Creating Field Extractions

Document Usage Guidelines

- Should be used only for enrolled students
- Not meant to be a self-paced document, an instructor is needed
- Lab Exercise slides reference the hands-on lab exercise guide
- Do not distribute

Course Goals

- Determine when the field extraction process occurs
- Utilize the Field Extractor (FX)
- Compare regex and delimited field extractions
- Create and use a regex field extraction
- Create and use a delimited field extraction
- Modify extracted fields

Course Outline

- Use the Field Extractor
- Create Regex Field Extractions
- Create Delimited Field Extractions

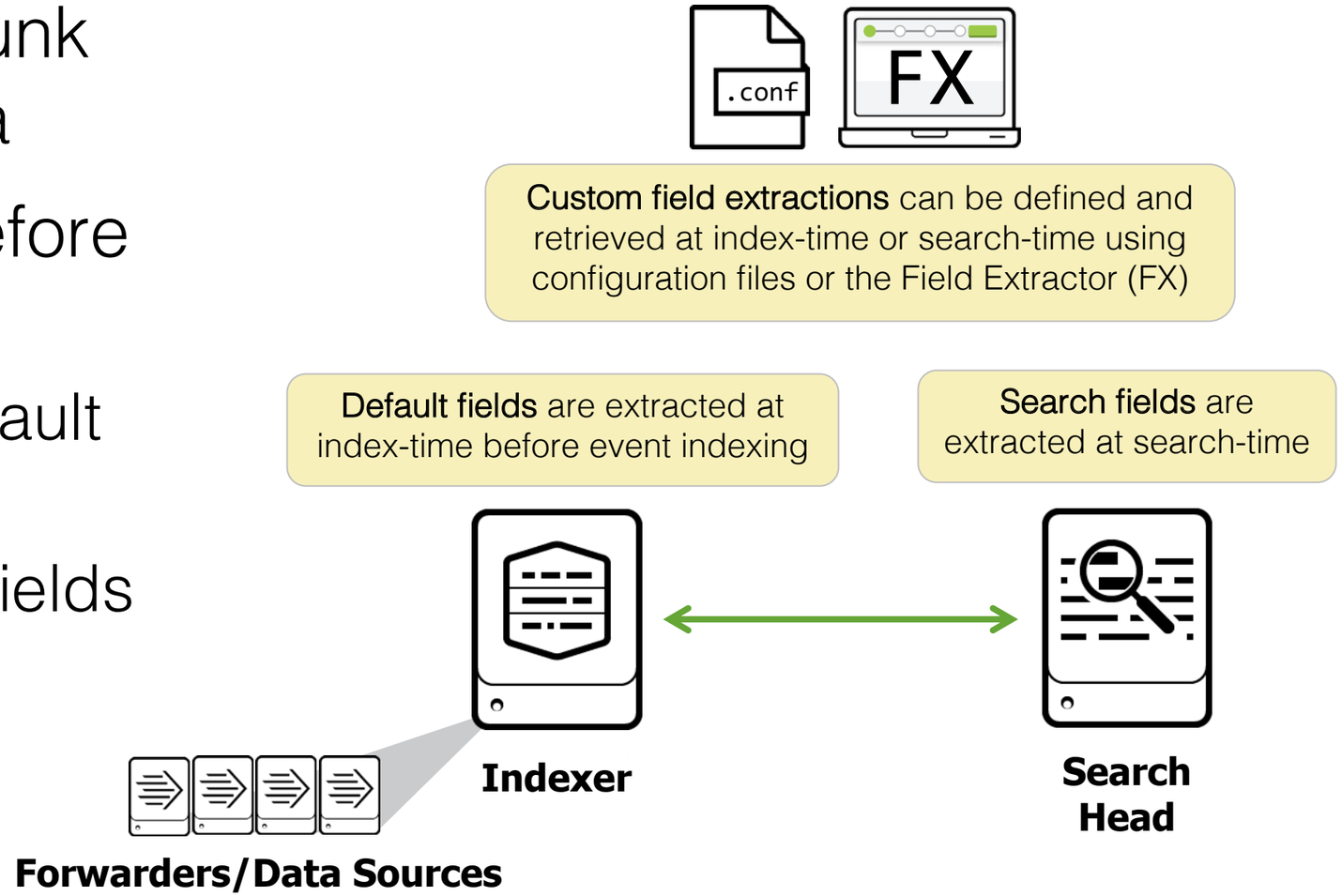
Use the Field Extractor

Topic Objectives

- Explore the different types of extracted fields and when they are extracted
- Define the Splunk Web Field Extractor (FX)

Explore Field Extraction

- The process by which Splunk pulls fields from event data
- Fields can be extracted before or after event indexing
  - Before: indexed fields (default and custom fields)
  - After: search and custom fields

