

JSON Validator Tool

Introduction

In systems that rely on structured JSON data, such as IoT platforms, manufacturing environments, or data ingestion pipelines, ensuring the accuracy and integrity of incoming data is critical. However, issues like schema violations, incorrect data types, misspelled keys, or improper timestamps can lead to downstream errors, data loss, or misinterpretation.

There is a need for a robust validation mechanism that not only checks JSON data against a predefined schema but also:

- Detects and suggests corrections for minor key spelling errors
- Verifies timestamp consistency (e.g., ensuring 13-digit millisecond format)
- Confirms the structural integrity of nested arrays and objects like device measurements.

Scope of the Project

JSON Validator Tool is meant for internal teams and stakeholders participating in the development, QA, and release processes of Sensoyo's data systems. It ensures that data flowing into the system is properly structured, valid, and conforms to organizational standards.

Specifically, the tool is designed to:

1. **Validate JSON Data Against Sensoyo Schema:** It confirms that the received JSON structure conforms strictly to the known schema definition, such as required fields, types, nested structures, and enumerations
2. **Detecting and Proposing Fixes for Typing Errors in JSON Keys:** Employing fuzzy matching algorithms, it goes through the keys of the JSON files and checks them against the expected list of key names. When a mismatch occurs (e.g., due to a typo), it provides suggestions to enable the developer to fix the problem
3. **Identify Missing or Invalid Fields:** It verifies whether all the required fields are available and if the data types and values of such fields are valid. For instance, it raises an error alert if a field intended to hold an integer has a string value or if a mandatory field is missing in its entirety
4. **Give Line-Specific Feedback to Developers:** Every error or warning produced contains a reference to its precise location in the JSON file, allowing developers to easily find and correct issues quickly
5. **Ensure Uniformity and Cleanliness of Data:** This tool functions as a gatekeeper. This minimizes the chances of consuming faulty, missing, or malformed data and ensures consistency among datasets

What the Tool Will Check For

Sr. No.	Check Performed	Description
1.	JSON validity	Is the input a valid JSON object?
2.	Required fields	Are all necessary keys like deviceList, mfg, etc. present?
3.	Correct data types	Are values of expected types (string, integer, array)?
4.	Timestamp format	Is timestamp a 13-digit integer (epoch ms)?
5.	Spelling suggestions	If a key is misspelled, the tool will suggest the correct one
6.	Nested structure checks	Ensures deviceList and measurements are well-formed
7.	All errors shown together	Outputs full list of errors, not just the first
8.	Line/Field reference	Indicates where in the JSON the error occurred (as key path)

Project Dependencies and Licensing Information

Sr. No.	Dependencies	Command	Version	License/Doc Link
1.	Python	Install via https://www.python.org/downloads And configure the Path in environment variables	3.13.2	https://github.com/python/cpython?tab=License-1-ov-file
2.	JSON	(Build-in python module)	3.13.2	https://docs.python.org/3/library/json.html
3.	jsonschema	pip install jsonschema	4.23.0	https://github.com/jsonrainbow/json-schema/blob/master/LICENSE
4.	difflib	(Build-in python module)	3.13.2	https://docs.python.org/2.7/library/difflib.html

