[Cloud Talent Solution](https://cloud.google.com/talent-solution/) (https://cloud.google.com/talent-solution/) Resources

Quotas and limits

This document contains current field and overall request quotas for Cloud Talent Solution. This page will be updated to reflect any changes to these restrictions and usage limits.

Note: For all Cloud Talent Solution APIs, the size of each request cannot exceed 10MB.

Request Limits

Job Search

The current API usage limits for Cloud Talent Solution - Job Search are as follows (averaged over a 100s time interval and are subject to change):

Type of Limit	Usage limit
Requests per 100 seconds for UserfacingGroup QPS (search, histogram and GET API calls)	500*
Requests per 100 seconds for PeriodicTaskGroup QPS (searchForAlerts API call)	10*
Requests per 100 seconds for defaultGroup QPS (indexing and List API calls)	500*
Requests per day	Unlimited
Maximum number of companies at any time*	500,000*
Maximum number of open jobs at any given time*	500,000*

* Quotas are projected up to the specified caps. It doesn't guarantee the availability of the quotas displayed above. For more information on requesting additional quota, <u>contact support</u>

(https://cloud.google.com/talent-solution/docs/support).

Profile Search (Beta)

The current API usage limits for Cloud Talent Solution - Profile Search (Beta) are as follows (averaged over a 100s time interval and are subject to change):

Type of Limit	Usage limit
Requests per 100 seconds for Indexing Profiles (Create, Update and Delete API calls)	2000*
Requests per 100 seconds for Search QPS (Search API call)	1000*
Maximum number of Profiles at any time*	20,000,000*

* Quotas are projected up to the specified caps. It doesn't guarantee the availability of the quotas displayed above. For more information on requesting additional quota, <u>contact support</u> (https://cloud.google.com/talent-solution/docs/support).

Increasing Limits

To increase QPS limits associated with Job Search API Create, Update, and Delete (CUD) usage, customers can send a written request to the Cloud Talent Solution Support team. Requests are granted up to the limits defined below, subject to availability and the validation of customers' Job Search API integration by the Cloud Talent Solution Support Team.

Customers will be eligible for an increase in CUD QPS based on the table below:

Monthly Spend Tiers		CUD Queries per 100 seconds limit
$ -	$ 1,000	500
$ 1,001	$ 10,000	2000
$ 10,001	$ 50,000	10000
$ 50,001+		20000

Cloud Talent Solution is committed to responding to QPS increase requests within 60 days of the initial request.

Queries per Second (QPS) and document restrictions

Cloud Talent Solution restricts the number of queries a customer can send to the service within a 100s time interval. Cloud Talent Solution also restricts the overall number of jobs and companies (any job or company is considered a document) that a customer can index at any given time based on the above restrictions.

Field Limits

The Cloud Talent Solution has many types of fields and thus has variations in the number of characters that can be included in each field. Fields typically support between 160-10,000 characters, depending on the field. See more information in the description of each field in the reference documentation (https://cloud.google.com/talent-solution/job-search/docs/reference/rest).

Expired jobs

Cloud Talent Solution stores a number of expired jobs in the system. By default, the retention period of these expired jobs is 60 days. However, in order to get a higher bandwidth for open jobs, customers should reach out to the support team (https://cloud.google.com/talent-solution/docs/support) to decrease or increase this limit.

These limits apply to each Cloud Talent Solution developer project and each user, and are shared across all applications and IP addresses related to a given Google Cloud project.

Sending Batch requests

Cloud Talent Solution allows you to batch up to 25 requests in a single call using the Google Cloud API batch functionality. Any batch call that includes more than 25 requests will be

rejected/inconsistently processed. It is up to the API consumer to ensure that the maximum batch size does not exceed the limit specified.