



Optimisasi Prompt LLM: Migrasi dari JSON ke TOON format

Lombok Dev Meetup 13

Perkenalan



ABU AMMAR

Software Engineer

Specializing in
Fullstack & AI Automation

Key Experience

- Mobile Developer, PT. Unreal Management Indonesia (Feb 2025 – Sep 2025)
- Modernized legacy codebase, improved performance & stability.
- Front End Developer, PT. Lembaga Sinergi Analitika (Jun 2023 – Dec 2024)
- Translated UI/UX into responsive web apps, integrated with back-end APIs.
- Top 12 Finalist, Amartha x GDG Jakarta Hackathon 2025



Lombok, Indonesia



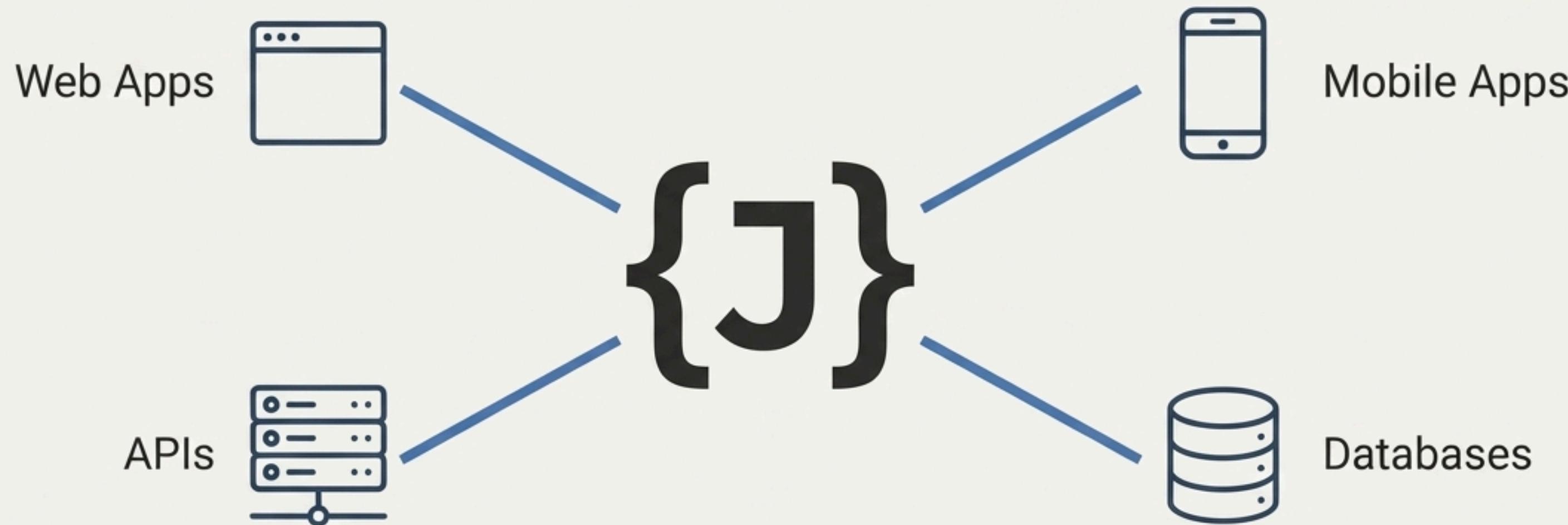
github.com/amaralkaff



amangly.com

LOMBOK{DEV}
Meetup

JSON: The Lingua Franca of the Web



- The undisputed standard for two decades.
- Human-readable, machine-parseable.
- **Optimized for:** Deterministic Parsers.

