Q&A API

FACTSET SEE THE ADVANTAGE

Q&A API

Contents

1. Overview

The Q&A API is a machine learning service that provides the ability to ask questions about the information contained in text documents. The service will return answer(s) based on the information contained in the text. This service uses natural language processing techniques to extract answers from the text.

This service can be used to find the same type of information across multiple documents. For example, you may query several documents with questions such as "what are new products?", "what are new plant-based products?", or "what does this company make?"

The Q&A API accepts user queries through a Swagger-based restful API (details below).

a. Methodology

The Q&A API uses an extractive approach to provide answers from documents. It extracts answers drawing on words from the original text. This method works well for more generic questions that can be extracted from the text or across multiple documents.

The service can return multiple answers for each question asked and will return a confidence score to denote the proposed usefulness of each answer. Confidence scores are dependent on the nature of the question asked and how the answer is phrased in the document. Direct questions such as “How many employees does FactSet have?” will return a higher confidence score than an open-ended question such as “Who is in charge of FactSet?” Confidence scores will increase if the answer in the text is specifically phrased similarly to the question (e.g. “FactSet has 10,000 employees”). Any instances where the model may have to infer the proper synonyms/word mappings may return a lower confidence score. For example, if the text reads, “FactSet has 10,000 staff”, the model would return the correct answer, but would have less confidence in it.

b. Sample Q&A Functionality

Given a selection of unstructured text:

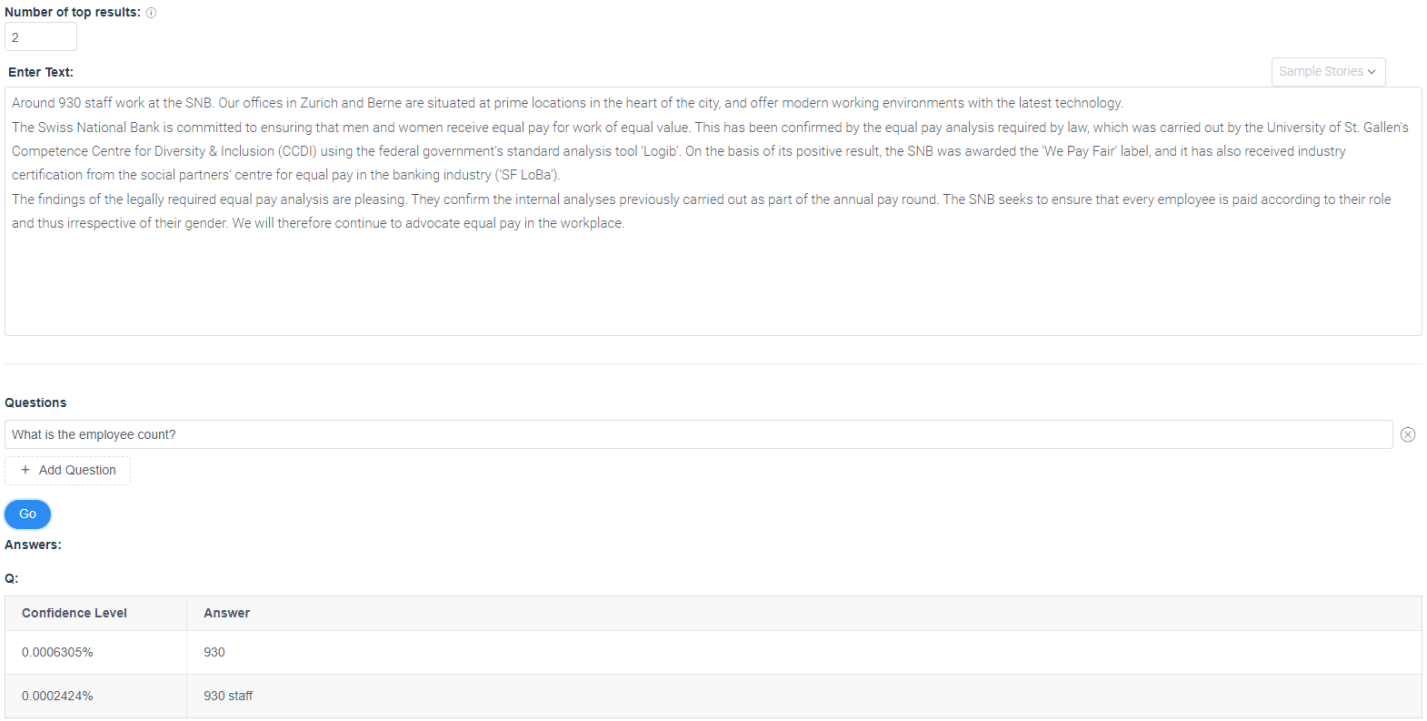
Around 930 staff work at the SNB. Our offices in Zurich and Berne are situated at prime locations in the heart of the city, and offer modern working environments with the latest technology.