How to Use the JSON Validator Tool

1. For Windows:

a. Download the Required Files

- Download the zip file for JSON_Validator.zip
- Zip file: Contains tool.py (python script for validation) and json_input.txt (text file for sample input)
- Unzip the file by below command:
`tar -xf JSON_Validator.zip -C targetfolder`

```
PS C:\Users\SONAM\Desktop\Internship\Sensoyo> mkdir JSON_Validator

Directory: C:\Users\SONAM\Desktop\Internship\Sensoyo

Mode                LastWriteTime         Length Name
----                -
d-----          07-04-2025      12:37             JSON_Validator

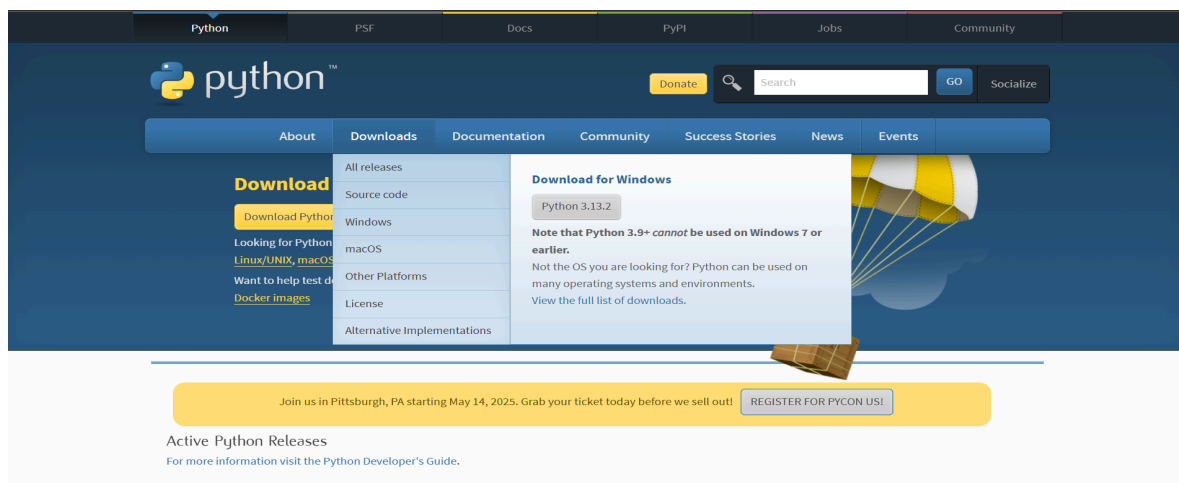
PS C:\Users\SONAM\Desktop\Internship\Sensoyo> tar -xf JSON_Validator.zip -C JSON_Validator
```

- Make sure you have the following files in the same directory
- tool.py → (Contains the validation logic)
- json_input.txt → (Contains the input to be checked for validation)

b. Set Up Your Environment

i. Install Python (if not already installed)

- Go to the official Python website: <https://www.python.org/downloads>
- Click the “Download Python 3.x.x” button



- Open the downloaded .exe file
- Important: Check the box "Add Python to PATH" at the bottom of the installer window
- Click "Install Now"
- Verify Installation: Open Command Prompt and enter the command
`python --version`

```
C:\Users\SONAM>python --version
Python 3.13.2
```

ii. Install Required Python Packages

- `pip install jsonschema`

```
PS C:\Users\SONAM\Desktop\Internship\Sensoyo> pip install jsonschema
Requirement already satisfied: jsonschema in c:\users\sonam\appdata\local\programs\python\python313\lib\site-packages (4.23.0)
Requirement already satisfied: attrs>=22.2.0 in c:\users\sonam\appdata\local\programs\python\python313\lib\site-packages (from jsonschema) (25.3.0)
Requirement already satisfied: jsonschema-specifications>=2023.03.6 in c:\users\sonam\appdata\local\programs\python\python313\lib\site-packages (from jsonschema) (2024.10.1)
Requirement already satisfied: referencing>=0.28.4 in c:\users\sonam\appdata\local\programs\python\python313\lib\site-packages (from jsonschema) (0.36.2)
Requirement already satisfied: rpds-py>=0.7.1 in c:\users\sonam\appdata\local\programs\python\python313\lib\site-packages (from jsonschema) (0.24.0)

[notice] A new release of pip is available: 24.3.1 -> 25.0.1
[notice] To update, run: python.exe -m pip install --upgrade pip
PS C:\Users\SONAM\Desktop\Internship\Sensoyo>
```

- Note: Diffliib library is an in-built library but if the if the script throws an error for not recognizing it, then execute below command
`pip install diffliib`

c. Run the Tool Using Sample Inputs

- Open tool.py
- Run the script directly by below command:
`python tool.py`

```
PROBLEMS  OUTPUT  DEBUG CONSOLE  TERMINAL  POSTMAN CONSOLE

PS C:\Users\SONAM\Desktop\Internship\Sensoyo> python tool.py
Enter the path of the JSON file: C:\Users\SONAM\Desktop\Internship\Sensoyo\json input.txt
```

- When prompted, paste the path to your JSON file: Enter the path to the JSON file: `/path/to/your/input.json`
- The tool will read the file, validate its contents, and show results in the terminal

2. For Linux:

a. Download the Required Files

- Download the zip file for JSON_Validator.zip
- Zip file: Contains tool.py (python script for validation) and json_input.txt (text file for sample input)
- Unzip the file by below command:
`unzip JSON_Validator.zip`

```
leapscale@leapscale:~/Downloads/Json_Validator$ unzip JSON_Validator.zip
Archive:  JSON_Validator.zip
  inflating: tool.py
  inflating: tool_main.py
```

