Sumo Logic – How to create Field Extraction Rules (FERs) and what to be aware of

dev.classmethod.jp/articles/202212-sumologic-fer-creation-and-notes

佐久間昇吾

December 11, 2022

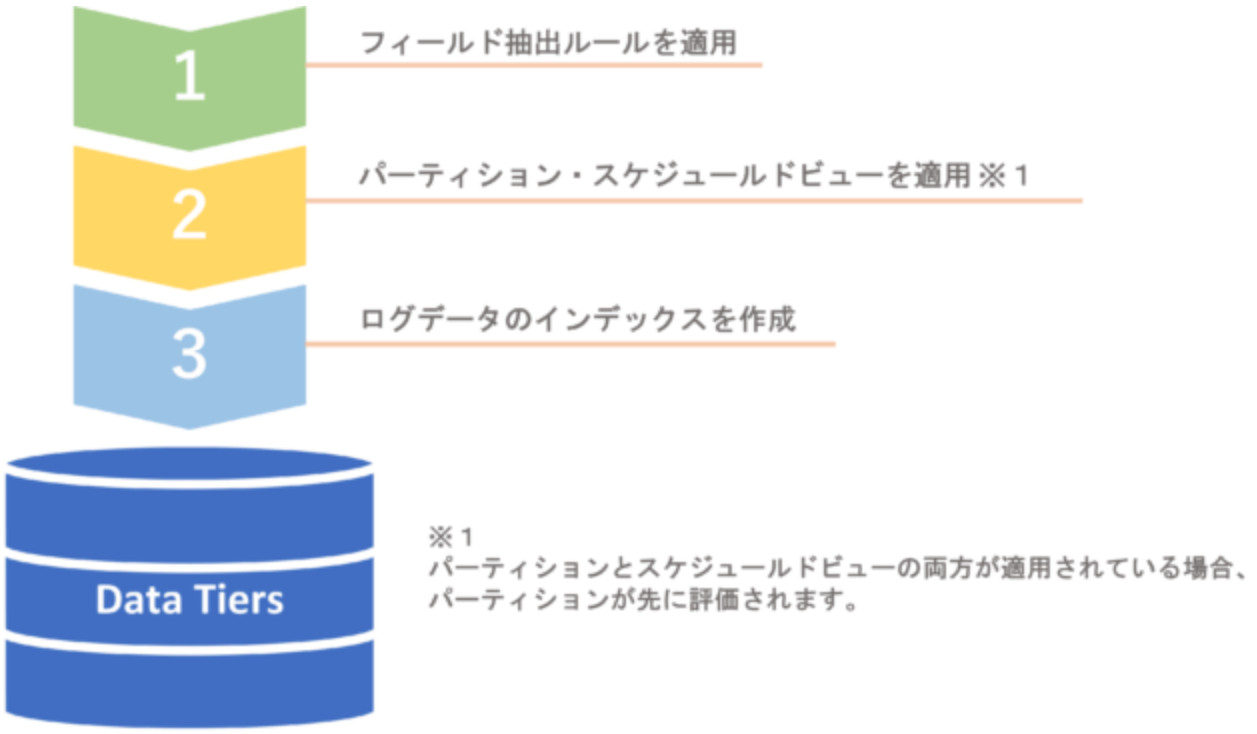
If the content of the article is outdated, please also check the official website.

For more information about Sumo Logic, please see below.

- Sumo Logic official website
- Classmethod - Cloud-native log management and analytics SaaS "Sumo Logic"