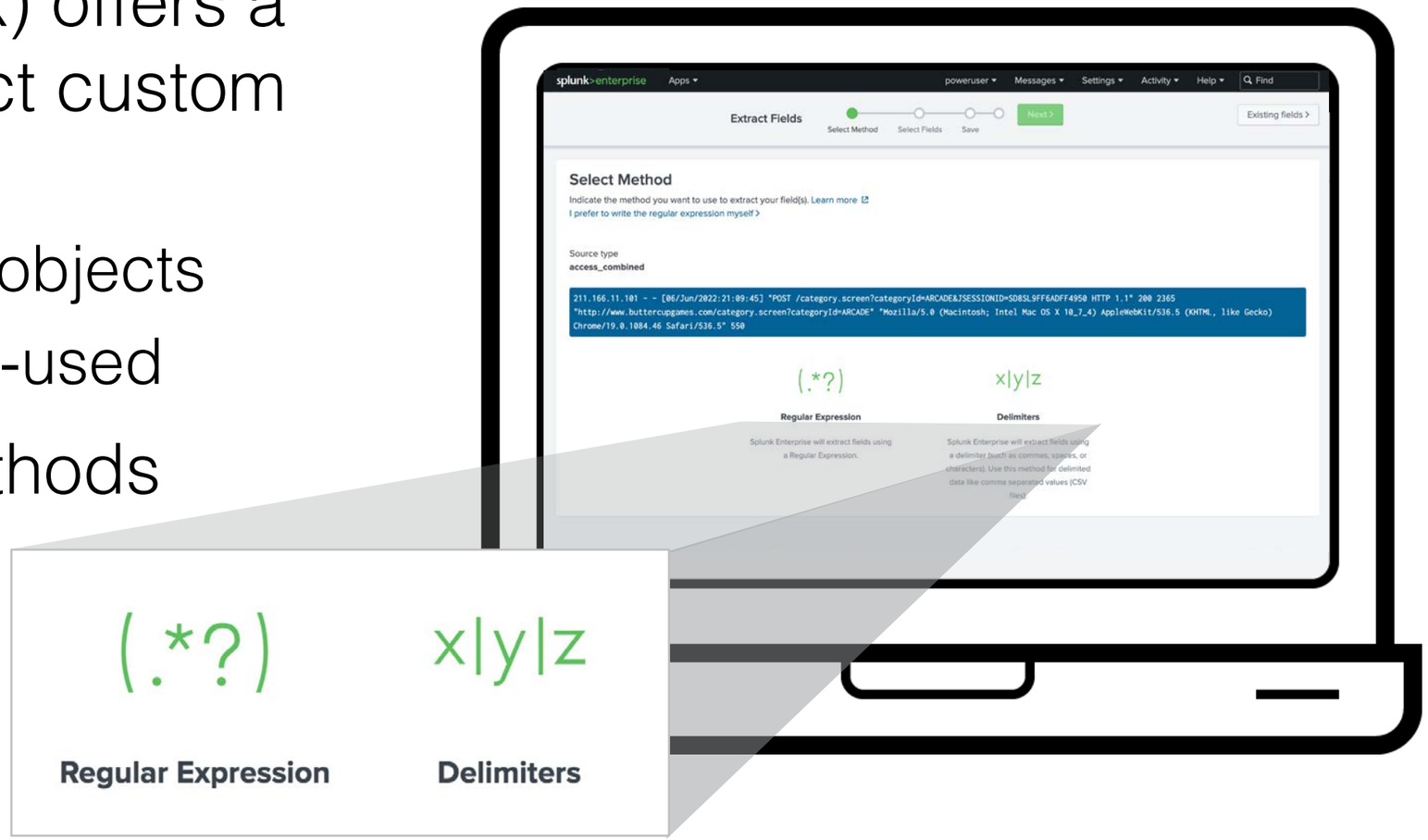


Note

These elements are all components of the Splunk data pipeline.

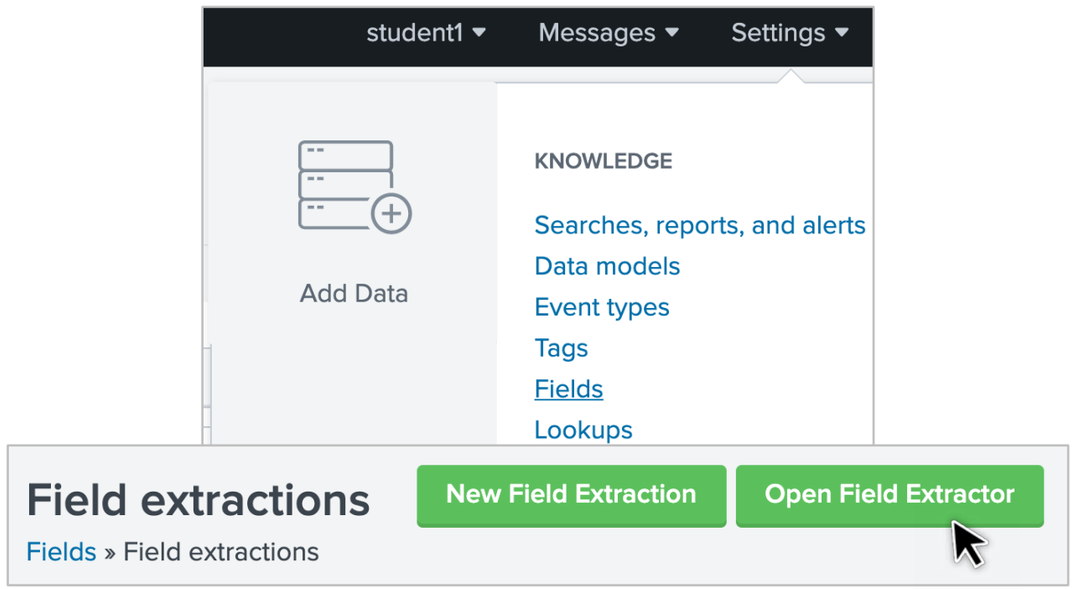
The Field Extractor

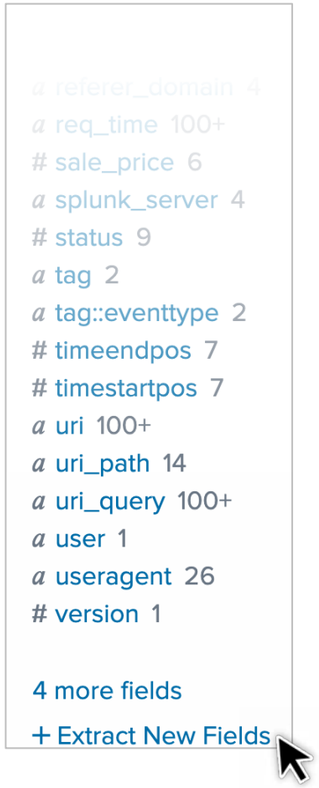
- The Field Extractor (FX) offers a user interface to extract custom fields that:
  - Persist as knowledge objects
  - Can be shared and re-used
- The two extraction methods offered by the FX are **Regular Expression** and **Delimiters**

