Artificial intelligence governance policy

Policy statement

Bow Valley College provides Information Technology Resources (ITR) to enable its learning, teaching, administration, and research activities. At the same time, the college is obligated to ensure these same resources are used in a secure, effective, reliable, lawful, ethical, and respectful manner. The college, through its policies and its operations, will endeavor to balance these diverse needs.

Purpose

The aim of this policy is to define the appropriate use of Artificial Intelligence (AI) systems within the college.

AI governance is required to ensure that AI technologies are developed and used responsibly, ethically, and safely. While addressing risks such as bias, privacy breaches, and ethical dilemmas, in order to promote transparency, accountability, and public trust in AI systems used at the college.

Scope

This policy applies to all users who access, for any duration, through any means, or from any location, the college's ITR, and the consumption of AI services by employees or learners on behalf of the college.

Principal objectives

The policy will provide a description of what is meant by AI, identify roles and responsibilities related to the consumption and contracting of AI services, and establish an understanding that AI services must be engaged in a way that:

- supports cybersecurity best practices.
- supports information privacy best practices.
- engages Executive Team member responsible for artificial intelligence, governance, and technical support of the services.
- supports dataset integrity throughout the use of AI.
- supports the accuracy of AI system outputs.
- supports the transparency of decision-making and outputs of AI systems;
- defines and adheres to appropriate risk assessment processes, risk management frameworks and governance;
- assesses the social and ethical impact on the college, its stakeholders and the community at large throughout the use of AI.

1 Cybersecurity and Privacy Best Practices

1.1 Internally developed AI services, and AI service providers must adhere to industry-accepted Information Security and Privacy frameworks.

1.2 Departments will remain accountable for the confidentiality, integrity and availability of information that is internally developed or developed by AI providers on behalf of the college.

1.3 AI services must comply with the college's policies regarding information security, electronic communications, privacy, and identity management.

2 Executive, Governance and Technical Support
- 2.1 Engages the Executive responsible for Corporate Services, Director of Information Technology Services (ITS), Director responsible for information technology security and the Artificial Intelligence Governance Committee in the assessment, selection, and procurement of AI services.
- 2.2 Engages the Instructional and Learning Technology Governance Committee in the selection of AI services that pertain to Instructional Learning and Design.
- 2.3 Engages ITS and IT Security in the determination of support for security, hosting, maintenance, and technical support required for deployment of AI services.

3 Dataset Integrity
- 3.1 Initial assessment of datasets for integrity of data must occur prior to deployment of AI systems within the college.
- 3.2 Recurring assessment of datasets must occur periodically during the use of the AI systems within the college.
- 3.3 If dataset integrity is determined compromised the college will take necessary action to address the issue, including but not limited to informing the AI Governance Committee, conducting a cybersecurity review, conducting a privacy review, informing the AI Service Provider of the compromise, or a pre-determined alternate process will be deployed for the duration of operations until the compromise is remediated.

4 AI System Output Accuracy
- 4.1 Initial assessment of the accuracy of AI system outputs must occur prior to deployment of AI systems within the college and must determine what level of accuracy is deemed acceptable.
- 4.2 Recurring assessment of accuracy must occur periodically during the use of the AI systems within the college.
- 4.3 If the AI systems output is determined to have decreased in accuracy below acceptable levels, a method of recourse is implemented, including but not limited to informing the AI Governance Committee, or a pre-determined alternate process will be deployed for the duration of operations until the issue is remediated.

5 Transparency of decision making and outputs
- 5.1 Initial assessment of the transparency of AI systems, their decision-making processes and outputs must occur prior to the deployment of AI systems within the college.
- 5.2 AI system providers must provide details regarding the decision-making processes of their systems in an open and transparent manner, at the initial deployment of the AI system within the college; any time the AI system or underlying model changes.