First

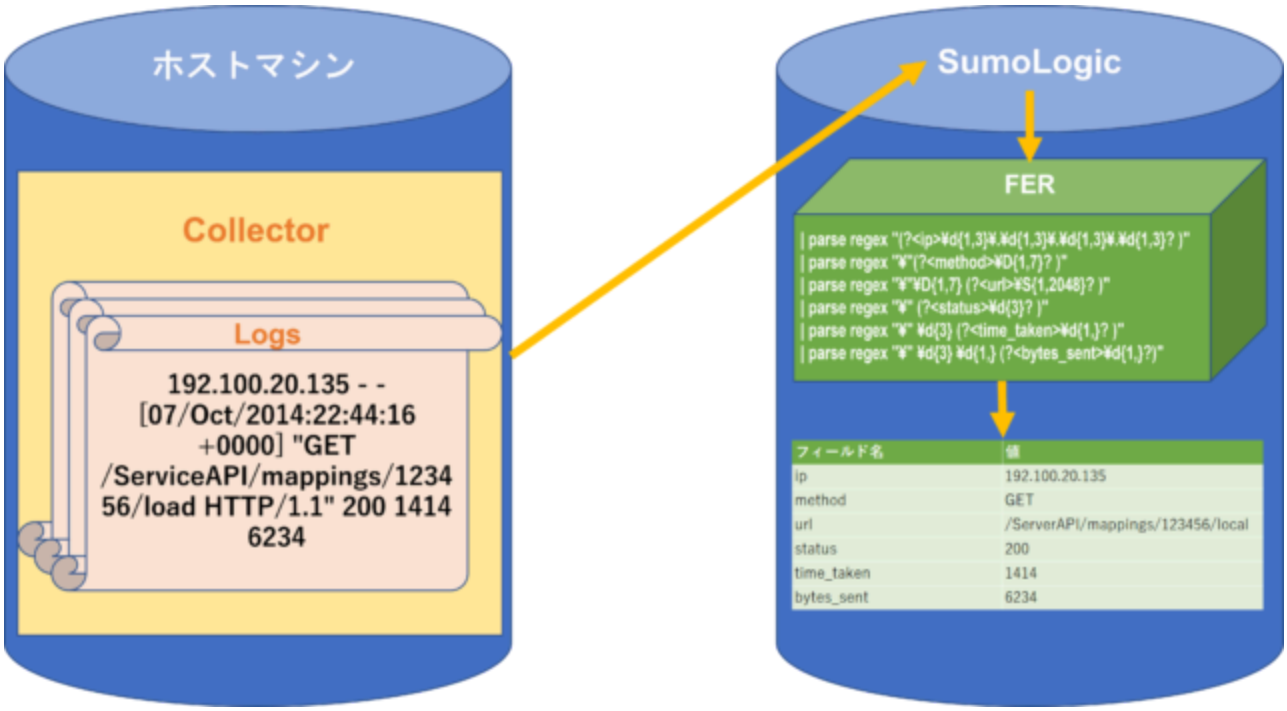
When Sumo Logic receives message data (logs and metrics), it evaluates the data in the following order:



***This time we will introduce the Field Extraction Rules!**

About FER (Field Extraction Rules)

Log messages delivered to Sumo Logic are in data structures such as JSON, XML, CSV, etc.
Field Extraction Rules (FER) parse this raw data and extract only the necessary parts as "field:value" pairs.



The data extracted by FER will be indexed in a later step.

When to create an FER:
The FER applies to message data from the time of creation onwards. Therefore, it is recommended to set it early. Also, when you create or change an FER, the rules are applied immediately.

Points to note when creating an FER

There are some restrictions and best practices for creating an FER:

Roles required to create an FER

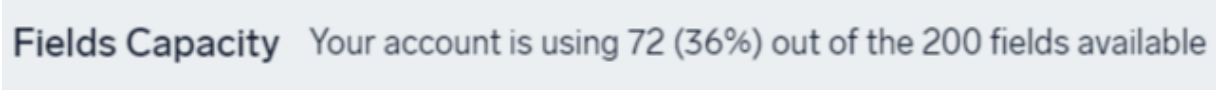
- **Manage field extraction rules**
  Permissions for "Manage Fields," "View Fields," and "View Field Extraction Rules."

FER Limitations

- **The maximum number of FERs is 50.**
  The limit, number of uses, and usage rate are displayed at the bottom of the FER management screen.

  Field Extraction Rules Capacity Your account has 15 (30%) enabled field extraction rules out of 50 available

- **The maximum number of fields that can be specified in an FER is 200.**
  Field limits are per account.
  The limit, number of uses, and usage rate are displayed at the bottom of the Field management screen.



- **Expressions can be up to 16k (16,384) characters long.**

FER Best Practices

- **Include exact keywords**
  Create a narrower scope of the message data itself to identify a subset of data. Defining a broader scope increases the number of fields to parse and increases the chance of extracting fields you don't need.

- **Create and apply multiple rules for different purposes**
  Rather than defining complex rules in a single FER, we recommend creating multiple rules with finely divided granularity. Rules created for clear purposes are easier to manage. For example, the impact of modifications can be minimized.