- Make sure you have the following files in the same directory
- tool.py → (Contains the validation logic)
- json_input.txt → (Contains the input to be checked for validation)

b. Set Up Your Environment

i. Install Python (if not already installed)

- Most Linux distributions come with Python pre-installed. To check if Python is already installed, open a terminal and run:
`python3 --version`

```
leapscale@leapscale:~/Downloads/Json_Validator$ python3 --version
Python 3.10.4
```

- If it's not installed, install it using the package manager for your distribution:
`sudo apt install python3 python3-pip`

```
rare@rare-All-Series:~$ sudo apt install python3 python3-pip
[sudo] password for rare:
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
python3 is already the newest version (3.10.6-1~22.04.1).
python3 set to manually installed.
The following additional packages will be installed:
  javascript-common libexpat1-dev libjs-jquery libjs-sphinxdoc
  libjs-underscore libpython3-dev libpython3.10-dev python3-dev
  python3-distutils python3-setuptools python3-wheel zlib1g-dev
Suggested packages:
  apache2 | lighttpd | httpd python-setuptools-doc
The following NEW packages will be installed:
  javascript-common libexpat1-dev libjs-jquery libjs-sphinxdoc
  libjs-underscore libpython3-dev libpython3.10-dev python3-dev
  python3-distutils python3-pip python3-setuptools python3-wheel
  python3.10-dev zlib1g-dev
0 upgraded, 14 newly installed, 0 to remove and 55 not upgraded.
Need to get 8,015 kB of archives.
After this operation, 34.2 MB of additional disk space will be used.
Do you want to continue? [Y/n] y
Get:1 http://in.archive.ubuntu.com/ubuntu jammy/main amd64 javascript-common all 11+nmu1 [5,936 B]
Get:2 http://in.archive.ubuntu.com/ubuntu jammy-updates/main amd64 libexpat1-dev amd64 2.4.7-1ubuntu0.5 [147 kB]
Get:3 http://in.archive.ubuntu.com/ubuntu jammy/main amd64 libjs-jquery all 3.6.0.dfsg+~1-2 [221 kB]
Get:4 http://in.archive.ubuntu.com/ubuntu jammy/main amd64 libjs-sphinxdoc all 3.6.0.dfsg+~1-2 [2,211 kB]
Get:5 http://in.archive.ubuntu.com/ubuntu jammy/main amd64 libjs-underscore all 1.13.4-1 [11.1 kB]
Get:6 http://in.archive.ubuntu.com/ubuntu jammy/main amd64 libpython3-dev amd64 3.10.6-1~22.04.1 [1,111 kB]
Get:7 http://in.archive.ubuntu.com/ubuntu jammy/main amd64 libpython3.10-dev amd64 3.10.6-1~22.04.1 [4,111 kB]
Get:8 http://in.archive.ubuntu.com/ubuntu jammy/main amd64 python3-dev amd64 3.10.6-1~22.04.1 [1,111 kB]
Get:9 http://in.archive.ubuntu.com/ubuntu jammy/main amd64 python3-distutils amd64 3.10.6-1~22.04.1 [111 kB]
Get:10 http://in.archive.ubuntu.com/ubuntu jammy/main amd64 python3-setuptools amd64 65.5.0-1 [391 kB]
Get:11 http://in.archive.ubuntu.com/ubuntu jammy/main amd64 python3-wheel amd64 0.40.0-1 [31.1 kB]
Get:12 http://in.archive.ubuntu.com/ubuntu jammy/main amd64 python3-pip amd64 23.0.1+repack-1 [1,111 kB]
Get:13 http://in.archive.ubuntu.com/ubuntu jammy/main amd64 python3.10-dev amd64 3.10.6-1~22.04.1 [4,111 kB]
Get:14 http://in.archive.ubuntu.com/ubuntu jammy/main amd64 zlib1g-dev amd64 1:1.2.13.dfsg-1 [111 kB]
Fetched 8,015 kB in 2s (4,008 kB/s)