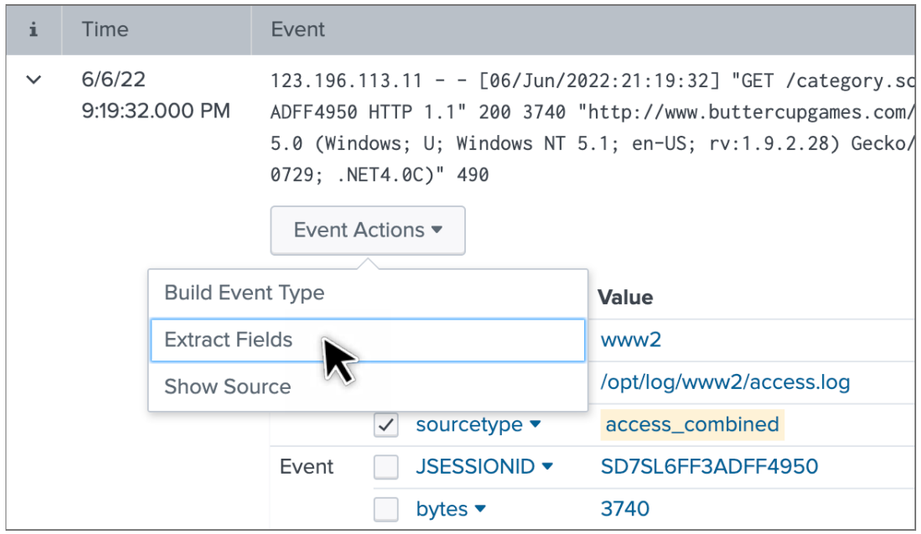


Access the Field Extractor

The FX user interface can be accessed via Settings, the Fields Sidebar, or the Event Actions menu







Create Regex Field Extractions

Topic Objectives

- Identify basics of regular expressions (regex)
- Explore the regex field extraction workflow
- Edit regex for field extractions

Field Extractor Method: Regex

- FX extracts fields using a Regular Expression (regex)
  - Generates regex from user-selected event data
  - Regex can be manually edited for more accuracy
- Use this option when your event contains unstructured data (e.g., a system log file)

What Is Regex?

- A regex (regular expression) is a case-sensitive sequence of characters defining a pattern
- Each character is either a regular character (with literal meaning) or a metacharacter (with special meaning)
- Widely used in programming and scripting languages for a variety of string processing tasks

regex example for U.S. email addresses

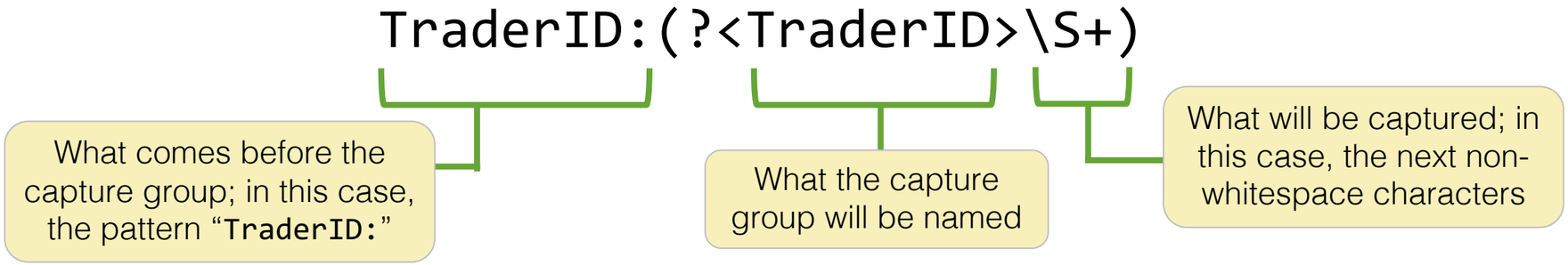
```
\b[a-zA-Z0-9._%+-]+@[a-zA-Z0-9_-]+\.[a-zA-Z]{2,}\b
```


Basics of Regex

regex	`cat`  Regular characters are treated literally	`c.t`  A `.` is treated as a wildcard and will match any character	`c\.t`  A `\` is used to “escape” characters so they can be treated as literal characters
matches	cat	cat cut c1t c#t …any many others…	c.t

Regex Captures

- Regex can “capture” part of the matching pattern by using `()`
- You can reference the capture by giving it a name using: `?<name>`



- With some older versions of regex, a “P” must be inserted in order to perform named captures `(?P<>  )`