- **Do not extract unnecessary fields**
  After narrowing the scope, narrow down the fields to be extracted. We recommend extracting only the minimum amount of data necessary.

- **Test the scope before creating the rule**
  Please check whether you can successfully retrieve the data that corresponds to the field in Log Search.

- **Verify that the required fields are present in the defined FER**
  Even if you create a field extraction rule, if the field does not exist, you will not be able to retrieve the data. Please check whether the fields in the extracted message data have been deleted or disabled.

- **Do not target the same field name in the same scope + same message data in multiple FERs**
  In this case, only one FER is randomly applied to the target field name.

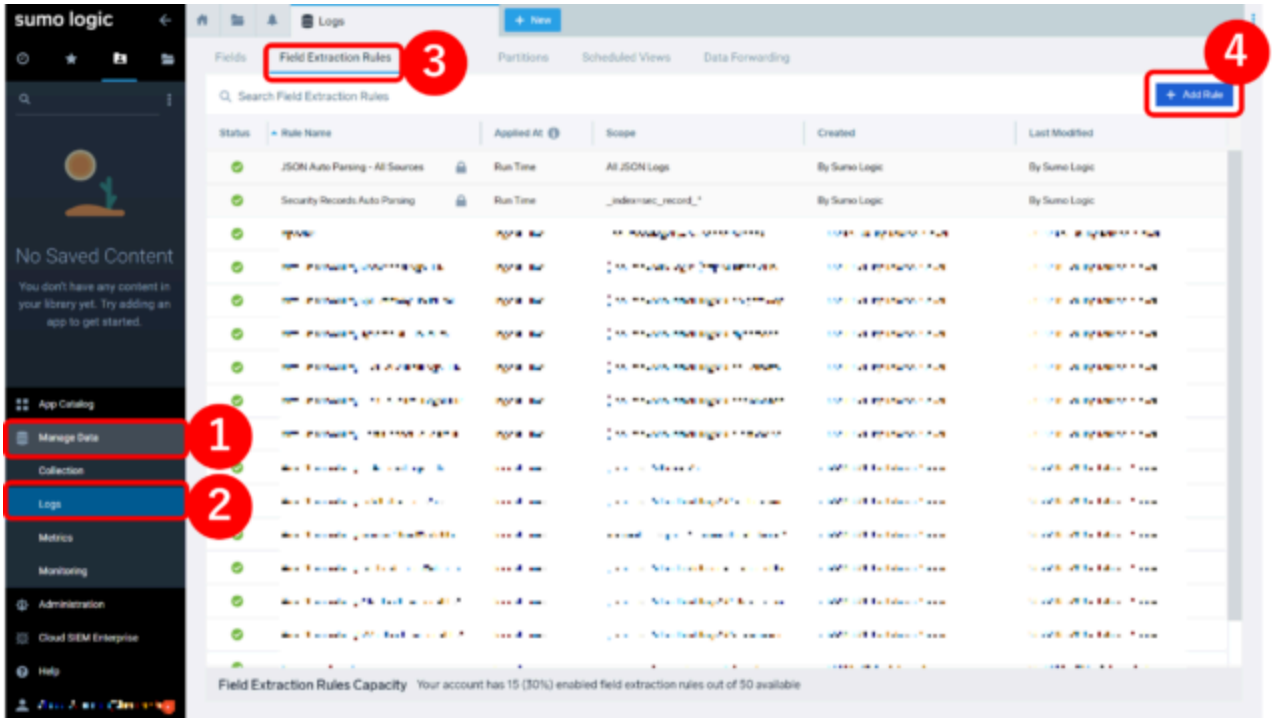
  * It is possible to specify the same field name in multiple FERs for different message data with separate scopes.

- **Use the parse nodrop statement**
  It allows you to filter data that either matches or does not match a specified condition. It can be used in rule expressions as an option for the parse operator.