The Swiss National Bank is committed to ensuring that men and women receive equal pay for work of equal value. This has been confirmed by the equal pay analysis required by law, which was carried out by the University of St. Gallen's Competence Centre for Diversity & Inclusion (CCDI) using the federal government's standard analysis tool 'Logib'. On the basis of its positive result, the SNB was awarded the 'We Pay Fair' label, and it has also received industry certification from the social partners' centre for equal pay in the banking industry ('SF LoBa').

The findings of the legally required equal pay analysis are pleasing. They confirm the internal analyses previously carried out as part of the annual pay round. The SNB seeks to ensure that every employee is paid according to their role and thus irrespective of their gender. We will therefore continue to advocate equal pay in the workplace.

The Q&A API will allow you to ask questions from the text. For example, if you were to ask, “What is the employee count?”, the Q&A API will return an answer drawn from the text along with an associated confidence level. You can also choose the number of results/answers for each question.

Below is an example of the Q&A API in action using an internal FactSet user interface. This example is for illustrative purposes only.



c. Accessing the Q&A API

The FactSet Q&A API is accessible to authorized FactSet clients only. When accessing the service, please make sure that your account has been enabled for the Q&A API. If not, calls to the Q&A API will fail. Please contact your FactSet representative if you encounter authentication issues while calling the Q&A API.

d. Document Size and Other Specifications

Currently, this service only supports English language documents in plain text format.

For performance reasons, document size is limited to 10,000 characters.

The first request made to the Q&A API will take 10-15 seconds longer to generate a response than subsequent requests as the service needs to start up. You will be provided with a unique ID immediately from

the **answers** endpoint but will have to wait to the extra time for output from the **ID** endpoint. The **status** endpoint is available to let you know when the request has completed.

2. Submitting a Request to the Q&A API

The Q&A API supports a restful HTTP-based API.

a. Q&A Endpoint

Production URLs for the service:

https://api.factset.com/cognitive/nlp/v1/qna/answers

https://api.factset.com/cognitive/nlp/v1/qna/answers/{id}/status

https://api.factset.com/cognitive/nlp/v1/qna/answers/{id}

The service supports HTTP POST for the **answers** endpoint and HTTP GET for the **ID** and **status** endpoints.

b. Building a Request-Payload for the Q&A API

A sample JSON payload to the FactSet Q&A service follows the template below:

Data {

```
    question*
                        String (required)
                        Example: What is the employee count?

                        Question to be asked from the text.

    input*
                        String (required)
                        Example: Around 930 staff work at the SNB. Our offices in Zurich
                        and Berne are situated at prime locations in the heart of the
                        city, and offer modern working environments with the latest
                        technology. The Swiss National Bank is committed to ensuring
                        that men and women receive equal pay for work of equal value.
                        This has been confirmed by the equal pay analysis required by
                        law, which was carried out by the University of St. Gallen's
                        Competence Centre for Diversity & Inclusion (CCDI) using the
                        federal government's standard analysis tool 'Logib'. On the
```

```
                    basis of its positive result, the SNB was awarded the 'We Pay
                    Fair' label, and it has also received industry certification
                    from the social partners' centre for equal pay in the banking
                    industry ('SF LoBa'). The findings of the legally required equal
                    pay analysis are pleasing. They confirm the internal analyses
                    previously carried out as part of the annual pay round. The SNB
                    seeks to ensure that every employee is paid according to their
                    role and thus irrespective of their gender. We will therefore
                    continue to advocate equal pay in the workplace.

                    English plain text to extract answers from.

  num_of_answers*
                    Integer (required)
                    Example: 2

                    Number of answers the service should provide to the question
                    that was asked.

  }
```

Sample Payload

Below is a sample input of unstructured text for the Q&A API:

Around 930 staff work at the SNB. Our offices in Zurich and Berne are situated at prime locations in the heart of the city, and offer modern working environments with the latest technology.

The Swiss National Bank is committed to ensuring that men and women receive equal pay for work of equal value. This has been confirmed by the equal pay analysis required by law, which was carried out by the University of St. Gallen's Competence Centre for Diversity & Inclusion (CCDI) using the federal government's standard analysis tool 'Logib'. On the basis of its positive result, the SNB was awarded the 'We Pay Fair' label, and it has also received industry certification from the social partners' centre for equal pay in the banking industry ('SF LoBa').

The findings of the legally required equal pay analysis are pleasing. They confirm the internal analyses previously carried out as part of the annual pay round. The SNB seeks to ensure that every employee is paid according to their role and thus irrespective of their gender. We will therefore continue to advocate equal pay in the workplace.