Regex Examples

`user\s(\w+)` captures the word following `user`

`Failed password for invalid user fpass from 211.24.4.4`

Successfully captures `fpass`

`Failed password for invalid user jean-luc from 211.25.4.4`

Doesn't successfully capture `jean-luc` because "`-`" isn't a "word" character

`user\s(\S+)`

`Failed password for invalid user jean-luc from 211.25.4.4`

Successfully captures `jean-luc`

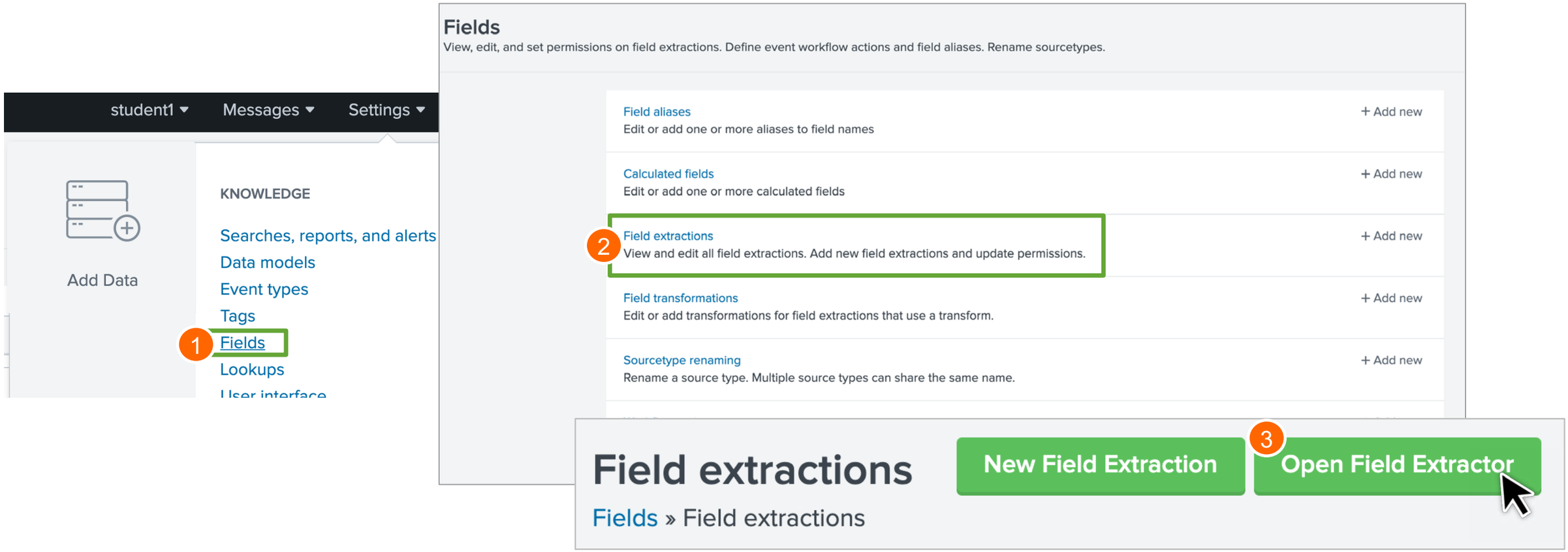
 |

Regex Best Practices

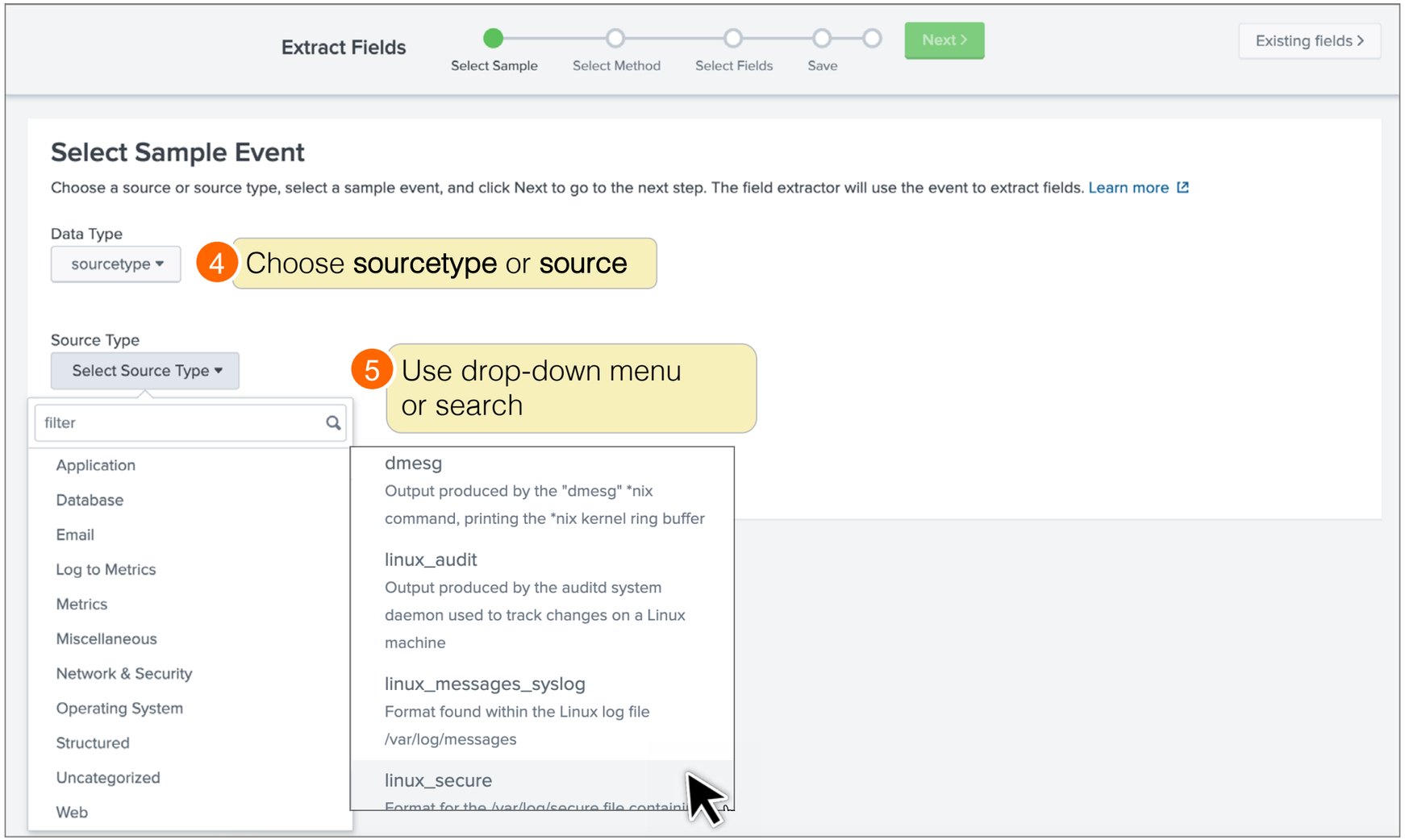
> Best Practice

- Avoid “backtracking” by writing simple and concise regex
  - **Backtracking** occurs when the regex engine must return to a previously saved state to continue its search for a match, causing the engine to make multiple passes
- Quantifiers (e.g. *) and alternation constructs (e.g. |) are powerful but can slow performance by causing backtracking
  - Use + rather than *
  - Avoid using multiple `.*` matches
  - Avoid greedy operators (`.*`), use non-greedy (`.*?`) instead