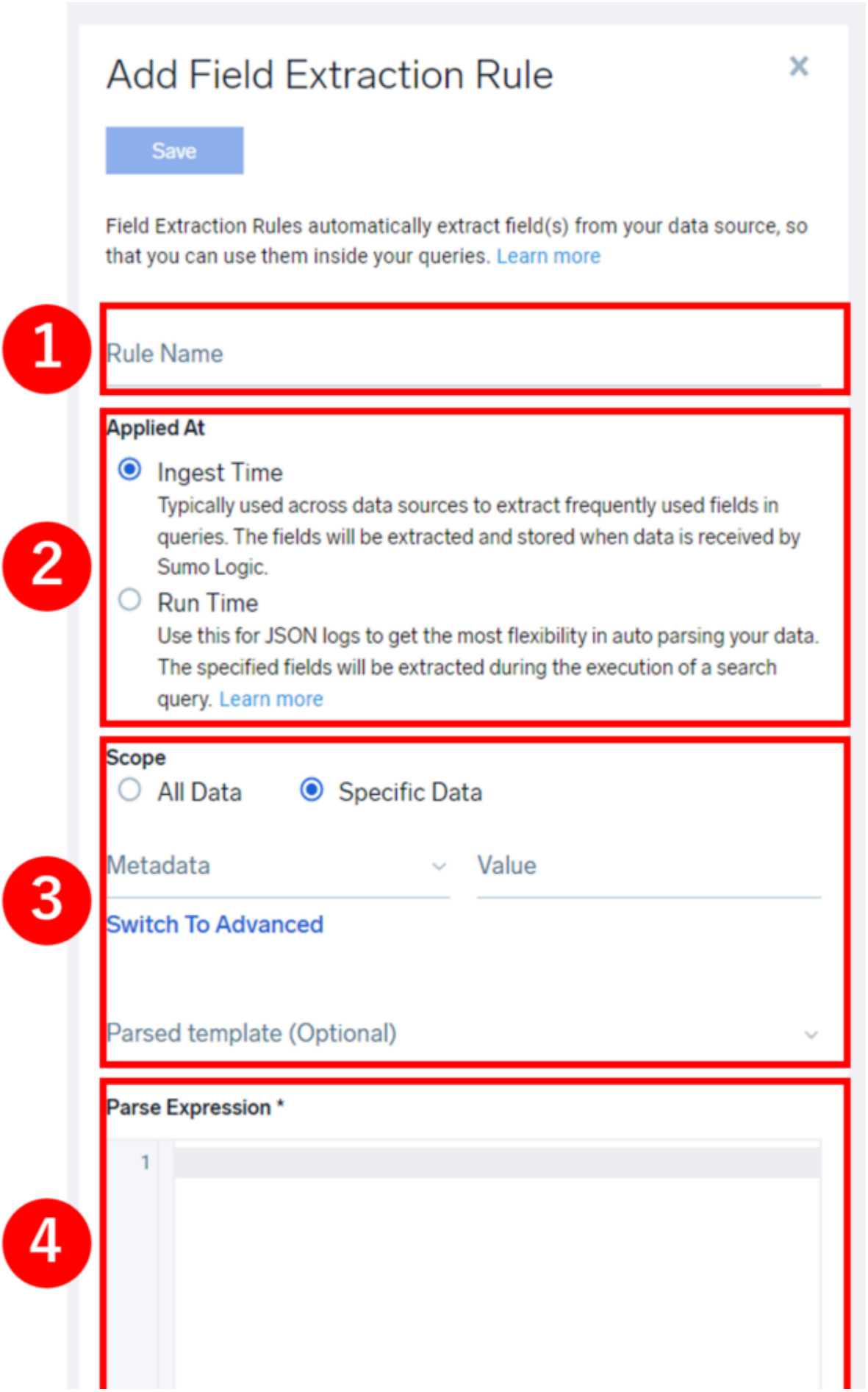
How to create an FER

Follow the steps below to go to the FER creation page:

① Manage Data > ② Logs > ③ Field Extraction Rules > ④ + Add Rule



The FER creation menu below will then appear on the right.


Add Field Extraction Rule
Save
Field Extraction Rules automatically extract field(s) from your data source, so that you can use them inside your queries. Learn more
1
Rule Name
Applied At
2
Ingest Time
Typically used across data sources to extract frequently used fields in queries. The fields will be extracted and stored when data is received by Sumo Logic.
Run Time
Use this for JSON logs to get the most flexibility in auto parsing your data. The specified fields will be extracted during the execution of a search query. Learn more
Scope
All Data
Specific Data
3
Metadata
Value
Switch To Advanced
Parsed template (Optional)
Parse Expression *
1
4

① Rule Name

Specify an easy-to-understand rule name,
as you will create multiple FERs depending on their purpose.