A New Consumer Enters the Arena: The Tokenizer

```
 Lorem ipsum dolor sit amet,——  
consectet=asf———ins=avun  
-----<FDN nevbio=——  
valority———ABET——  
———[ah d>mesch——i——m29—  
———znhum quals——loce  
amtor———zie;sngone  
reve———bip——tellen-  
———oone=eecrad apals-  
notors]——  
fabattcv kae———lnot!use—  
———esmenz=oru——rds—  
———xent-on——[]——  
slip ss-asptnde———cjp  
ee=d——comout---+—tvu  
———[ib——texavpre  
———lon paixexexe osnat·purur  
pelic]——{  
———a]  
cewetsonet-sodults..
```



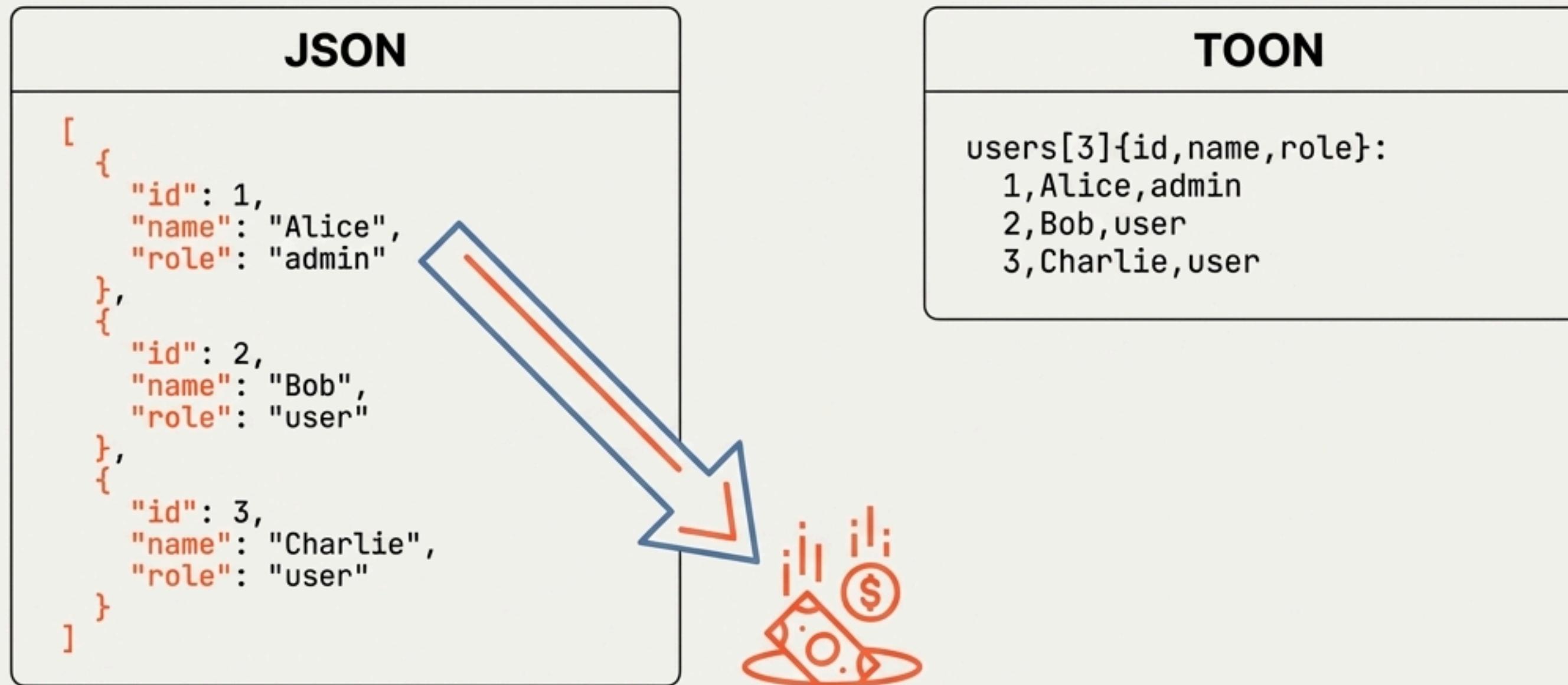
LLMs don't see text; they see tokens.

Every token has a direct API cost.

Every token has a performance cost (latency).

Suddenly, **every character—`{ }[]`,—is a potential tax.**

The 'Token Tax': JSON's Noise is Now a Financial Cost



- **Repetitive Keys:** Sending "name", "role" for every single item.
- **Structural Overhead:** Braces and quotes convey structure, not information.
- This is why 'JSON Prompting' is often called a '*stupid and bad hack*'.