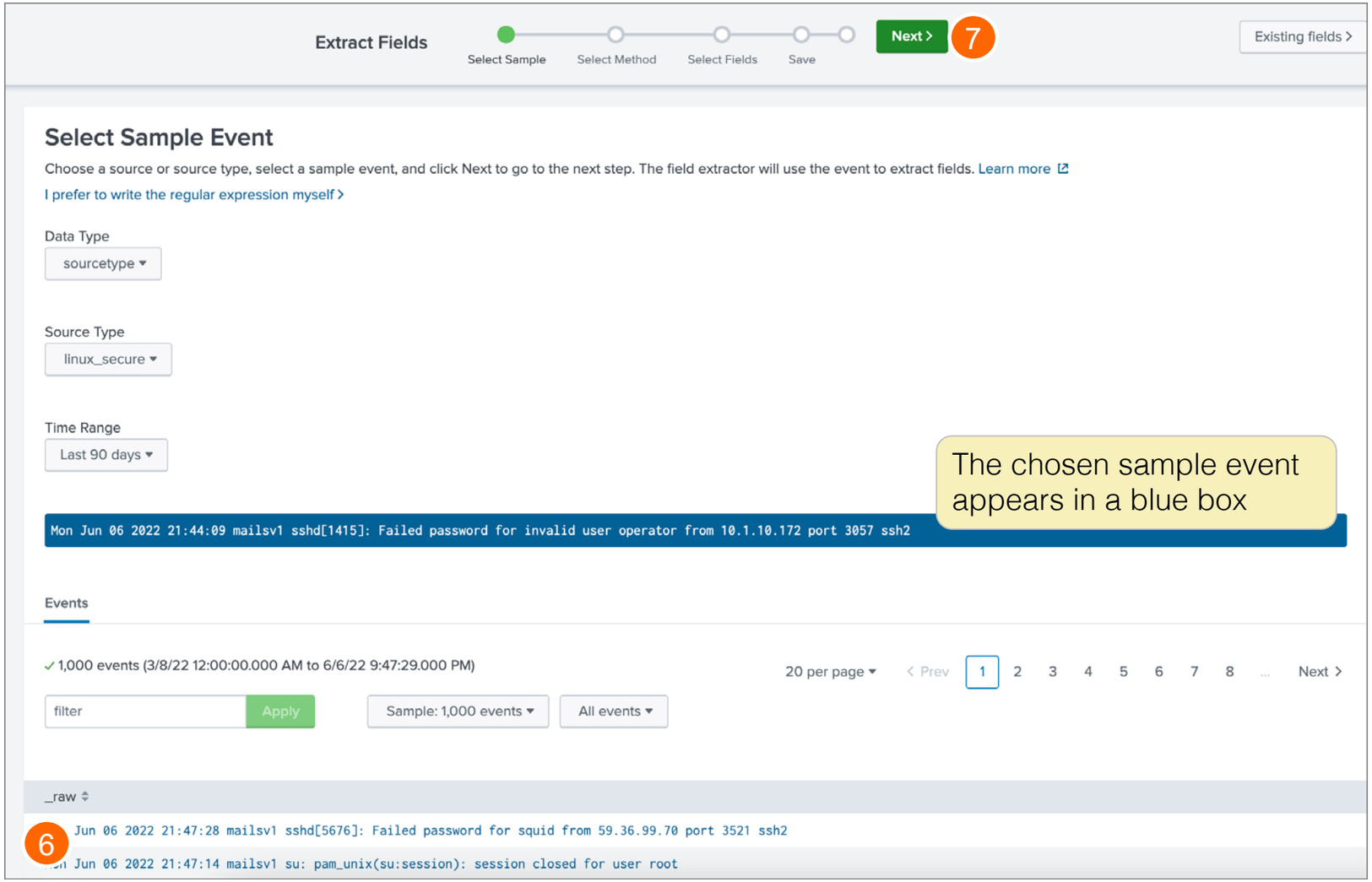
Regex Field Extraction Workflow



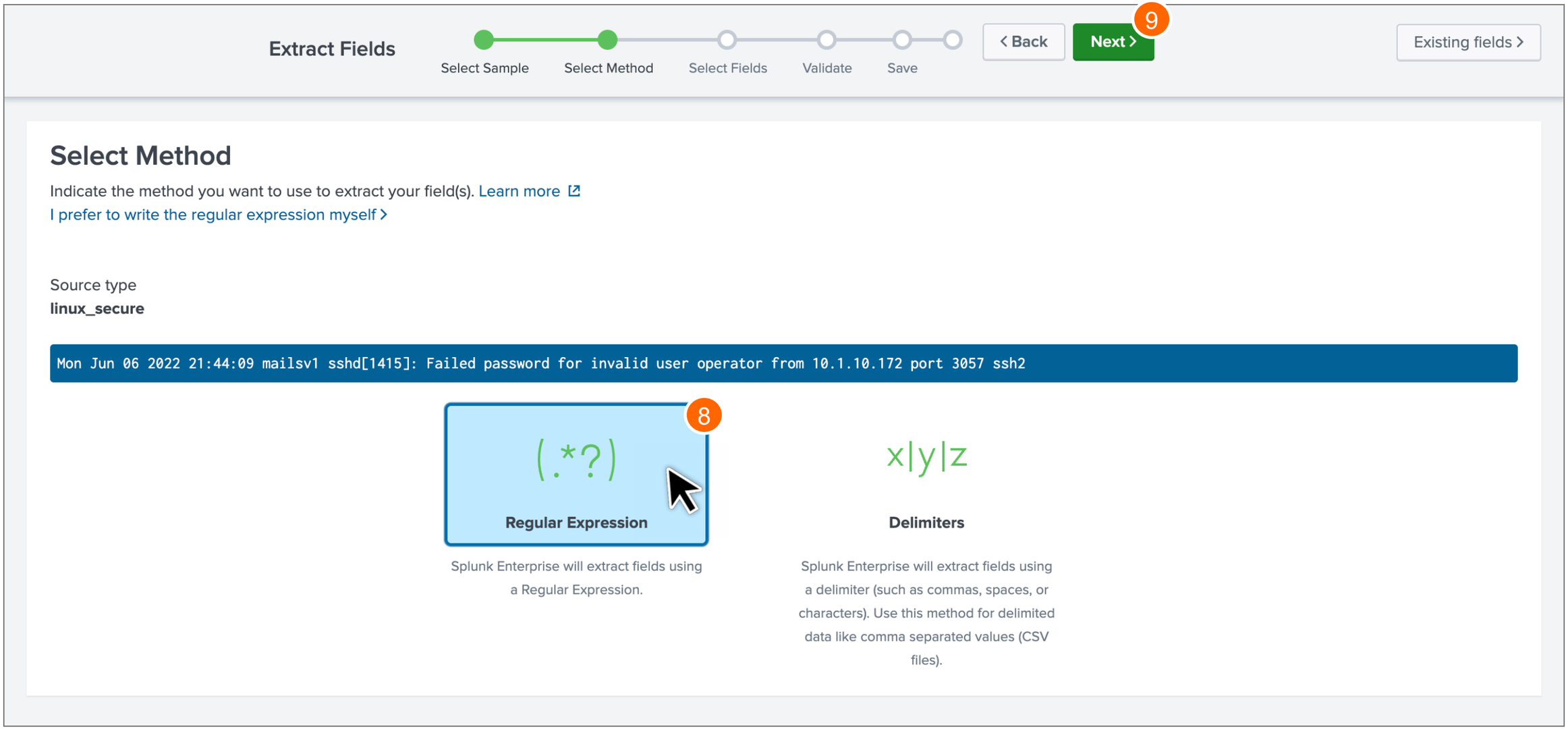
