



Company: BioNorth Solutions

Contact: Amber Kivisto (amber@bionorthsolutions.com)

Job Title: AI Intern, Environmental Engineering Data Extraction

Job Description: We are seeking a motivated and curious AI/ML Intern to support a 15-week project focused on developing automated methods to extract structured data from environmental engineering reports. These reports often contain complex technical language, tables, figures, and regulatory information. The intern will help design, build, and evaluate machine learning and natural language processing (NLP) pipelines to make this information more accessible and actionable.

This role is ideal for someone interested in applied AI, environmental data, and real-world document intelligence problems.

Key Responsibilities:

Document Understanding & Data Extraction

- Analyze environmental engineering reports to identify key data fields, terminology, and extraction targets.
- Assist in building annotation guidelines and labeling datasets.
- Perform manual data annotation to create training and evaluation sets.

Model Development

- Experiment with NLP and machine learning techniques (e.g., named entity recognition, text classification, table extraction, OCR).
- Train, validate, and test models using modern frameworks.
- Evaluate model performance and propose improvements.

Data Engineering & Preprocessing

- Clean, structure, and preprocess text from PDFs and other report formats.



Federal Economic Development
Agency for Northern Ontario

Agence fédérale de développement
économique pour le Nord de l'Ontario

Canada



- Work with tools for OCR, text parsing, and table extraction.

Research & Prototyping

- Review relevant literature on document AI, environmental data extraction, and regulatory reporting.
- Prototype workflows for automated extraction and summarization.

Required Qualifications:

- A degree or diploma in **Computer Science, Data Science, Environmental Engineering, or a related field**.
- Foundational knowledge of **Python** and common ML/NLP libraries (e.g., PyTorch, TensorFlow, scikit-learn, spaCy, HuggingFace).
- Strong analytical and problem-solving skills.
- Ability to read and interpret technical documents.
- Interest in environmental engineering, sustainability, or regulatory reporting.

Preferred Qualifications

- Experience with **OCR tools**, PDF parsing, or document AI frameworks.
- Familiarity with **NER models**, transformers, or LLM-based extraction.
- Coursework or experience in **environmental science or engineering**.
- Experience working with unstructured text datasets.

Other Information: The Youth Tech Impact Program requires candidates to fulfill the following requirements:

- Be between the ages of 18-29





- Be legally allowed to work in Canada
- Be living in Northwestern Ontario

Location of Work:

This position is in-person only.

While participating with the Youth Tech Impact Program, applicants must be willing to participate in all of the program's orientation, learning sessions, peer to peer sessions, retreat and graduation.



Federal Economic Development
Agency for Northern Ontario

Agence fédérale de développement
économique pour le Nord de l'Ontario

Canada