Introducing TOON: Token-Oriented Object Notation

TOON

Designed for the Tokenizer, Not the Parser.

- A lossless, drop-in representation of your JSON.
 - Maximizes **Signal**, minimizes **Syntax**.
 - Combines YAML's readability with CSV's density.

How It Works: Indentation for Objects, Tables for Arrays

```
{ "user": {  
    "name": "Alice"}}
```

```
user:  
  name: Alice
```

```
[  
  {"id": 1, "name": "Alice"},  
  {"id": 2, "name": "Bob"}]
```

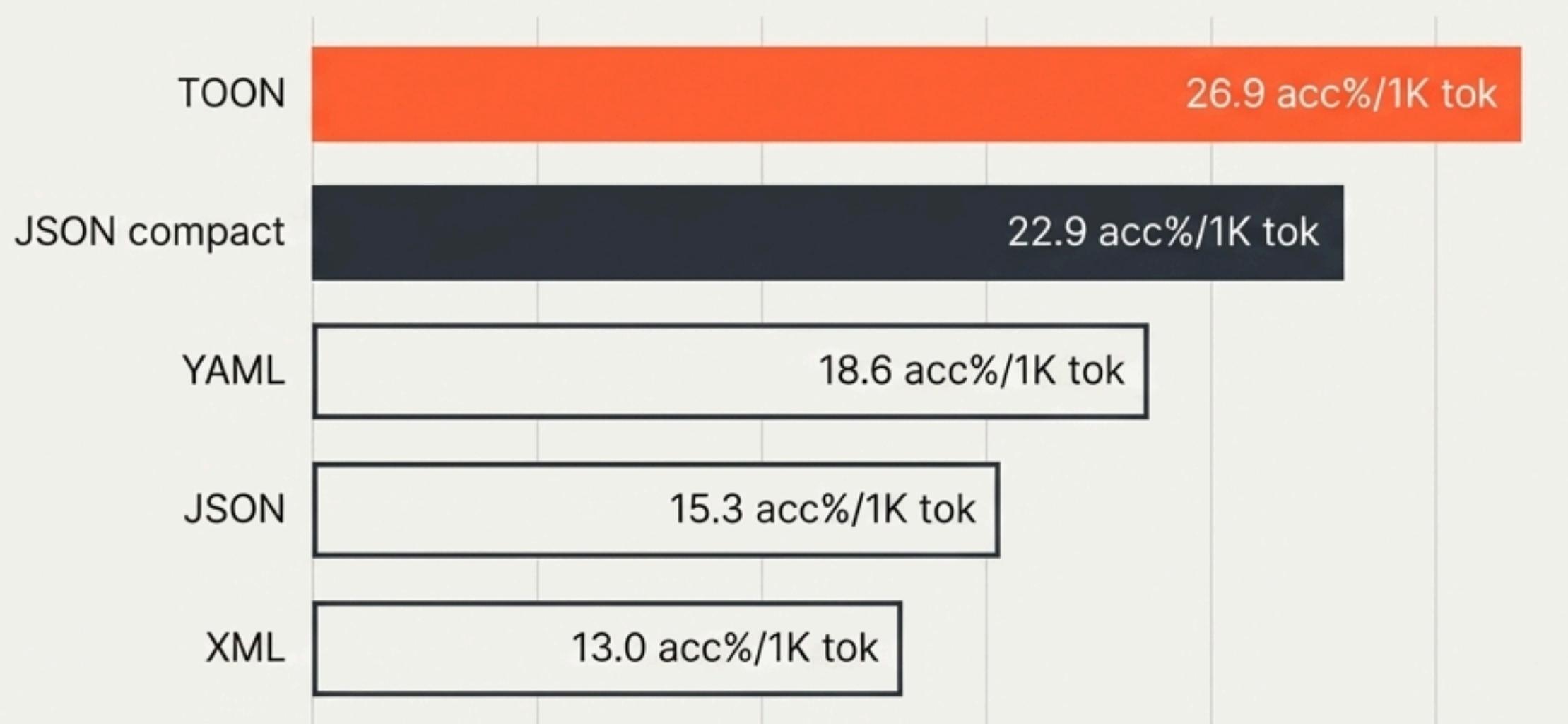
```
users[2]{id,name}:  
1,Alice  
2,Bob
```

Objects become YAML-like: Structure is defined by indentation. No more {}.