6 Risk Assessment, Management and Governance
- 6.1 Allows for the management of risk in the same or similar manner within both internally developed and AI service providers.
- 6.2 Security, Privacy and AI Governance risk assessments of new AI services and service providers must be conducted with IT Security, the Privacy Officer, the Risk and Insurance Manager, the AI Governance Committee, and the stakeholder department(s).
- 6.3 To ensure business continuity and to manage unacceptable risks, ITS and stakeholder departments will develop an appropriate exit strategy before using AI services.
- 6.4 The AI provider must provide preference to services and service providers with Canadian data residency for the AI service provided and manages risk in accordance with 600-1-4,

Enterprise Risk Management Policy. In addition must align risks to the risk categories defined in *section 9*.

7. Social and Ethical Impact
   - 7.1 Initial assessment of the social and ethical impact of AI systems must occur prior to the deployment of AI systems within the college.
   - 7.2 Recurring assessment of social and ethical impacts must occur periodically during the use of the AI systems within the college. This includes, but is not limited to bias impacts (explicit and implicit), diversity, equity and inclusion impacts, workforce impact and training impacts, college value and cultural impacts, societal impacts, future AI impact, technical measures that relate to social and ethical impact, and provide expert guidance on ethical dilemmas and ensure that AI solutions align with ethical principles.
   - 7.3 Provide scope and definition of AI systems informed within this policy for use within the college. In-scope domains are defined as generative AI, expert systems, vision, robotics, planning, and speech recognition.

8. Define alignment with related Regulatory Compliance and Legal Acts
   - 8.1 Maintain alignment with the risk-based approaches defined in and compliance with
     - The proposed federal Artificial Intelligence and Data Act (AIDA); currently listed in Bill C-27.
     - The European Union's (EU) AI Act.
     - When using or developing AI content or tools, all learners, employees and contractors are expected to comply with all applicable laws, regulations, and related College policies.

9. Conduct all risk-based decisions as defined through the EU AI Risk Level Classifications & General-Purpose AI systems definitions. Classifications are, unacceptable risk (prohibited), high risk, limited risk, and minimal risk.

   Systems defined in the risk classification:
   - 9.1 Unacceptable will be prohibited from use, also:
   - 9.2 Systems as defined in 300-2-4: Acceptable Use of Information Technology Resources Policy section 1.12;
   - 9.3 Utilization of systems defined as not fit for use by the AI Governance Committee;
   - 9.4 Entering content that contains data as defined in 300-2-18 Data Classification Policy section 3.2.3/PII, or the College's or third parties' intellectual property and copyrighted materials into open model or unapproved AI services.
   - 9.5 High and limited will require appropriate Risk Assessment, Management and Governance processes as defined in section 6.2.
   - 9.6 Minimal will be unregulated.
   - 9.7 General-Purpose AI models (GPAI) systems must adhere to all section 1 requirements and align to defined activities for section 9.5 (high risk) classifications.

10. Definition of Oversight and Accountability of AI systems
    - 10.1 The AI Governance Committee is responsible for the assessment, selection, and procurement of AI services.
    - 10.2 All AI systems, or integrations with AI systems, must provide end-users with a system to explicitly consent to use the system as automated decision-making or AI systems. In addition, must explicitly notify of the use of AI systems to end-users.

10.3 Must define a business support contact and an AI Governance Committee contact address for end users to elicit feedback, report an incident, and provide a conduit for transparency concerns.

11. In alignment with related AI Acts, AI Awareness Training must be provided to all stakeholders, for the following domains: understanding AI and its impacts, legal and regulatory frameworks, ethical AI use, risk mitigation, technical skills development, and ongoing education.

12. Define Technical and Manual Quality and Safety Controls, including monitoring for all AI services in use at the College; including but not limited to, data poisoning, tuning, bias, deception/trust rot, hallucination and monitoring.

13. Define research & development requirements for the use of AI systems, or the research conducted on AI systems, including Research responsibilities, as defined in 500-3-5 Research Administration Policy, and 500-3-2: Ethical Conduct for Research Involving Human Participants Policy; adherence to all relevant legal, ethical, and documentation standards.

