Sumo Logic – How to create metadata fields and what to be aware of

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

佐久間昇吾

December 3, 2022

First

If the content of the article is outdated, please also check the official website.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

About Metadata

Sumo Logic has built-in metadata for viewing and analyzing logs and metrics.
You want to define metadata fields with understandable names for searching.

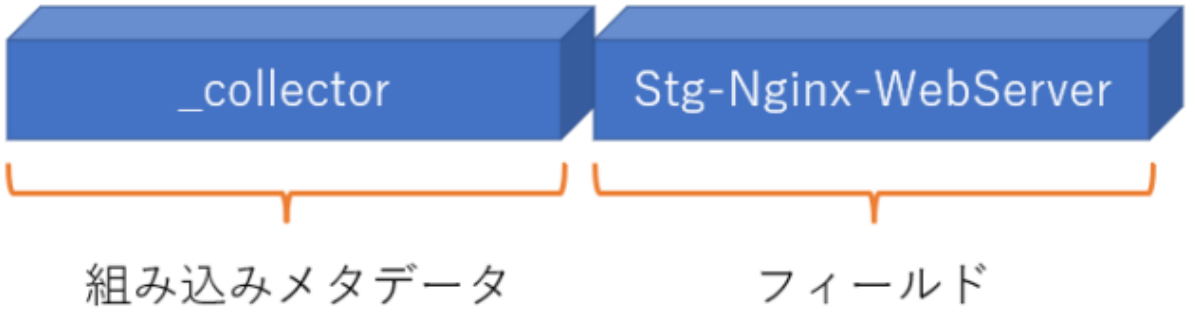
Built-in Metadata Field Names

- **_collector**
- **_source**

- **_sourceCategory**
- **_sourceHost**
- **_sourceName**

If you can imagine it, please skip this.

- **Metadata and fields**



Metadata can be thought of as a box to hold fields.
Fields are actually named values. When searching, you specify the field name.
Figure showing an actual query search

Things to keep in mind when creating metadata fields

- **Use alphanumeric characters**
  Must be alphanumeric.

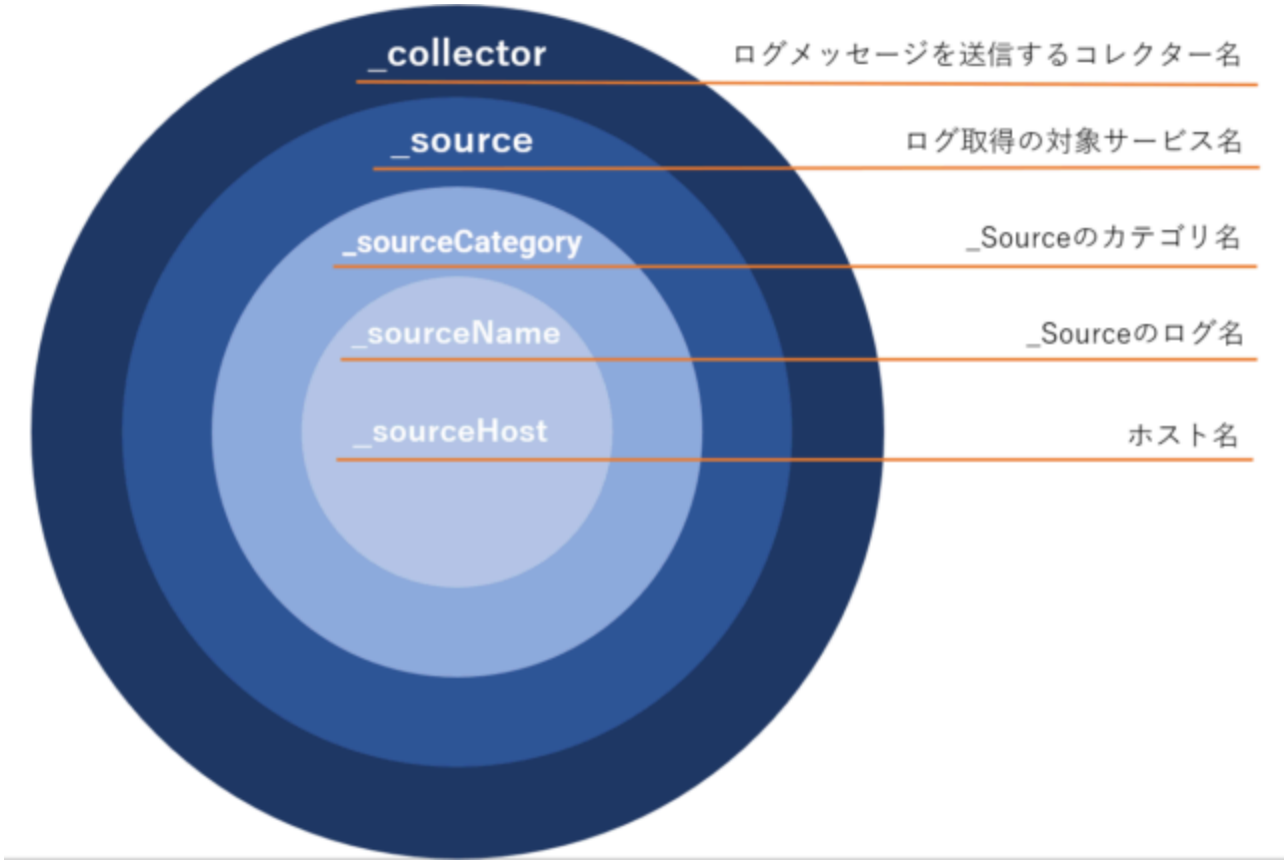
- **You can use punctuation marks such as underscores, hyphens, and periods.**
  With these in place, you can create readable field names.