Arrays become CSV-like: Keys {...} are declared once. Data follows. This is where the **massive token savings** originate.

The Proof: Up to 40% Fewer Tokens, ~4% Higher Accuracy

Efficiency Ranking (Accuracy per 1K Tokens)



~40%

TOKEN SAVINGS
(vs. formatted JSON)

+4%

RETRIEVAL ACCURACY
(average across models)

The Right Tool for the Right Data Structure



SWEET SPOT (Gunakan TOON!)

- Uniform arrays of objects (database results, API lists).
- Flat or semi-flat data structures.

****Killer Use Case**:** Time-series data → **59% token reduction!**



STICK WITH JSON (Gunakan JSON)

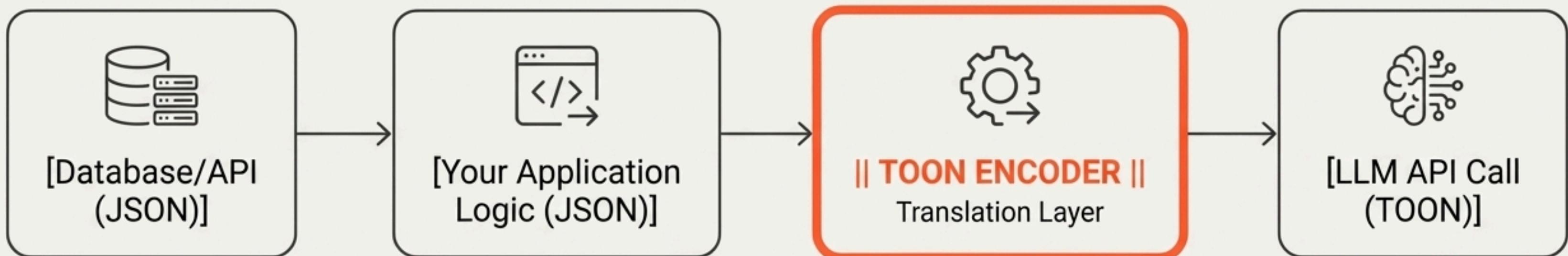
- Deeply nested, non-uniform data (e.g., complex configs).
- Situations where JSON-compact is more token-efficient.
- When your existing pipelines are latency-critical and optimized for JSON.

Analyze Your Savings in Seconds

```
# Analyze token savings instantly
$ npx @toon-format/cli data.json --stats

✓ Encoded data.json → output.toon
ℹ Token estimates: ~15,145 (JSON) → ~8,745 (TOON)
✓ Saved ~6,400 tokens (-42.3%)
```

The Architectural Pattern: Don't Rewrite, Just Translate



1. **Keep JSON** for your database, internal APIs, and business logic.
2. **At the edge**, right before calling the LLM, encode your data to TOON.
3. Reap the cost and performance benefits without disrupting your stack.

Ringkasan



Think in Tokens. In the AI era, the 'Token Economy' is the new reality. Every character is a cost.

Roboto Regular is the AI and AI era, the 'Token Economy' is the new reality. Every character committee that inction models token contents to levit to total cost. Image ov massine savings and shrarete to lots of nonlova value.



TOON is purpose-built for this economy. It offers massive savings and higher accuracy for structured, uniform data.

Roboto Regular TOON is purpose-built for this innable economy. to ensure it benefit and innations, complones sector. It offers massive savings and higher accuracy for structured, uniform data.



Implement as a 'Translation Layer.' Optimize for your LLM at the edge without rewriting your core application logic.

Roboto Regular is in a 'Translation layer.' Optimize for your LLM at the edge — without rewriting your core application logic.

Terima Kasih!

Questions?

*The future of AI integration belongs to those
who use their tokens most wisely.*

Abu Ammar
github.com/amaralkaff
www.amangly.com