debconf: delaying package configuration, since apt-utils is not installed
Selecting previously unselected package javascript-common.
(Reading database ... 123456 files and directories currently installed.)
Preparing to unpack .../0-javascript-common_11+nmu1_all.deb ...
Unpacking javascript-common (11+nmu1) ...
Selecting previously unselected package libexpat1-dev.
Preparing to unpack .../1-libexpat1-dev_2.4.7-1ubuntu0.5_amd64.deb ...
Unpacking libexpat1-dev (2.4.7-1ubuntu0.5) ...
Selecting previously unselected package libjs-jquery.
Preparing to unpack .../2-libjs-jquery_3.6.0.dfsg+~1-2_all.deb ...
Unpacking libjs-jquery (3.6.0.dfsg+~1-2) ...
Selecting previously unselected package libjs-sphinxdoc.
Preparing to unpack .../3-libjs-sphinxdoc_3.6.0.dfsg+~1-2_all.deb ...
Unpacking libjs-sphinxdoc (3.6.0.dfsg+~1-2) ...
Selecting previously unselected package libjs-underscore.
Preparing to unpack .../4-libjs-underscore_1.13.4-1_all.deb ...
Unpacking libjs-underscore (1.13.4-1) ...
Selecting previously unselected package libpython3-dev.
Preparing to unpack .../5-libpython3-dev_3.10.6-1~22.04.1_amd64.deb ...
Unpacking libpython3-dev (3.10.6-1~22.04.1) ...
Selecting previously unselected package libpython3.10-dev.
Preparing to unpack .../6-libpython3.10-dev_3.10.6-1~22.04.1_amd64.deb ...
Unpacking libpython3.10-dev (3.10.6-1~22.04.1) ...
Selecting previously unselected package python3-dev.
Preparing to unpack .../7-python3-dev_3.10.6-1~22.04.1_amd64.deb ...
Unpacking python3-dev (3.10.6-1~22.04.1) ...
Selecting previously unselected package python3-distutils.
Preparing to unpack .../8-python3-distutils_3.10.6-1~22.04.1_amd64.deb ...
Unpacking python3-distutils (3.10.6-1~22.04.1) ...
Selecting previously unselected package python3-setuptools.
Preparing to unpack .../9-python3-setuptools_65.5.0-1_amd64.deb ...
Unpacking python3-setuptools (65.5.0-1) ...
Selecting previously unselected package python3-wheel.
Preparing to unpack .../10-python3-wheel_0.40.0-1_amd64.deb ...
Unpacking python3-wheel (0.40.0-1) ...
Selecting previously unselected package python3-pip.
Preparing to unpack .../11-python3-pip_23.0.1+repack-1_amd64.deb ...
Unpacking python3-pip (23.0.1+repack-1) ...
Selecting previously unselected package python3.10-dev.
Preparing to unpack .../12-python3.10-dev_3.10.6-1~22.04.1_amd64.deb ...
Unpacking python3.10-dev (3.10.6-1~22.04.1) ...
Selecting previously unselected package zlib1g-dev.
Preparing to unpack .../13-zlib1g-dev_1%3A1.2.13.dfsg-1_amd64.deb ...
Unpacking zlib1g-dev (1:1.2.13.dfsg-1) ...
Setting up javascript-common (11+nmu1) ...
Setting up libexpat1-dev (2.4.7-1ubuntu0.5) ...
Setting up libjs-jquery (3.6.0.dfsg+~1-2) ...
Setting up libjs-sphinxdoc (3.6.0.dfsg+~1-2) ...
Setting up libjs-underscore (1.13.4-1) ...
Setting up libpython3-dev (3.10.6-1~22.04.1) ...
Setting up libpython3.10-dev (3.10.6-1~22.04.1) ...
Setting up python3-dev (3.10.6-1~22.04.1) ...
Setting up python3-distutils (3.10.6-1~22.04.1) ...
Setting up python3-setuptools (65.5.0-1) ...
Setting up python3-wheel (0.40.0-1) ...
Setting up python3-pip (23.0.1+repack-1) ...
Setting up python3.10-dev (3.10.6-1~22.04.1) ...
Setting up zlib1g-dev (1:1.2.13.dfsg-1) ...
```