14. Define Copyright and Intellectual Property for the use in AI systems, or as the output of AI systems, including:
    - Unacceptable use, as defined in section 9.1
    - AI Outputs that can contain copyrighted information or others' intellectual property. While ownership in many of these cases is unclear, learners, employees and contractors should refrain from using any AI output that contains material believed to be under copyright protection.
    - Learners, employees and contractors must protect the College from claims against copyright infringement and/or theft of intellectual property. All AI-generated content must be reviewed by the individual using the Generative AI service, and a declaration of use and/or citation included when being used for work purposes.

15. Define limitations on pedagogy, learner acceptable use and honest, and classroom use, including:
    - Other than the limitations listed above, no limitations will be placed on these uses, except those as defined in the Academic AI Working Group Recommendations and Policy, and previously defined policies as defined under 500-1 – Learner Practices

Compliance

Members of the college community are expected to adhere to the policies, procedures, and standards established by the college. These policies are designed to foster a safe, respectful, and inclusive environment that supports learning, teaching, and professional integrity.

Non-compliance may create risk for the college and will be addressed accordingly through applicable college policies, procedures, and contracts. This may result in disciplinary action, up to and including termination for employees and expulsion for students.

Definitions

Artificial Intelligence (AI):
Technology that enables computers and machines to simulate human intelligence and problem-solving capabilities that can perform tasks that would otherwise require human intelligence or intervention.

Data Poisoning:
AI data poisoning is a type of cyber attack where an attacker deliberately manipulates the data used to train or influence an AI system, aiming to corrupt its outputs or degrade its performance. This can involve introducing subtly altered or entirely fabricated data points into a training dataset, with the intent to cause the AI to learn incorrect patterns, make errors, or exhibit biased behavior. Data poisoning targets the learning process itself, potentially compromising the integrity of the AI without direct interference with its code or operation.

Data Tuning:
AI tuning, also known as model tuning or hyperparameter optimization, involves adjusting the parameters that govern the learning process of an AI model to improve its performance. These parameters, which are not learned directly from the data but are set prior to training, can significantly influence how well an AI model trains and generalizes to new data. AI tuning aims to find the optimal set of these parameters to enhance the model's accuracy, efficiency, and effectiveness in solving specific tasks.

Data Bias:
AI data bias refers to the presence of prejudiced assumptions or partialities within the training data used for machine learning models, leading these models to systematically and unfairly discriminate against certain individuals or groups. This bias can result from non-representative or incomplete data samples, historical inequalities, or flawed data collection methods, and it often manifests in the model's decisions, predictions, or behavior, perpetuating or amplifying existing societal biases.

Deception/Trust Rot:
AI deception or trust rot refers to the erosion of trust in AI systems caused by instances where these systems intentionally or inadvertently deceive users. This can occur through the generation of misleading, inaccurate, or biased outputs, or when AI behaves in unpredictable or unexplainable ways. Trust rot undermines confidence in AI technologies, impacting their reliability and the willingness of users to adopt and interact with these systems.

Expert Systems:
Artificial intelligence technologies that simulate the decision-making ability of a human expert. These systems are designed to solve complex problems by reasoning through bodies of knowledge, represented mainly as if-then rules rather than through conventional procedural code. The core components of an expert system include a knowledge base, which contains accumulated knowledge and rules about a specific domain, and an inference engine, which applies the rules to the given data to derive conclusions and make decisions.

Hallucination:
AI hallucination refers to instances where an AI system generates false or fabricated information in its outputs, despite being presented with accurate data. This phenomenon typically occurs in generative AI models, such as those used for text or image creation, where the AI might produce outputs that are unconnected to or inconsistent with the input data or known facts. Hallucinations can be a result of model overfitting, lack of sufficient training data, or errors in the model's learning process.

Information Technology Resources (ITR):
Consists of business information created, stored, viewed, displayed, printed, or transmitted in whole, or in part, using the College's or personal technology resources; business information is information that supports the College's mandate. Technology resources are stand-alone or networked computer and telecommunication systems, and current or future technology or techniques used to access College business information. A College technology resource is technology the College owns, leases, hosts, maintains, supports or is under legal obligation to manage in part or in whole.