② Applied At

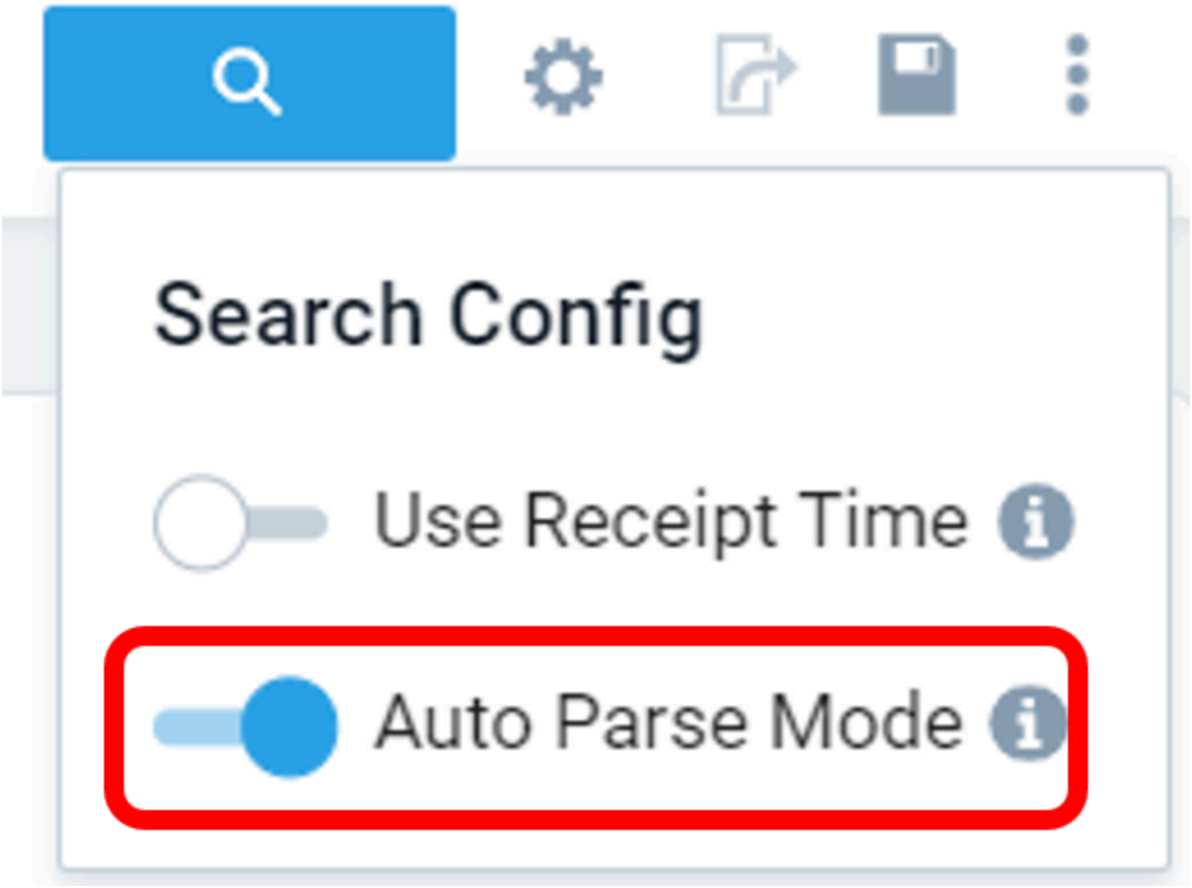
Select one of the following two types of perspective method.

- **Ingest Time**
  Supports any structured data type. Requires writing an expression to parse.
  There is a limit of 50 FERs and 200 fields.
  This applies to message data after the FER is created.

- **Run Time**
  Only JSON data types are automatically parsed.
  There are no restrictions on the number of FERs or fields.
  Only JSON data is parsed when using Auto Parse Mode.

  - Auto Parse Mode -
  Sumo Logic has the ability to dynamically parse only JSON types when searching

logs.

