



ZYTE PAPER

Recruitment Case Study

Mining web data to uncover job listing & wage trends

How one of the leading
e-commerce companies uses
Data on Demand to make
better hiring decisions.



Overview

The human resources (HR) industry is experiencing huge change thanks to the increasing use of data in decision making. Many HR teams are sitting on a untapped data gold mine that includes everything from recruitment data, career progression data, training data, absenteeism figures, productivity data, personal development reviews, competency profiles and staff satisfaction data.

What's more with the increasing use of data extraction technologies HR teams can now mine and analyze external data sources such as job boards and social media profiles for critical insights that can be used to improve decision making, make employees happier, optimize processes, and add value to the company.

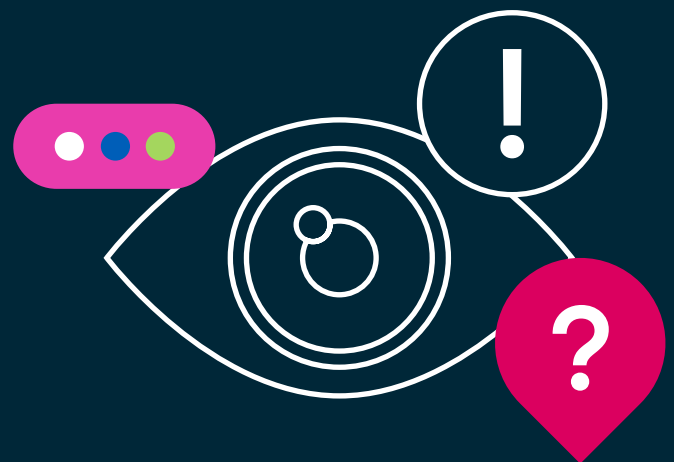
This trend has given rise to what is known as "intelligent HR", which in reality is data-driven HR. Through the increased use of data, companies can find the best talent, boost employee retention, and make better hiring decisions in an increasingly competitive labor market.

In this case study, we will show you how we helped one of the leading e-commerce giants gather the job and wage data they needed to discover hiring trends in key strategic fulfilment (FBA) locations and understand regional salary requirements, so they can better optimise their hiring of blue-collar seasonal workers.

This customer needed to be able to extract job and wage data from hundreds of jobs sites every week and combine it with their own internal demand and labour forecasts to make accurate predictions of seasonal hiring and wage rate trends in each of their fulfilment centre locations.

With this data their data science team has been able to give hiring managers accurate forecasting on the available pool of contractors in each location and the associated wage levels. Enabling them to make accurate hiring plans months in advance, whilst simultaneously reducing their personnel costs for these seasonal hires.

Note: Due to the sensitive nature of the information in this case study we won't name the customer or give any identifying information for confidentiality and competitive advantage reasons.



Ours Customer's objective

The company in question is one of the leading e-commerce companies in the world, employing thousands of staff in its fulfilment locations.

However, due to the nature of e-commerce, there is a strong seasonal component to online purchases. During key periods of the year, (Thanksgiving, Christmas, etc.) e-commerce stores experience huge surges in online transactions requiring them to hire additional workers for these periods of very high demand.

To make better hiring decisions and reduce their personnel costs the customer wanted to accurately forecast seasonal hiring and wage rate trends in each of their fulfilment centre locations.

With this data the company's data science team would be able to give operations and hiring managers accurate forecasting on the available pool of contractors in each location and wage bracket. Enabling them to make accurate hiring plans months in advance and lower their personnel costs.

The customer had a dedicated internal web scraping teams in house, however, for this project the data science team decided it would be more efficient to outsource the design and development of this data extraction system to an external company. They had attempted

to build their own data extraction system to extract this data, however, after building the initial prototype they determined that allocating more engineering resources to build and maintain the infrastructure wouldn't be a good allocation of their time. To make better hiring decisions and reduce their personnel costs the customer wanted to accurately forecast seasonal hiring and wage rate trends in each of their fulfilment centre locations.