A request payload for the sample text is constructed as a JSON-formatted query that can be saved as a file (qna.json):

```
{
  "data": {
    "input": " Around 930 staff work at the SNB. Our offices in Zurich and Berne are situated at prime locations in the heart of the city, and offer modern working environments with the latest technology. The Swiss National Bank is committed to ensuring that men and women receive equal pay for work of equal value. This has been confirmed by th
```

```
e equal pay analysis required by law, which was carried out by the University of St. Gallen's Competence Centre for Diversity & Inclusion (CCDI) using the federal government's standard analysis tool 'Logib'. On the basis of its positiv e result, the SNB was awarded the 'We Pay Fair' label, and it has also received industry certification from the social p artners' centre for equal pay in the banking industry ('SF LoBa'). The findings of the legally required equal pay analy sis are pleasing. They confirm the internal analyses previously carried out as part of the annual pay round. The SNB seeks to ensure that every employee is paid according to their role and thus irrespective of their gender. We will the refore continue to advocate equal pay in the workplace.",
    "num_of_answers": 2,
    "question": "What is the employee count?"
  },
  "meta": {}
}
```


c. Sending a Sample Request to the Q&A API

Below is a command using cURL to send a POST request to the service using the sample input payload from a JSON file (qna.json).

```
curl -s https://api.factset.com/cognitive/nlp/v1/qna/answers -d@qna.json
-H "Content-Type: application/json"
```

d. Output Response from the Q&A API

In response to a valid request, the Q&A API **answers** endpoint service returns a unique ID {qna.id} from the request in JSON format:

Response

```
{
```

```
started_at          string<date-time>
                    example: 2022-10-31T16:31:30.423086
                    The date/time when the task started in UTC.

status              string
                    example: queued
                    Status of the task.
id                  String<uuid>
                    example: ecdf284b-2db8-4499-898e-8d6b59f28d2a
```

```
                    Unique ID to be used to retrieve the answers
                    extracted from the text.
}
```

Sample Response from the Q&A API

For the given sample payload, the Q&A API will return a unique ID in the following format:

```
{
  "data": {
    "id": " ecdf284b-2db8-4499-898e-8d6b59f28d2a",
    "started_at": null,
    "status": "queued"
  },
  "meta": {}
}
```

e. Using the Unique ID to Determine the Status of the Request

Most documents will be processed immediately, but longer documents may take a few seconds to process. The average processing time is about one second for every 1,000 words sent to the service. To determine whether or not the job has finished, use cURL to send a GET request to the **status** endpoint using the unique ID.

```
curl --request GET \
  --url https://api.factset.com/cognitive/nlp/v1/qna/answers/ecdf284b-2db8-4499-898e-8d6b59f28d2a/status \
  --header 'Content-Type: application/json'
```

Sample Response from the Status Endpoint

For the given unique ID, the **status** endpoint will return the current status of the job. When the job reaches a status of “finished”, the answers are ready to be retrieved using the **ID** endpoint.

Please Note: For security reasons, results expire within 500 seconds of being generated. Please retrieve your results using the **ID** endpoint as soon as possible.

```
{
  "data": {
    "id": "ecdf284b-2db8-4499-898e-8d6b59f28d2a",
    "started_at": "2022-11-09T16:03:33.379729",
```

```
    "status": "finished"
  },
  "meta": {}
}
```


f. Using the Unique ID to Retrieve Answers

Once the status of the job is “finished”, use cURL to send a GET request to the **ID** endpoint using the unique ID.

```
curl --request GET \
  --url https://api.factset.com/cognitive/nlp/v1/qna/answers/ecdf284b-2db8-4499-898e-8d6b59f28d2a \
  --header 'Content-Type: application/json'
```

Sample Response from the ID Endpoint

For the given unique ID, the **ID** endpoint will return the number of answers requested along with a confidence score for each answer:


```
{
  "data": [
    {
      "answer": "930 staff work at SNB",
      "score": 0.8831543922424316
    },
    {
      "answer": "930",
      "score": 0.705894410610199
    }
  ],
  "meta": {}
}
```

3. A Sample Client for the Q&A API

Sample Python code for a simple client utilizing the FactSet Q&A service

```
import time

import requests
from requests.auth import HTTPBasicAuth

auth = HTTPBasicAuth('user-name', '<password>')

def extract_from_text(doc_text,
                      fds_qna_url='https://api.factset.com/cognitive/nlp/v1/qna/answers'):

    payload = {'text': doc_text}
    resp = requests.post(f"{fds_qna_url}/summary",
                         json=payload, auth=auth)
    if not resp:
        status_code = resp.status_code if (resp is not None) else -1
        raise ValueError(
            f'Received unexpected response from service: status_code: {status_code}')

    result_id = resp.json()["data"]["id"]

    while(True):
        resp = requests.get(
            f"{fds_qna_url}/{result_id}", auth=auth)
        if not resp:
            status_code = resp.status_code if (resp is not None) else -1
            raise ValueError(
                f'Received unexpected response from service: status_code: {status_code}')
        if resp.json() != 'Processing':
            return resp.json()
        time.sleep(1)
```