- Verify Installation: Open Command Prompt and enter the command
`python3 --version`

ii. Install Required Python Packages

- **pip3 install jsonschema**

```
leapscale@leapscale:~/Downloads/Json_Validator$ pip install jsonschema
Defaulting to user installation because normal site-packages is not writeable
Collecting jsonschema
  Downloading jsonschema-4.23.0-py3-none-any.whl (88 kB)
    88.5/88.5 KB 1.5 MB/s eta 0:00:00
Collecting attrs>=22.2.0
  Downloading attrs-25.3.0-py3-none-any.whl (63 kB)
    63.8/63.8 KB 1.2 MB/s eta 0:00:00
Collecting rpds-py>=0.7.1
  Downloading rpds_py-0.24.0-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (389 kB)
    389.5/389.5 KB 864.5 kB/s eta 0:00:00
Collecting referencing>=0.28.4
  Downloading referencing-0.36.2-py3-none-any.whl (26 kB)
Collecting jsonschema-specifications>=2023.03.6
  Downloading jsonschema_specifications-2024.10.1-py3-none-any.whl (18 kB)
Collecting typing-extensions>=4.4.0
  Downloading typing_extensions-4.13.1-py3-none-any.whl (45 kB)
    45.7/45.7 KB 729.0 kB/s eta 0:00:00
Installing collected packages: typing-extensions, rpds-py, attrs, referencing, jsonschema-specifications, jsonschema
WARNING: The script jsonschema is installed in '/home/leapscale/.local/bin' which is not on PATH.
Consider adding this directory to PATH or, if you prefer to suppress this warning, use --no-warn-script-location.
Successfully installed attrs-25.3.0 jsonschema-4.23.0 jsonschema-specifications-2024.10.1 referencing-0.36.2 rpds-py-0.24.0 typing-extensions-4.13.1
```

- Note: Difflib library is an in-built library but if the if the script throws an error for not recognizing it, then execute below command

pip3 install difflib

c. Run the Tool Using Sample Inputs

- Open tool.py
- Run the script directly by below command:

python3 tool.py

```
leapscale@leapscale:~/Downloads/Json_Validator$ python3 tool.py
Enter the path of the JSON file: /home/leapscale/Downloads/json_input.txt

Validation Results:
=====
JSON is valid.
```

- When prompted, paste the path to your JSON file: Enter the path to the JSON file: **/path/to/your/input.json**
- The tool will read the file, validate its contents, and show results in the terminal

Output Cases

Input file - data.json(wrong input):

```
data = {
  "mgf": "Interns Leap&Scale",
  "devList": [
    {
      "deviceId": "IN03091228",
      "deviceType": "IN_005",
      "mode": "Auto",
      "timestamp": 17428600274000,
      "createdT": 1742860274000,
      "status": "Okay",
      "measure": {
        "1":{
          "paramName": "Temperature",
          "value": 31.1,
          "multip": 1,
          "unit": "celsius"
        },
        "2":{
          "paramName": "Humidity",
          "val": 27.9,
          "multiplier": "1",
          "unit": "Percentage"
        }
      }
    }
  ]
}
```

Output:

```
PROBLEMS  OUTPUT  DEBUG CONSOLE  TERMINAL  POSTMAN CONSOLE

PS C:\Users\SONAM\Desktop\Internship\Sensoyo> python tool.py
Enter the path of the JSON file: C:\Users\SONAM\Desktop\Internship\Sensoyo\json input.txt

Validation Results:
=====
line 2: mgf is spelled wrongly, change spelling to mfg
line 3: devList is spelled wrongly, change spelling to devicelist
Schema Error: root - 'mfg' is a required property
Schema Error: root - 'devicelist' is a required property
PS C:\Users\SONAM\Desktop\Internship\Sensoyo> █
```

Input file - data.json (correct input):

```
data = {
  "mfg": "Interns Leap&Scale",
  "deviceList": [
    {
      "deviceId": "IN03091228",
      "deviceType": "IN_005",
      "mode": "Auto",
      "timestamp": 1742860274000,
      "createdTime": 1742860274000,
      "status": "Okay",
      "measurements": [
        {
          "paramName": "Temperature",
          "value": 31.1,
          "multiplier": 1,
          "unit": "celsius"
        },
        {
          "paramName": "Humidity",
          "value": 27.9,
          "multiplier": 1,
          "unit": "Percentage"
        }
      ]
    }
  ]
}
```

Output:

```
PROBLEMS  OUTPUT  DEBUG CONSOLE  TERMINAL  POSTMAN CONSOLE

PS C:\Users\SONAM\Desktop\Internship\Sensoyo> python tool.py
Enter the path of the JSON file: C:\Users\SONAM\Desktop\Internship\Sensoyo\json input.txt

Validation Results:
=====
JSON is valid.
PS C:\Users\SONAM\Desktop\Internship\Sensoyo> 
```