

British Board of Film Classification

Job Title Data Engineer
Department Product & Business Services
Reports to Head of Business Services

Company Overview:

Independent and not-for-profit, the BBFC is here to help everyone in the UK – especially children and families – choose age-appropriate films, videos and websites, wherever and however they watch or use them.

We are committed to using AI to evolve and leverage over 110 years of experience in media compliance and regulation. We are on a mission to harness this extensive knowledge to create scalable solutions that align with our core mission of supporting audiences to make safe and informed viewing choices.

As a Data Engineer in our organisation, you'll play a crucial role in shaping this innovative new technology, with the aim of supporting the industry to meet the changing ways that audiences now enjoy content. As part of a dynamic team, the successful candidate will provide integral support on the two main projects currently in progress: an AI solution that identifies and tags content issues, as well as a separate tool that will determine and assign international age ratings. Further information about these solutions is available on our website [here](#).

Working closely with influential collaborators, including the University of Bath (UoB) and Amazon Web Services (AWS), this is an outstanding opportunity for the ideal candidate to develop their skills and expedite their career growth.

Responsibilities:

1. Data Acquisition and Integration:

- Develop and implement robust data pipelines for acquiring, ingesting, and integrating diverse datasets.
- Utilise Python for scripting and automation to ensure efficient data extraction and integration.

2. Web Scraping:

- Leverage web scraping techniques to gather relevant data from various online sources.
- Ensure data quality and integrity in the web scraping process.

3. Data Preparation and Cleaning:

- Work closely with data scientists and analysts to understand data requirements.
- Perform data cleaning, transformation, and enrichment to prepare high-quality datasets for analysis.

4. Data Processing:

- Design and implement scalable data processing solutions for large volumes of structured and unstructured data.
- Utilise SQL for efficient querying and manipulation of databases.

5. MLOps:

- Implement and maintain MLOps practices to streamline the deployment and management of machine learning models.
- Collaborate with data scientists to operationalise and scale machine learning solutions.

6. Data Pipelines:

- Design, build, and maintain end-to-end data pipelines to support AI for good projects.
- Implement best practices for orchestration, scheduling, and monitoring of data pipelines.
- Automate data pipeline processes to improve efficiency and reliability.

7. Cloud and API Integration:

- Utilise AWS services, including SageMaker, for deploying and managing machine learning models.

- Implement cloud storage solutions and leverage APIs for seamless data integration.

8. Image Recognition:

- Understand and work with image recognition models to process and analyse visual data.
- Familiarity with data tagging and classification techniques for image datasets.

9. Software Engineering:

- Possess experience in software engineering practices.
- Contribute to the development and maintenance of software applications related to data processing and analytics.

10. Collaboration:

- Collaborate with cross-functional teams, including data scientists, software engineers, and business analysts, to understand project requirements and deliver effective solutions.
- Act as a key contributor to our AI for good initiatives, contributing your expertise to projects aimed at making a positive impact on society.

Requirements:

1. Bachelor's or Master's degree in Computer Science, Data Science, or a related field or equivalent .
2. Proven experience as a Data Engineer or similar role.
3. Proficiency in some or all of the following technologies:
 - Python for scripting and automation.
 - SQL for database querying and manipulation.
 - Web scraping tools and techniques.
 - Data processing frameworks and tools.
 - MLOps practices and tools.
 - AWS services, including SageMaker.
 - Cloud storage solutions and API integration.
 - Strong understanding of development tools and principles, e.g. GitHub, DevOps
4. Strong understanding of data architecture, data modelling, and database design principles.
5. Experience with building and maintaining data pipelines for large-scale data processing.
6. Familiarity with image recognition models and data tagging/classification techniques.
7. Software engineering experience, with an understanding of software development best practices.
8. Experience in automating data pipelines for efficiency and reliability.
9. Excellent problem-solving skills and the ability to work independently as well as part of a collaborative team.
10. Strong communication skills and the ability to convey complex technical concepts to non-technical stakeholders.
11. Passion for leveraging technology, specifically AI, for social good.

What We Offer:

- Competitive salary and benefits package.
- A challenging and dynamic work environment.
- Opportunities for professional growth and development.
- The chance to be part of an amazing team working on projects that make a positive impact.

We look forward to welcoming you to our team as we collectively strive to make a positive impact through AI for good.

