

OTC Website Find Technologies Contact Us

Request Information Permalink

# Streaming JSON Data With Bit-Parallel Fast-Forwarding

Tech ID: 33334 / UC Case 2023-977-0

#### **FULL DESCRIPTION**

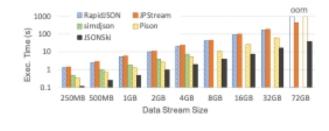
#### **Background**

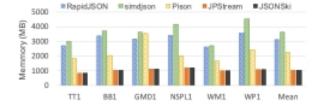
Semi-structured data is a hybrid data type that possesses some level of structure in the form of tags, labels, etc. that help identify and categorize the data. Common semi-structured data formats include XML, JSON, CSV. There are two typical processing schemes for querying semi-structured data - preprocessing schemes and streaming schemes. Preprocessing schemes first parse the data into an in-memory structure before querying - therefore, require a large memory footprint. Streaming schemes locate and extract the data on the fly and thereby reduce the memory footprint - though they still need to parse the entire data stream character by character which limits the efficiency of querying.

# Technology

Prof. Zhijia Zhao and his team have developed a novel querying framework for streaming JSON data called JSONSki - by considering the possibility and practicality of fast-forwarding the data stream to accelerate the query processing. JSONSki offers a set of application programming interfaces (APIs) that can be invoked during the streaming to dynamically fast-forward over different cases of irrelevant substructures - irrelevant to the query.

# Images





The images above present a comparison of JSONSki performance with state-of-the-art techniques - (top) execution time of different solutions as a function of data stream size; (bottom) memory footprint comparison.

# **ADVANTAGES**

- ▶ By fast-forwarding irrelevant data segments, JSONSki can query the streaming JSON data much faster than state-of-the-art techniques.
- ► Minimal memory footprint
- ▶ Simplified parser design.
- ▶ Offers a set of APIs that can be naturally integrated into analytics of streaming data.
- Dutperforms state-of-the-art techniques such as JPStream and simdjson on real-world data including small and large datasets.
- All instructions used by JSONSki are commonly available in modern CPUs.

#### CONTACT

Venkata S. Krishnamurty venkata.krishnamurty@ucr.edu tel:

#### OTHER INFORMATION

#### **KEYWORDS**

json, parser, semi-structured data, bitparallel algorithm, data analytics, streaming data, jsonski

#### **CATEGORIZED AS**

**▶** Computer

Software

RELATED CASES

2023-977-0

JSON data querying is valuable in a wide variety of industries including e-commerce, healthcare, manufacturing, media & entertainment and retail - to name a few.

## **STATE OF DEVELOPMENT**

Working prototype that demonstrates the effectiveness and efficiency of the invention - tested on real-world datasets. Working prototype evaluated against state-of-the-art solution.

#### **INVENTOR INFORMATION**

- ▶ Please see recent press coverage of Prof. Zhao.
- ▶ JSONSki was awarded the best paper award at the recent ASPLOS 2022 Conference.
- Please visit Prof. Zhao profile page to learn more about his research.

## **RELATED MATERIALS**

▶ JSONSki: Streaming Semi-structured Data with Bit-Parallel Fast-Forwarding

## **PATENT STATUS**

Patent Pending

University of California, Riverside

Office of Technology Commercialization

200 University Office Building,

Riverside, CA 92521

otc@ucr.edu

research.ucr.edu/

Terms of use | Privacy Notice | © 2023, The Regents of the University of California