Regex Field Extraction Workflow (cont.)



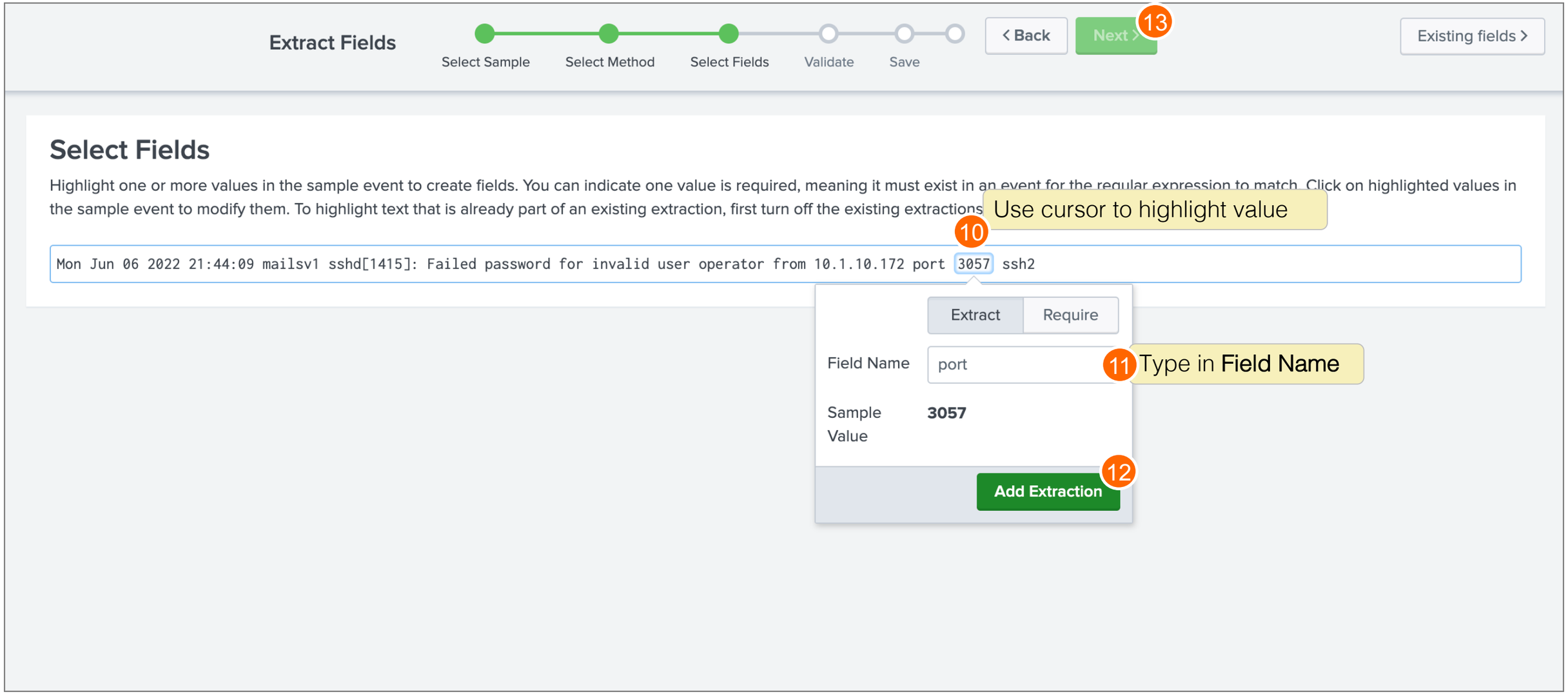
Regex Field Extraction Workflow (cont.)