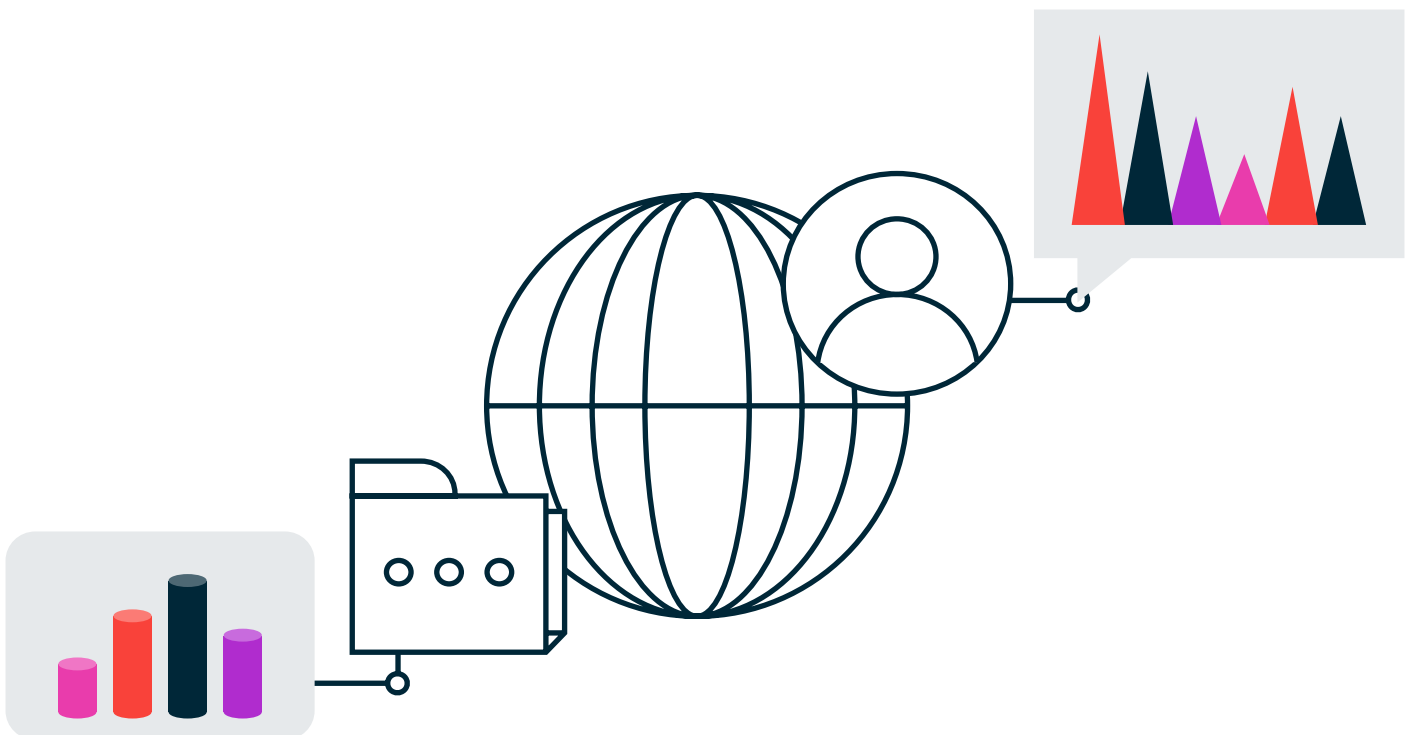
Project scope

After a number of calls with our Solution Architecture team, the customer and our team came to the conclusion that their primary interest was in the underlying data, not the data extraction infrastructure itself.

They had no interest in maintaining web crawlers or managing proxies. All they wanted was a reliable data feed of clean and accurate job data for their data science team so that they could make geographical hiring forecasts which hiring and operations managers could use to make more informed hiring decisions.

The less input they had to give to the data extraction process the better. Their ideal scenario was to be able to simply name the data they wanted and Zyte would take care of the rest.

As a result of these conversions, it became clear that Data on Demand, Zyte's data extraction as a service offering, was the best solution for their needs. With Data on Demand, the customer could simply subscribe to a data feed for the data they want and Zyte would deliver it to them how they wanted it. The customer didn't need to worry about building spiders or maintaining proxies, freeing them up to focus on analysing and using the data.

