



# Introduction to Data Extraction Services

by Apurva Chaturvedi, Account Manager, Content Engineering  
& Amit Siwach, Manager, Content Engineering

# Introduction



E-commerce (also known as electronic commerce or internet commerce) refers to the cyber purchase and sale of goods or services.

In today's highly competitive e-commerce world, an organization's most valuable asset is its captured data. The use of data extraction services ensures that information is used effectively. At GlobalLogic, our data extraction process includes various strategies to enhance our clients' businesses and improve their overall online presence and growth. We use programming robots that collaborate with and pull data from unique sites and convey it in real-time, continually refreshing and incorporating your framework's information.

# Data Extraction Tools & Programs

Data extraction services help to:

- Improve your accuracy
- Leverage competitive research
- Consistently update your product offerings
- Boost your SEO
- Perform more in-depth client research
- Increase productivity
- Study financial markets
- Generate leads
- Monitor your brands

Our dedicated data extraction team engages with some of the most advanced datasets to optimize, automate, and simplify the processes.

Data extraction powers smoothly running systems and generates more content, providing a higher return on investment. Here are some tools and programs to support our clients' needs:

**Web crawling:** A web crawler is an automated, programmed script that browses the web. The automated collection leaves more time for data analysis and verification.

**Data extraction:** We collect or retrieve disparate types of data from various sources, even if they are poorly-organized or completely unstructured.

**Dashboards:** We create a dashboard that allows us to analyze, track and display metrics performance, which shows an organization's growth.

**Reports:** We create reports in an organized format for specific audiences and purposes.

**Data quality:** We measure data quality based on reliability, consistency, completeness, and accuracy.

**Quality team:** Our quality team conducts meticulous tests throughout the process, identifying possible issues using data reliability and site investigation. They report any concerns to the development team for resolution.

# Understanding the Processes of Data

## Web Crawling

We provide automated crawling technology with a user-friendly interface that guides you through the whole extraction process, performing the tedious, often repetitive work in an efficient manner without the need for additional experts. Any extracted data is presented in a number of formats, including HTML, CSV or Excel formats.

We set up a web crawler to identify and collate partial or total content from websites and provide all the information in a manageable format.

- Results are accurate, refined, relevant, and precise.
- Lowers the total cost of ownership and return on investment (ROI).
- Flexible: resolves issues within a specified time frame.

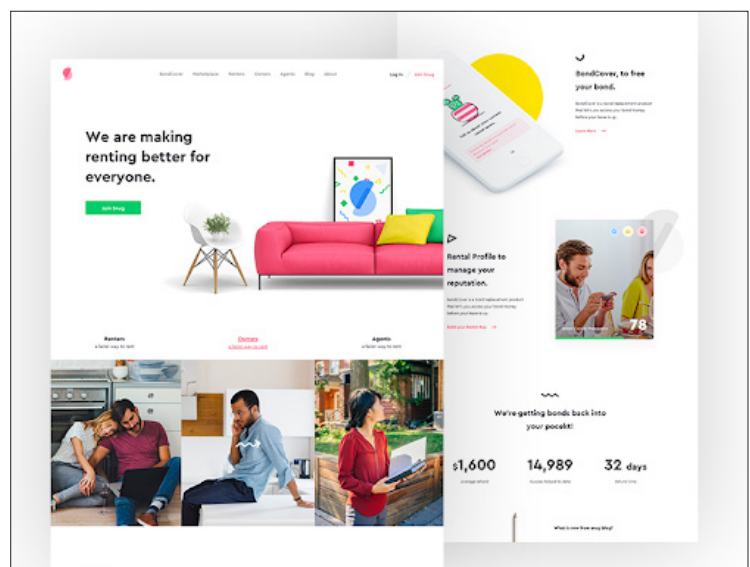
## Extracting data from dynamic web sites

We can extract data from sites with multiple levels of navigation, including pagination and product pages.