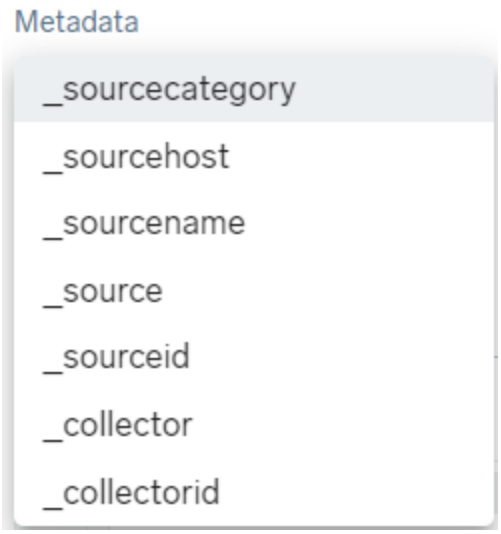


③ Scope

You can determine the range of logs to analyze.
The Parsed template (described below) will be applied to the logs within the range you set.

- **All Data**
  All message data is in scope.
- **Specific Data**

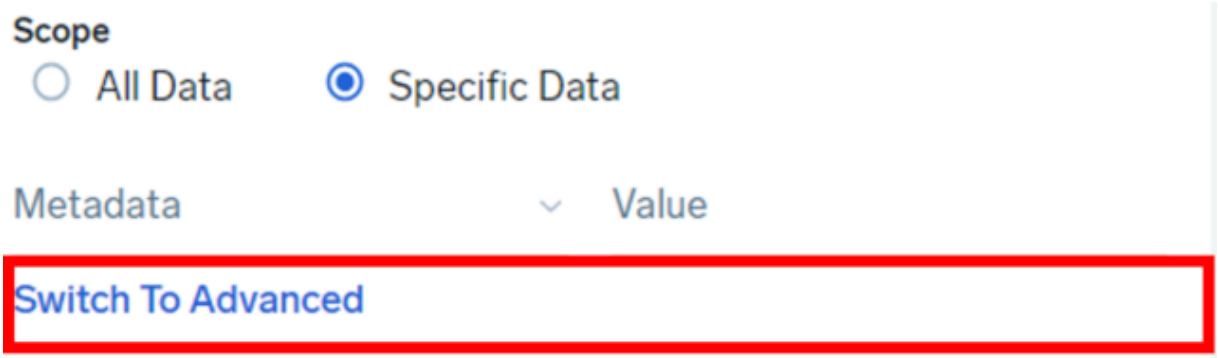
- MetadataSpecify
the following built-in metadata as the log extraction range.



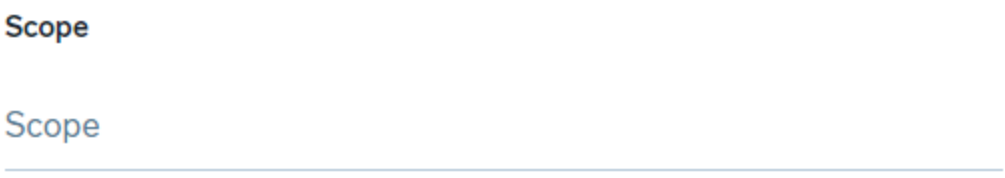
- ValueSpecify
the Metadata field value.
Example) _sourcecategory=/Prod/NginxWebServer/Access

*You can use search expressions such as Boolean values and wildcards in the fields.
For details, see [Keyword Search Expressions .]

- Switch To Advanced -



Selecting Switch To Advanced will take you to a form where you enter everything manually.



- **Parsed Template (Optional)**
② Applied At is displayed only if you specify Ingest Time.
Templates used for Parse Expression (described later).

④ Parse Expression

② Applied At is displayed only if you specify Ingest Time.
Write an extraction formula for fields and values using supported analysis and search operators.

example)

Parse Expression *

```
1  parse regex "^(?
   <src_ip>\d{1,3}\.\d{1,3}\.\d{1,3}\.\d{1,3})"
2  | parse regex "(?<method>[A-Z]+)\s(?
   <url>\S+)\sHTTP/[\d\.]+\"\s(?
   <status_code>\d+)\s(?<size>[\d-]+)\s\"(?
   <referrer>.*?)\"\s\"(?<user_agent>.+?)\".*"
```

For details about operators, please see below.
・Search Operators
・Parse Operators

