



## **DigiEduHack Challenge Data Exploitation**

# **How can we use datasets to track the evolution of knowledge?**

## **How can we use datasets to track the evolution of knowledge?**

Venture into the historical datasets released by the National Library of Scotland to track the development of education and knowledge across the first eight editions of the Encyclopaedia Britannica, from 1771 to the mid-nineteenth century, making use of your text analysis & critical thinking skills.

### **University of Edinburgh Digital Education Hack**

Join our DigiEdHack Hack at the University of Edinburgh to connect with other people interested in digital education, innovation, and entrepreneurship to tackle the challenges facing education today, including interconnectivity, interdisciplinarity and sharing of knowledge. Participate in the challenges, meet new and exciting people, and turn your vision into reality! You will have the opportunity to showcase your solutions on a European and global scale, and the best teams can win €5,000 in a global award contest!

## **How to prepare for the challenge?**

The National Library of Scotland will soon be publishing a text-based dataset containing the first eight editions of the Encyclopaedia Britannica, from 1771 to the mid-nineteenth century. We will let you know when the dataset is published so you can have a look around and familiarise yourself with it. In the meantime, you can brush up on your natural language processing skills, or look at the existing information on the collection to begin thinking about interesting questions you can ask of the data.

## **Why is it important to find a solution to this challenge?**

The analysis of text-based data is becoming an increasingly important part of data science and digital scholarship. Digital scholars are using new techniques in natural language processing to make sense of enormous datasets and gain more knowledge about our past, present and future. By contributing to this challenge, you will be developing solutions that will help future scholars gain more information on how to work with these datasets in productive ways.

## **What is the possible impact of a good solution?**

A good solution for this challenge will not only allow you to begin creating solutions for working with historical databases, but it might help learn more about how education has been evolving for the past two centuries.

## **What is the expected outcome of this challenge?**

The expected outcome of this challenge will be the development of ways in which we can learn more from text-based historical datasets. This doesn't have to be limited to code that processes the data, it can also be a set of questions that can be asked from the data to better understand the information contained in it and build the foundations for further analysis.