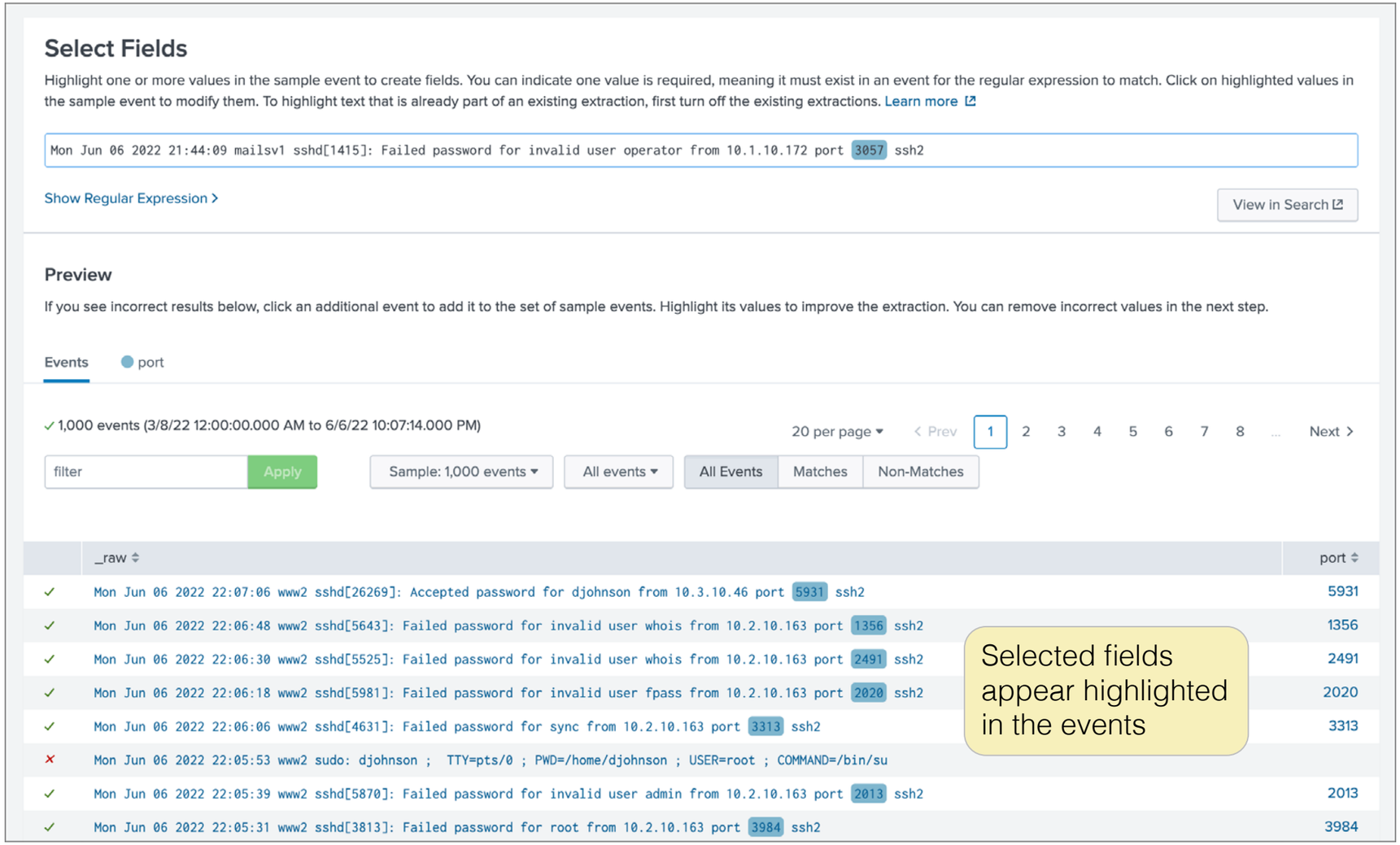
Regex Field Extraction Workflow (cont.)