## Landing pages

Websites display individual products on web pages known as landing pages. The landing page contains product information, and we can extract the required data, such as:

- Price
- Availability
- Description
- Image
- Brand
- SKU
- MPN
- GTIN
- Breadcrumb



## Dashboard

We create customizable dashboards to meet the specific needs of a department and company.

In the background, a dashboard connects with your APIs, documents, attachments, and services.

In the foreground, it shows this information as charts, tables, and bar graphs.



## Reports

Reports can display weekly, monthly, quarterly or yearly data and can be created in both tabular form or as graphic displays.

	A	B	C	D	E	F	G	H	I	J
1	Weekly Project Status Report Template									
2										
3	Completed items									
4	Project	Task	Team Members	Estimation	Note					
5	Billing Table update	Requirements	Umair	5-Jul	Got Stakeholder Interview Done					
6	Billing Table update	Design	Kiran	12-Jul						
7	Billing Table update	Code	Umar	25-Jul						
8	Customer Svc Connecto	Design	Umair	15-Jul						
9	In Progress									
10										
11	Project	Task	Team Members	Estimation	Note					
12	Billing Table update	Requirements	Umair	5-Jul						
13	Billing Table update	Design	Kiran	12-Jul						
14	Assigned but Not Started									
15										
16	Project	Task	Team Members	Estimation	Note					
17	Billing Table update	Requirements	Hafaz Iqbal	5-Jul						

## Data quality

Our emphasis on data quality in enterprise systems has increased as data processing and analytics have become more intricately linked with business operations. This includes:

- Data quality reports
- Data monitoring
- Consistency
- Accuracy
- Completeness

## Outcomes

### **Improve Search Engine Rankings.**

You can monitor improvements in website ranking through increased click-through rate (CTR), which helps us track the percentage of total clicks versus total impressions.

### **Increase Targeted Traffic.**

If you generate the right kind of traffic to your website and design your website to deliver the content customers need, you'll end up with happy, satisfied visitors. This will boost your brand's online reputation.

### **Generate Revenue.**

As you optimize your website, you will see your sales increase and your revenue grow.



### **More Leads.**

If customers purchase something during their first visit to your website, there is a strong possibility that they will buy again in the future. A continuous flow of superior and high-quality content will lead to exceptional user experience, increasing traffic to your website.

## Conclusion

The information obtained from a data extraction service has the potential to improve the way companies run their operations. Smart methodologies that envision a client's needs help businesses run more smoothly and more effectively.

Data exists everywhere and in many different formats. With data extraction services, we obtain this data with a crawling team in an efficient and effective way.

Our team provides dashboards and in-depth reports, ensuring the accuracy of data-driven statements from our quality team. We believe that businesses should be forward-thinking so they can become leaders when new trends appear, and they should revise and review their plans regularly.



## About the Authors

**Apurva Chaturvedi** is an account manager for content engineering (e-commerce and travel) operations at GlobalLogic with 17+ years of experience in the International Call Centre, Outsourcing Services, and ITES Business. Apurva leads in client servicing, operations management, business development, team management, analytics, process management, and quality management. Apurva holds a bachelor's degree in commerce from Delhi University and a diploma in web design.

**Amit Siwach** is an operations manager at GlobalLogic. He has 16 years of experience in content engineering, with specialized expertise in process improvements and delivery management. He holds a bachelor's degree in engineering from MDU and a PG diploma in business administration from Symbiosis Pune.



GlobalLogic is a leader in digital product engineering. We help our clients design and build innovative products, platforms, and digital experiences for the modern world. By integrating strategic design, complex engineering, and vertical industry expertise,— we help our clients imagine what's possible and accelerate their transition into tomorrow's digital businesses. Headquartered in Silicon Valley, GlobalLogic operates design studios and engineering centers around the world, extending our deep expertise to customers in the communications, automotive, healthcare, technology, media and entertainment, manufacturing, and semiconductor industries.



[www.globallogic.com](http://www.globallogic.com)