- **Avoid using spaces**
  Spaces must be enclosed in quotation marks when searching queries, and spaces in configured field names cannot be omitted. Furthermore, wildcards themselves do not work when enclosed in quotation marks, so it is best to avoid using spaces.

- **Uppercase and lowercase letters**
  Since case is not recognized when searching queries, we recommend creating queries using a consistent case.

The meaning of various metadata

Below is a high-level overview of the metadata:



_collector

In Sumo Logic, the Java software that collects log and metric data is called a collector,

and the name you specify when installing the collector becomes the field name.

- Note -
It is a good idea to define a name for metadata_collector **that** will identify which host machine the collector is installed on.
Example:
Prod-NginxWebServer

_source

To receive data, you must associate a source with a collector.
The name you specify when creating the source becomes the field name.

- Note -
It is a good idea to define a name for metadata_source **that** will let you know what data you are sending.
Example:
NginxWebServer

_sourceCategory

The category name for the Source, which is
the name you specify when you create the Source.

- Note -
A collector can contain multiple sources.
Therefore, it is recommended to use a hierarchical name for the metadata **_sourceCategory** (Component 1/Component 2/Component 3) that indicates the category of data being sent to the _source.
And the benefit of this hierarchy is that it allows you to narrow down your queries like this:
_sourceCategory=Networking/Firewall/* (all firewall data)
_sourceCategory=Networking/*/Cisco/* (all Cisco data)

Example:
Prod/NginxWebServer/Access

_sourceName

The log file name. The metadata **_sourceName** is set to the file path you entered when creating the source.

_sourceHost

The source hostname. The metadata **_sourceHost** is set to the OS-level hostname.
* With Installed Collectors, you can define a name other than the hostname only when the source is a local file on the host.
However, it is better to use a clear name to identify the hostname on which the collector is installed.

Example) Ja/MySQL/Pri Ja/FW/Sec Where/what do you do/what role

summary

Optimized built-in metadata is easy to understand when executing queries, providing appropriate analytical data insights. In particular, care must be taken when naming _sourceCategory, as this affects the convenience of query searches. We recommend that you define it by referring to the official page at the URL below.

- Sumo Logic - Best Practices for Data Collection

Reference source

- Sumo Logic - Metadata Naming Conventions
- Sumo Logic - Built-in Metadata
- Sumo Logic - Best Practices for Data Collection