General-purpose AI model or system (GPAI):
Model: An AI model, including when trained with a large amount of data using self-supervision at scale, that displays significant generality and is capable to competently perform a wide range of distinct tasks regardless of the way the model is placed on the market and that can be integrated into a variety of downstream systems or applications.

Generative AI:
Artificial intelligence technologies that have the ability to generate new content, including text, images, audio, and video, based on their training data. These systems utilize advanced machine learning models and algorithms, primarily deep learning techniques such as neural networks, to analyze and learn from large datasets. Once trained, generative AI can produce outputs that are novel yet similar in structure and content to the original data it has learned from.

Planning:
Artificial intelligence technologies that focuses on the selection and organization of actions and tasks to achieve specific goals. These systems use algorithms to devise strategies or plans for executing tasks under constraints and varying conditions, often in complex and dynamic environments.
AI planning systems utilize a planner or a planning engine that interprets a given set of objectives, understands the resources available, and then generates a sequence of actions that will lead to the desired outcome. This planning process involves reasoning about the temporal and logical aspects of action sequences, managing dependencies between actions, and optimizing resource allocation.

Robotics:
Robotics equipped with artificial intelligence technologies that enable them to perform tasks autonomously or with minimal human intervention. These systems integrate AI capabilities such as machine learning, computer vision, and natural language processing to interpret their environment, make decisions, and learn from their experiences.
The key components of AI robotics systems often include sensors for data acquisition, actuators for movement control, and software algorithms that process input data to drive actions. The intelligence of these robots allows them to navigate complex environments, recognize objects, interact with humans and other machines, and adapt to changing conditions.

Speech Recognition:
Also known as automatic speech recognition (ASR), is a AI technology that enables computers to understand and process human speech into a written format. This AI-driven capability involves analyzing the sound waves of speech, interpreting the content, and converting it into text. The process entails several complex steps including audio signal processing, phonetic and linguistic analysis, and leveraging deep learning models, particularly neural networks, to accurately decipher spoken words regardless of accent, speed, or background noise.
AI speech recognition systems are characterized by their ability to learn and adapt to new accents, dialects, and vocabularies over time, improving their accuracy and usability. As these systems continue to advance, they play a critical role in enhancing human-computer interaction, making technology more accessible, and streamlining communication processes across various sectors.

Vision:
Artificial intelligence technologies that enable computers and systems to derive meaningful information from digital images, videos, and other visual inputs. They use machine learning algorithms, particularly deep learning models like convolutional neural networks, to interpret the visual world in a manner similar to human sight.
The capabilities of AI vision systems include object detection, image classification, facial recognition, scene reconstruction, event detection, and image restoration among others. These systems learn from vast amounts of visual data, identifying patterns and features that are used to make decisions or recommendations.

Data sheet

Accountable officer
Executive Team member responsible for Information Technology Services

Responsible officer
Associate Director, ITS Security

Approval
Executive

Contact area
ITS Security

Relevant dates

Approved	Executive: EXT241029-02
Effective	January 2, 2025
Next review	October 2025
Modification history	
Verified By	Office of the President, December 2024

Associated policy(ies)
Acceptable Use of Information Technology Resources Policy (300-2-4)
Cloud Computing Policy (300-2-15)
Enterprise Risk Management Policy (600-1-4)
Ethical Condict for Research Involving Human Participants Policy (500-3-2)
Information Security, and Identity Management Policy (300-2-11)
Privacy and Access Policy (300-2-10)
Data Classification Policy (300-2-18)
Research Administration Policy (500-3-5)
500-1 – Learner Practices

Directly related guideline(s) (if any)
List any guidelines that support the purpose of this policy. List in alphabetical order.

Related legislation
Bill C-27 Artificial Intelligence and Data Act (Canada)
EU Artificial Intelligence Act

Attachments (optional)
Forms
FAQ
Matrix