION F: Anticipated Risks

	Outline any anticipated risks that may adversely affect any of the participants, the researchers and/or the University, and the steps that will be taken to address them.	
1	**Risks to participants** For example: sector-specific health & safety, emotional distress, financial disclosure, physical harm, transfer of personal data, sensitive organisational information. If you have identified in section D that there are no participants then enter N/A and go skip to question 3.	
	Risk to participants:	*How you will mitigate the risk to participants:*

	There are always risks. Do not write N/A unless you have no participants. N/A — no human participants are involved in this research.	*(this box should expand as you type)*
2	If research activity may include sensitive, embarrassing or upsetting topics (e.g. sexual activity, drug use) or issues likely to disclose information requiring further action (e.g. criminal activity), give details of the procedures to deal with these issues, including any support/advice (e.g. helpline numbers) to be offered to participants. Note that where applicable, consent procedures should make it clear that if something potentially or actually illegal is discovered in the course of a project, it may need to be disclosed to the proper authorities	
	Not applicable — the research activity does not involve sensitive, embarrassing, or upsetting topics, and does not involve human participants.	
3	**Risks to the investigator** For example: personal health & safety, physical harm, emotional distress, risk of accusation of harm/impropriety, conflict of interest	
	Risk to the investigator: There are always risks. Do not write NA. Minor risk of time pressure due to project deadlines. No physical, emotional, or health and safety risks are anticipated, as all work is conducted using a personal laptop and publicly available data in a controlled, isolated software environment.	*How you will mitigate the risk to the investigator:* Time risk mitigated through a structured project plan and regular supervision meetings with the Director of Studies.
4	**University/institutional risks** For example: adverse publicity, financial loss, data protection	
	Risk to the University: There are always risks. Do not write NA. Minimal institutional risk. No use of University systems for data storage beyond standard student cloud storage (OneDrive), and no processing of personal or sensitive organisational data.	*How you will mitigate the risk to the University:* All research materials will be stored in the student's University-provided OneDrive account with restricted access.
5	**Environmental risks** For example: accidental spillage of pollutants, damage to local ecosystems	
	Risk to the environment: You may write NA if there are no research-related environmental risks. Driving to the university does not count as a risk. N/A — this research involves no physical or environmental activity	*How you will mitigate the risk to environment:* N/A

SECTION G: Feedback, Consent and Confidentiality

If you have identified in section D that there are no participants then enter skip this section and continue to section H.

1	**Feedback** What de-briefing and feedback will be provided to participants, how will this be done and when?
	You don't need to email your participants with your final report. A good alternative is to set up an email address that they will be able to contact for further details or results. Not applicable — this research does not involve any human participants.
2	**Informed consent** Describe the arrangements to inform potential participants, before providing consent, of what is involved in participating. Describe the arrangements for participants to provide full consent before data collection begins. If gaining consent in this way is inappropriate, explain how consent will be obtained and recorded in accordance with prevailing data protection legislation.
	If you are using a paper questionnaire then you should have the participants sign an appropriate consent form. These forms will count as personal data and should be noted as such in section J.

	If you are using an online questionnaire, then you should have a screen before the questions start that acts as a consent form, informing participants that by clicking on the NEXT button they are providing consent. Not applicable — no participants are involved; therefore no consent procedures are required.
3	**Confidentiality / Anonymity** Set out how anonymity of participants and confidentiality will be ensured in any outputs. If anonymity is not being offered, explain why this is the case.
	Do not collect names unless you really need them. Do not name participants or organisations in any research publications (including the thesis) without their explicit permission. Not applicable — no participants are involved. No personal or identifiable data is collected at any stage.

SECTION H: Data Protection and Storage

	Does the research activity involve personal data (as defined by the General Data Protection Regulation 2016 "GDPR" and the Data Protection Act 2018 "DPA")?	**YES**	**NO**
1	***"Personal data"*** *means any information relating to an identified or identifiable natural person ('data subject'). An identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person. Any video or audio recordings of participants is considered to be personal data.*	☐	☒
	If YES, provide a description of the data and explain why this data needs to be collected:		
2	This includes audio/video data of participants, but can also include IP addresses and usernames. Names, addresses and emails also count, as do consent forms. No personal data is collected. Data used is either the publicly available, fully anonymised CICIDS 2017 dataset, or network traffic independently captured by the researcher from their own personal machine, containing no third-party identifiable information.		
	Does it involve special category data (as defined by the GDPR)?	**YES**	**NO**
3	***"Special category data"*** *means sensitive personal data consisting of information as to the data subjects' –* *(a) racial or ethnic origin,* *(b) political opinions,* *(c) religious beliefs or other beliefs of a similar nature,* *(d) membership of a trade union (within the meaning of the Trade Union and Labour Relations (Consolidation) Act 1992),* *(e) physical or mental health or condition,* *(f) sexual life,* *(g) genetics,* *(h) biometric data (as used for ID purposes),*	☐	☒
	If YES, provide a description of the special category data and explain why this data needs to be collected:		
4	What counts as 'sensitive' will differ between cultures. Any information on behaviour that is not in accordance with cultural norms would count as sensitive personal data. Not applicable — no special category data is collected or processed.		