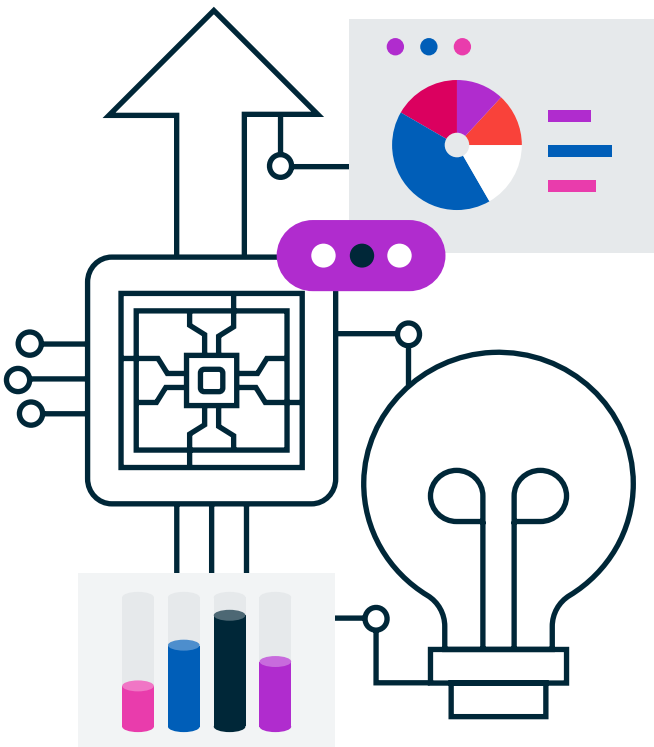


Project execution

During the consultation process, the customer outlined the types of data they wanted to extract from which sites, along with how often they needed the data to be extracted.

From there Zyte's team of engineers built the underlying data extraction infrastructure and set up a data feed to deliver the data to the customer each period in the format that best suited them.

For an initial proof of concept project, the customer decided that they wanted to extract data from 26 different job sites and have the data delivered to them weekly via their Amazon S3 bucket.



The execution of the project was as follows:



The customer outlined the data they needed along with the desired frequency.



Our Solution Architect team created a scope for the project and gave an estimated cost.



We delivered the customer sample data for their approval.



Upon approval we ramped up the data extraction and began delivering full data sets to the customer.



Our QA team continuously monitored the customer's data feeds to ensure they received consistent high-quality data.

Why Data on Demand?

For this client, Zyte has become a crucial partner for their data science team. We don't just supply web data to them, our Solution Architect team regularly works with them to identify new data sources and troubleshoot the data acquisition challenges they are facing →



Get Up And Running Straight Away

With Data on Demand the customer didn't have to wait for their internal engineers to develop the skills to build and maintain the data extraction infrastructure they needed. They just needed to tell us the data they needed, and they had an active data feed in a matter of days.



Guaranteed Uptime

After the data feed has been approved by the customer, Zyte's team of expert crawl engineers and QA specialists will maintain the data feed so even if website structures change the customer will always have a reliable stream of data for their business.

Entrusting Zyte with their entire data extraction process enabled them to focus 100% on data analysis and research, what they do best. Outside of being a valued data extraction partner, here are some of the key reasons why this customer choose a data feed from Zyte over building in-house or a different data feed provider.



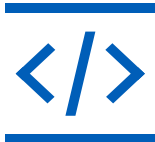
Guaranteed Data Quality

Every Data on Demand data feed comes with a data quality guarantee explicitly stated in the customers SLA that ensures they can confidently use the data in their mission-critical products and decision making without worrying about data inaccuracies or gaps.



Predictable Pricing

With traditional web scraping services, due to the number of line items (engineers, proxies, servers, QA, etc.) it is hard to accurately predict the cost of your data as you scale. With Data on Demand, data acquisition becomes a single line item. The customer could accurately forecast how much their data will cost at any scale as pricing is based on data delivered.



At Zyte we turn websites into data with industry leading technology and services.

Our solutions include:

- **Data Extraction Service**
Let our web scraping experts build and manage the bespoke data extraction solution for your business needs.
- **Automatic Extraction powered by AI**
Instantly access accurate web data through our user-friendly interface or various Extraction APIs and save time getting the data you need.
- **Smart Proxy Manager (formerly Crawlera)**
Forget about proxy lists. We manage hundreds of thousands of proxies, so you don't have to.
- **Data extraction platform**
Access developer tools, data extraction APIs and documentation, built and maintained by our world-leading team of over 100 extraction experts.

zyte

It's yours. The web data you need.

Access clean, valuable data with web scraping services that drive your business forward.

[Talk to us](#)

www.zyte.com
Copyright 2021 © Zyte

Cuil Greine House, Ballincollig Commercial
Park Link Road, Ballincollig / Co. Cork, Ireland