AI Cyber Risk Awareness Code of Practice

Executive Summary

The aim of this document is to define the purpose, direction and principles for an AI Cyber Risk Awareness Code of Practice (CoP) and outline the key terms and risks associated with the use of generative AI within a business environment.

CoP Statement

1. This CoP considers the responsible use of Generative Artificial Intelligence (Gen AI).

2. Gen AI tools are having a significant impact on business and education. These tools pose considerable challenges around data and cybersecurity governance, educational assessment and academic integrity. They also present opportunities, saving staff time by helping with the streamlining of business processes, the creation of learning materials, or presenting students with new tools to enhance the way they work.

Purpose

3. The purpose of this CoP is to set out the risks of the use of Gen AI in the workplace and to ensure that the benefits of Gen AI are maximised while minimising the risks associated with use of this technology.

4. Potential risks include, but are not limited to, data protection breaches, copyright issues, the protection of confidential information, ethical considerations, and compliance with wider legal obligations.

5. The terms of this CoP should be considered when using Gen AI to carry out business activities, whether in relation to your role or otherwise. It should be understood and implemented in conjunction with other relevant University policies and procedures.

Scope and Application

6. This CoP applies to;
   - the free or paid for/contractually covered use of Public Gen AI systems developed by third parties, such as Microsoft CoPilot;

   - any Private AI or Machine Learning (ML) models or systems that Leeds Beckett University develops internally.

7. It applies wherever the University has a vested interest in protecting;
   - data that is processed and stored using Gen AI;

- outcomes/decisions that are informed by the use of Gen AI.

8. This CoP applies to the following users;
   - all members of the university community, including students, academics and staff;
   - any individuals granted access to university resources, including consultants, contractors, volunteers, casual workers, agency workers, and associates.

Standards

1. Use of the University IT Environment is subject to law and regulation, and the University seeks to ensure compliance with these via this CoP.

2. The CoP also supports adherence to the requirements contained within various acts of parliament and other legislation and relevant certification schemes, including, but not limited to;

   Computer Misuse Act
   Data Protection Act
   Freedom of Information Act
   Copyright, Designs and Patents Act
   Regulation of Investigatory Powers Act
   Human Rights Act
   Electronic Communications Act
   Digital Economy Act
   Obscene Publications Act
   Counterterrorism and Security Act
   Cyber Essentials

DEFINITIONS

9. Artificial Intelligence (AI) - Machines, digital systems and software able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages.

10. Generative AI (Gen AI) - AI capable of producing a variety of novel, human-like output types, such as images, video, music, speech, text, software code and product designs. It consists of an algorithm (or algorithms) trained on existing data to generate new, realistic output that reflects the characteristics of the training data but doesn't repeat it.

11. AI model - The algorithm or set of algorithms used to interpret, assess, and create output based on the training it has received.

12. Training data set - Existing data on which an AI model learns to produce output.

13. AI system - The infrastructure surrounding the AI model through which data is inputted and output is produced based on interpretations and decisions made by the algorithm.

14. Public AI - An AI system that a vendor makes available to any user who wants access, and that collects and uses their inputs to improve the algorithm's performance. Unlike private AI systems, public systems send data outside the organization.

15. Private AI - A proprietary AI system developed, either exclusively or in collaboration with a third-party, and used by the University for which the University has ownership or responsibility.

16. Prompt - These are the inputs or queries that a user provides to the GenAI application to receive the required output. Prompts can be used by a GenAI application to further train their outputs.

17. Large Language Model - An LLM is a type of GenAI that can generate human like text in response to a prompt. They use deep learning techniques and massive data volumes to generate a response.

18. Hallucination - LLMs can produce outputs which may initially appear believable but are in fact highly inaccurate or fabricated. This is known as an hallucination.

RESPONSIBILITIES

19. The Vice Chancellor is responsible for ensuring that all necessary resources are available to support the implementation of this CoP across the University.

20. The University Executive Committee is accountable for approving and supporting this CoP.

21. The Director of Digital Services is responsible for the operational coordination of the CoP and for ensuring appropriate policies, codes of practice, procedures, guidance, advice, and training are in place to support the CoP's implementation and compliance, and for ensuring these artifacts are regularly reviewed at appropriate intervals.

22. The Head of Cyber Security is responsible for developing and maintaining the University's cyber security programme, coordinating risk assessments and incident response, and reporting on security posture to senior leadership.

23. Digital Services is responsible for implementing technical controls, monitoring systems, supporting secure infrastructure, and enabling disaster recovery.

24. Hardware and software purchases will go through the institutionally approved procurement processes, but will additionally be subject to the approval of the Director of Digital Services (or nominee).

25. The Information Governance Manager is responsible for implementing non-cyber security related information protections, such as records management and compliance with information law.

26. The University Registrar and Secretary has ultimate responsibility for business continuity.

27. Data Owners are responsible for ensuring appropriate classification and protection of data under their control.

28. All Users are responsible for complying with this CoP, and other relevant policies, Codes of Practice and procedures protecting university assets, reporting suspected security incidents or breaches, and completing mandatory security awareness training.

Consequences and Sanctions

29. The University may suspend or withdraw systems access to users who breach this CoP without notice to them and take further action as appropriate.

30. Breaches of this CoP by University staff or students may result in disciplinary procedures, up to and including dismissal.

31. Contractors and associates may also face sanctions such as termination of contracts or reporting to professional regulators.

32. The University reserves the right to report criminal offences to the appropriate authorities and pursue legal remedies.

OPERATIONAL APPLICATION

Responsibility Principles

All organisational use of Gen AI must consider the responsibility principles defined below:

33. **Privacy:** Individual privacy should be respected.

