

# Data Extraction from E-commerce Website using Web Scrapping with Python and Selenium

Unnati N. Changole, Sandesh B. Harne, Prof. Ashwini P. Ghatol

Students, Department of Computer Science & Engineering  
Sipna College of Engineering & Technology, Amravati, Maharashtra, India

**Abstract:** *The internet provides access to a vast array of information and data gathered by people. It will, however, include a vast array of disparate and poorly structured data, which will be difficult to obtain through physical means and tricky to use in mechanical operations. Procedures and numerous outfits have been established in the recent past to allow data collection and transformation into ordered information by B2C and B2B systems. Beginning with a basic introduction and a brief review of various online scraping applications and tools, this article will focus on various elements of web scraping. We also went through the process of web scraping, as well as the numerous types of web scraping techniques, before concluding with the benefits and drawbacks of web scraping, as well as a detailed discussion of the various sectors where it may be used. The possibilities for utilizing these data are vast, including Open Government Data, Big Data, Business Intelligence, aggregators and comparators, as well as the development of new apps and mash ups, to name a few.*

**Keywords:** Web scrapping, Data Extraction, Business Intelligence, Automation, Selenium, HTML Parsing

## REFERENCES

- [1]. Shakra Mehak, Rabia Zafar, Sharaz Aslam, Sohail Masood Bhatti “Exploiting Filtering approach with Web Scrapping for Smart Online Shopping” 2019 International Conference on Computing, Mathematics and Engineering Technologies – iCoMET 2019
- [2]. Vidhi Singrodia, Anirban Mitra, Subrata Paul “A Review on Web Scrapping and its Applications” 2019 International Conference on Computer Communication and Informatics (ICCCI -2019), Jan. 23 – 25, 2019, Coimbatore, INDIA
- [3]. Sarah Fatima , Shaik Luqmaan , Nuha Abdul Rasheed “Web Scrapping with Python and Selenium” IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 23, Issue 3, Ser. II (May – June 2021), PP 01-05
- [4]. K Usha Manjari, Syed Rousha, Dasi Sumanth, Dr. J Sirisha Devi “Extractive Text Summarization from Web pages using Selenium and TF-IDF algorithm” Proceedings of the Fourth International Conference on Trends in Electronics and Informatics (ICOEI 2020) IEEE Xplore Part Number: CFP20J32-ART; ISBN: 978-1-7281-5518-0
- [5]. Sanjay Kumar Malik1 , SAM Rizvi2 “Information Extraction using Web Usage Mining, Web Scrapping and Semantic Annotation” 2011 International Conference on Computational Intelligence and Communication Systems