Anti-Scraping Techniques

Shubham Sandeep Zujam¹ and Prof. Bhanudas Satam²

Student, Department of MCA¹
Mentor, Department of MCA²

Late Bhausaheb Hiray S. S. Trust’s Institute of Computer Application, Mumbai, India

Abstract: Web scraping refers to a software program that mimics human web surfing behavior by pointing to a website and collecting large amounts of data that would otherwise be difficult for a human to extract. A typical program will extract both unstructured and semi-structured data, as well as images, and convert the data into a structured format. The activity is deemed illegal, but the change in legality has not stopped people from doing the same. This paper aims to list challenges and proposes techniques to developing anti-scraping application.

Keywords: Web scraping

REFERENCES