⑤ Extracted Fields

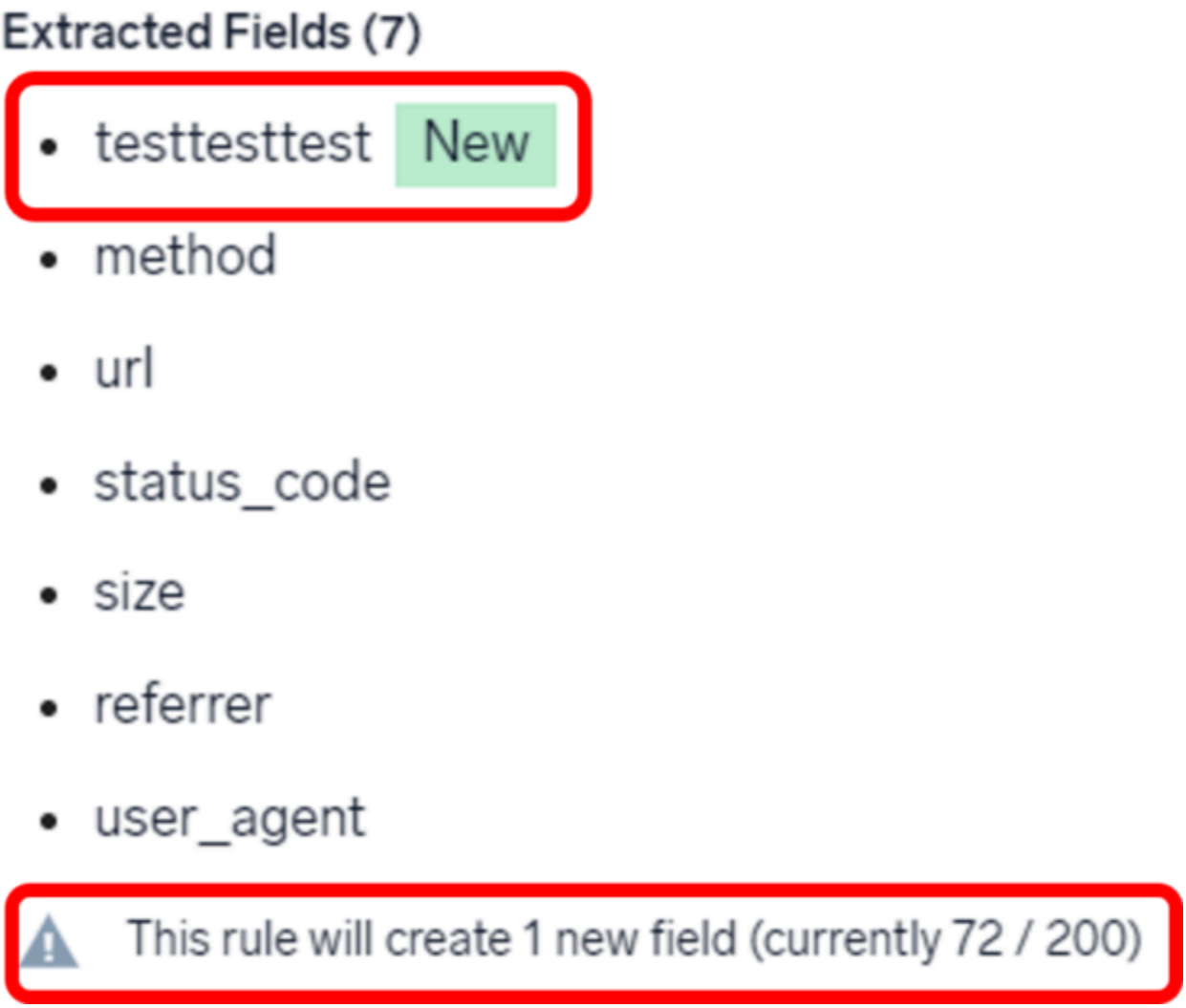
④ The field names specified in the Parse Expression are listed.

Extracted Fields (7)

- src_ip
- method
- url
- status_code
- size
- referrer
- user_agent

- Note -
Fields specified in a field extraction rule are automatically added to the field table schema

and become valid.



summary

After creating an FER, it is a function that you will rarely touch except when delivering new logs. It can easily become complicated when delivering new logs or when there are many logs being acquired/acquired in the first place. For this reason, it is best to thoroughly test and configure it in detail to reduce future management overhead.

Reference source

- Field Extractions | Sumo Logic
- Field Naming Convention | Sumo Logic
- Role Capabilities | Sumo Logic