	Will data from the research activity (collected data, drafts of the thesis, or materials for publication) be stored in any of the following ways?	**YES**	**NO**
5	Manual files (i.e. in paper form)?	☐	☒
6	University computers?	☐	☒

7	Private company computers?	☐	☒
8	Home or other personal computers?	☒	☐
9	Laptop computers/ CDs/ Portable disk-drives/ memory sticks?	☒	☐
10	"Cloud" storage or websites?	☒	☐
11	Other – specify:	☐	☒
12	For all stored data, explain the measures in place to ensure the security of the data collected, data confidentiality, including details of backup procedures, password protection, encryption, anonymisation and pseudonymisation:		
	If possible, save your data on computers that are secure and regularly backed up. Many cloud services only provide GDPR-compliant storage for business customers. An example of suitable text is given below. *All data will be kept in password protected cloud storage on the University Office 365 system which will not be shared. Audio/visual data will be transcribed and would be shown to participants to check accuracy of reporting. Any USB sticks used to store or transfer data will be password protected. All participants will be given a unique identifier to ensure confidentiality and this list will be kept securely in the password protected folder.* All research data, code, and drafts will be stored in the student's password-protected University OneDrive account, accessible only to the student and the supervisor. The personal laptop used for data capture and model development is password protected. No data is shared with third parties. Since no personal data is collected, no anonymisation or pseudonymisation procedures are required.		

	Data Protection		
	Will the research activity involve any of the following activities:	**YES**	**NO**
13	Electronic transfer of data in any form?	☐	☒
14	Sharing of data with others at the University outside of the immediate research team?	☐	☒
15	Sharing of data with other organisations?	☐	☒
16	Export of data outside the UK or importing of data from outside the UK?	☐	☒
17	Use of personal addresses, postcodes, faxes, emails or telephone numbers?	☐	☒
18	Publication of data that might allow identification of individuals?	☐	☒
19	If YES to any question, please provide full details, explaining how this will be conducted in accordance with the GDPR and Data Protection Act (2018) (and/or any international equivalent):		
	This includes data such as drafts of your thesis as well as experimental or survey data. An example of suitable text is given below. *All data will be encrypted and kept in password protected cloud storage on the University Office 365 system which will not be shared. Any USB sticks used to store or transfer data will be password protected. All data transfers will be encrypted and password protected. All participants will be given a unique identifier to ensure confidentiality and this list will be kept securely in the password protected folder. The data will be stored until the completion of the project and then deleted. In accordance with the DPA2018, participants will have the right to ask to see what data is held relating to them, and this data will be deleted immediately if the participant requests this, in which case the data will not be used in the project.* Not applicable in most respects, as no personal data is involved. Research files (code, datasets, and report drafts) will be transferred only between the student's personal laptop and their password protected University OneDrive account for supervisor review.		
20	List all who will have access to the data generated by the research activity:		

	Normally the principal researcher, possibly also the supervisor and, if the project has an industrial partner, a representative of that partner. Possibly also external examiner or second marker? The principal researcher (Habibullah Shaikh) and the Director of Studies/Supervisor (Suresh Shanmugasundaram), via shared OneDrive access.
21	List who will have control of, and act as custodian(s) for, data generated by the research activity:
	Usually the principal researcher. The principal researcher, Habibullah Shaikh.
22	Give details of data storage arrangements, including security measures in place to protect the data, where data will be stored, how long for, and in what form.
	All data will be encrypted and kept in password protected cloud storage on the University Office 365 system which will not be shared. Any USB sticks used to store or transfer data will be password protected, and will be reformatted at the end of the project in order to destroy the data. The data will be stored until the completion of the project and then deleted. All project files are stored in the student's password-protected UWTSD OneDrive account. Data will be retained until the project is fully marked and the module has concluded, after which working files will be deleted. No personal data is stored at any point.

22	Confirm that you have read the UWTSD guidance on data management (see https://www.uwtsd.ac.uk/library/research-data-management/)	☒
23	Confirm that you are aware that you need to keep all data until after your research has completed or the end of your funding	☒

SECTION I: Declaration

	The information which I have provided is correct and complete to the best of my knowledge. I have attempted to identify any risks and issues related to the research activity and acknowledge my obligations and the rights of the participants. In submitting this application I hereby confirm that I undertake to ensure that the above named research activity will meet the University's Research Ethics and Integrity Code of Practice which is published on the website: https://www.uwtsd.ac.uk/research/research-ethics/		
1	**Signature of applicant:**	Habibullah Shaikh	**Date:** **9/Aug/2026**
2	Director of Studies/Supervisor:	Suresh	**Date:** 9 Aug 2026
3	Signature:		

FOR INTERNAL USE ONLY:

	Ethical approval given		
1	**Signature of assessor:**		**Date:**
2	Name:		
3	Role:		