34. **Fairness and Bias Detection:** Unbiased data should be used to produce fair outputs.

35. **Explainability and Transparency:** Decisions or predictions based on the output of Gen AI should be explainable and the result of a transparent process.

36. **Safety and Security:** AI systems should be secure, safe to use, and robust.

37. **Validity and Reliability:** Plans should be made to monitor relevant data, models and systems.

38. **Accountability:** A person or organisation needs to take responsibility for any decisions that are made based on the model.

Cybersecurity Principles

Data Confidentiality

39. All existing data confidentiality controls and best practices must be in place and observed when using AI as part of a business process.

40. Data used to train private AI models should be classified as confidential and protected to secure against data exfiltration by a bad actor.

41. Privacy regulations, University policies and processes must be followed when entering data into an AI system or service, especially in cases involving public AI.

42. All suspected or confirmed cases of compromised data confidentiality must be reported via the University incident reporting process as soon as possible.

43. Information Asset Owners to give formal approval before sensitive University data is used in an AI system or service.

44. Knowledge and specific details about how a private AI model has been trained and how it works should be kept strictly confidential, with access to such information being granted on a need-to-know basis.

Data Integrity

45. Output from AI systems should be verified to meet quality standards before being incorporated into organizational data repositories to avoid degrading data integrity with erroneous or otherwise low-quality inputs.

46. AI-generated output should be labeled as such so it can be quickly located if associated data sets must be reviewed, corrected, adjusted, recalled, etc.

47. Private AI system data should be audited regularly to ensure it has not been tampered with and continues to meet relevant data-integrity standards.

Data Availability

48. Private AI models and training data should be backed up at appropriate, regular intervals.

49. Recovery time objectives (RTOs) and Recovery point objectives (RPOs) should also be tested regularly.

IT Controls

50. User access monitoring should be in place for relevant AI systems, models, and training data.

51. Appropriate data access controls, including Multi-Factor Authentication (MFA), should be in place when signing into AI systems and services or accessing Private AI models and training data.

52. All Confidential or Internal data used in conjunction with AI systems or services is to be protected appropriately according to its data classification, as described in the University's Information Classification Policy.

53. Encryption key management best practices are to be followed, in line with the University's Cryptography and Encryption Code of Practice**https://www.leedsbeckett.ac.uk/staffsite/services/it-services/it-help/it-security/it-security-policies/**.

54. Appropriate technical controls, such as Intrusion Prevention Systems, should be considered for Private AI models, systems, and training data repositories in line with existing University policies and procedures.

55. Public AI systems and services should be protected in line with university expectations of any 3rd party provided service.

56. AI-generated code should not be incorporated into any of the University's systems without proper authorization and integrity assessment.

Procedural Controls

57. Employees may use Gen AI for business processes (such as research, data analysis, and communications) if organisational standards to protect data confidentiality and integrity, as laid out in this CoP and elsewhere, are upheld.

58. Employees should seek advice and guidance from Information Asset Owners, Governance or Digital Services before using or entering sensitive university data into Public AI systems.

59. Approval for use of Gen AI tools, both public and private, should be sought following established university software or cloud service request processes. If unsure, contact the IT Service Desk in the first instance.

60. The use of AI systems and services should consider opt-out options relating to the use of University data for algorithm training purposes. It is recommended that these options are activated before first use. This will prevent data entered into AI services, in the form of prompts for example, from being used for training. If the opt out selection is unclear or not available on the AI application, please contact IT Service Desk for further clarification.

61. When using AI applications for business purposes, university email addresses must be used for registering and logging-in.

62. Private AI systems should only be used by those who have completed appropriate training to protect data confidentiality and integrity and who only use it as part of approved business processes.

63. Use of Gen AI systems must be lawful and not jeopardise the University's professional reputation or brand.

64. Users will be accountable for any issues arising from their elective use of Gen AI as part of business processes, including, but not limited to: copyright violations, sensitive data exposure, poor data quality, and bias or discrimination in outputs.

65. This CoP must be read in conjunction with other university IT Security policies, and any other university policies that apply.

COP MANAGEMENT

CoP Review

66. This CoP will be reviewed every two years. It is the responsibility of the Head of Cyber Security to ensure that these reviews take place and remains internally consistent.

67. Annual reviews resulting in minor changes shall be signed off by the Director of Digital Services prior to deployment.

68. Substantial changes to the CoP shall be approved by the Digital Services Senior Management Board prior to deployment.

69. Substantive changes to this CoP shall be communicated to all relevant Users.

Document Control

Organisation	Leeds Beckett University
Author(s)	Dominic Jennings
Developed in consultation with	Matthew Page, Legal Services
Owner	Leon Etherington
Target audience	All staff, students and all other relevant parties
Sensitivity	Public
EDI Assessment	The CoP outlines the principles of the university's approach to IT and Cyber Security, which include a commitment to comply with legal and regulatory requirements by adhering to applicable UK laws and regulations ensuring fair, transparent and equal implementation of controls. No significant EDI impacts have been identified. The CoP will continue to be monitored to ensure fairness, transparency, and equality of treatment for all users.
Approved by	Approved by Digital Services Board
Endorsed by	Director of Digital Services

Effective date	15-01-2025
Review Date	+2 years from effective date [01-2027]
Status	Live
External references	N/A
Links to other internal policies / procedures	University policies | Leeds Beckett University
Version reference	V1.2
Version History - summary of changes	V1.1 Title amended from 'AI Acceptable Use Policy' (25th April 2025) V1.2 Minor updates to document structure and change from policy to Code of Practice.