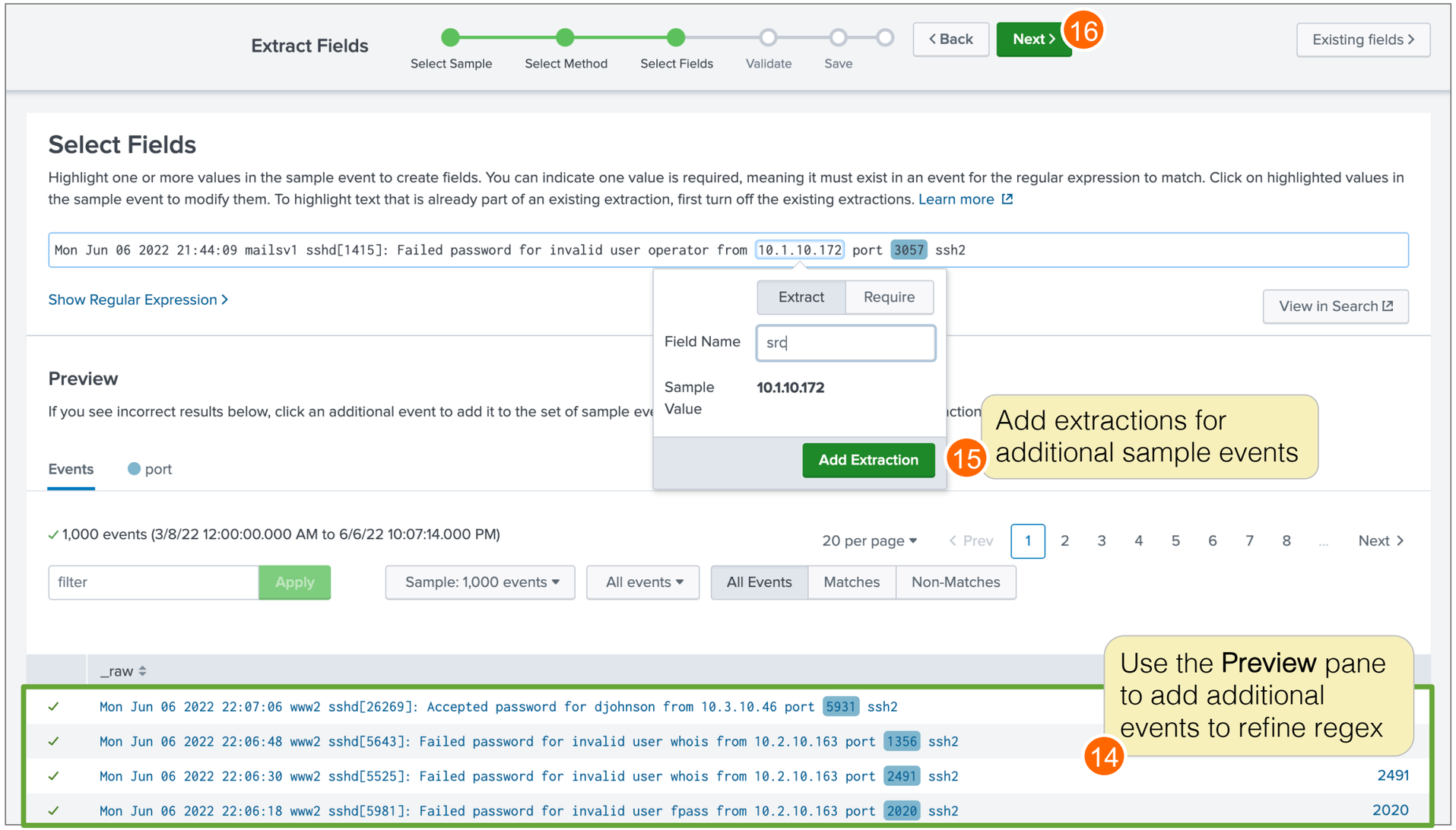
Regex Field Extraction Workflow (cont.)



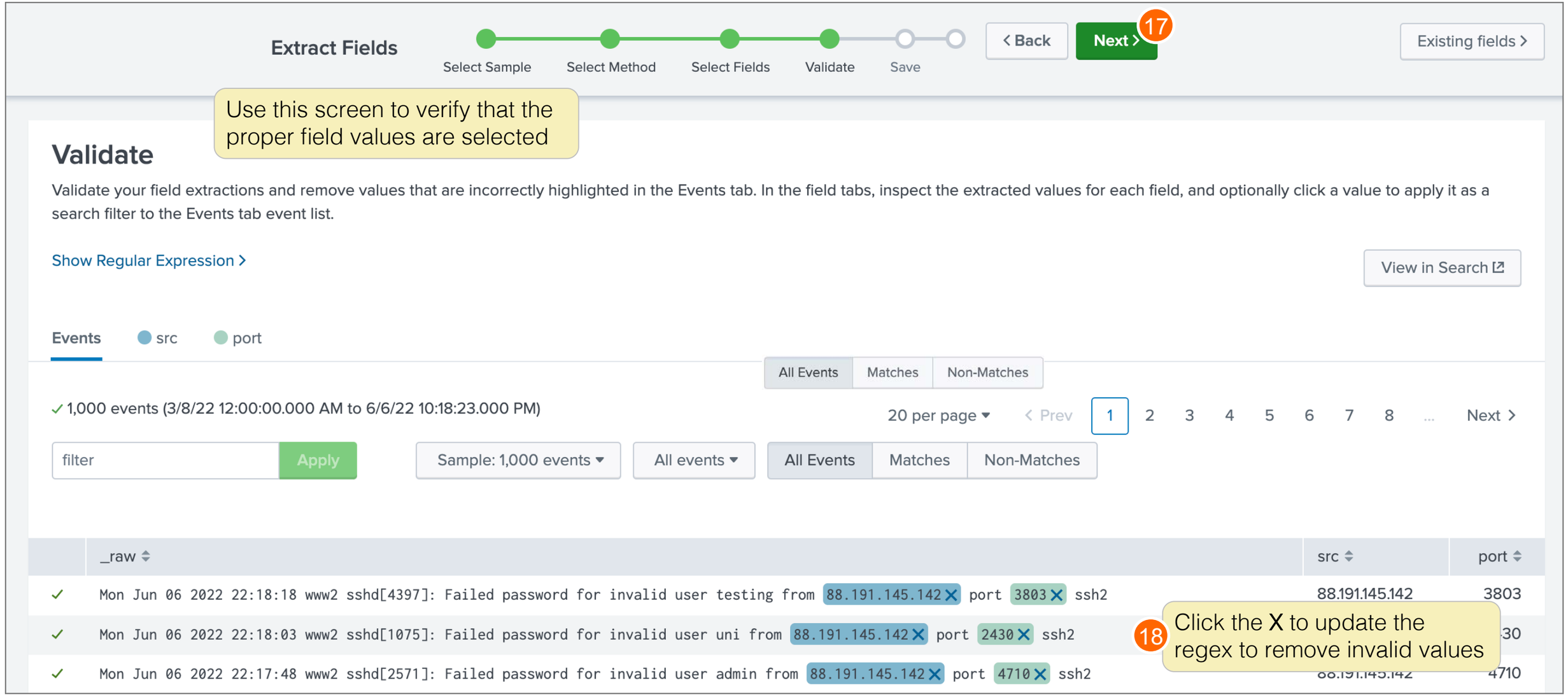
Regex Field Extraction Workflow (cont.)



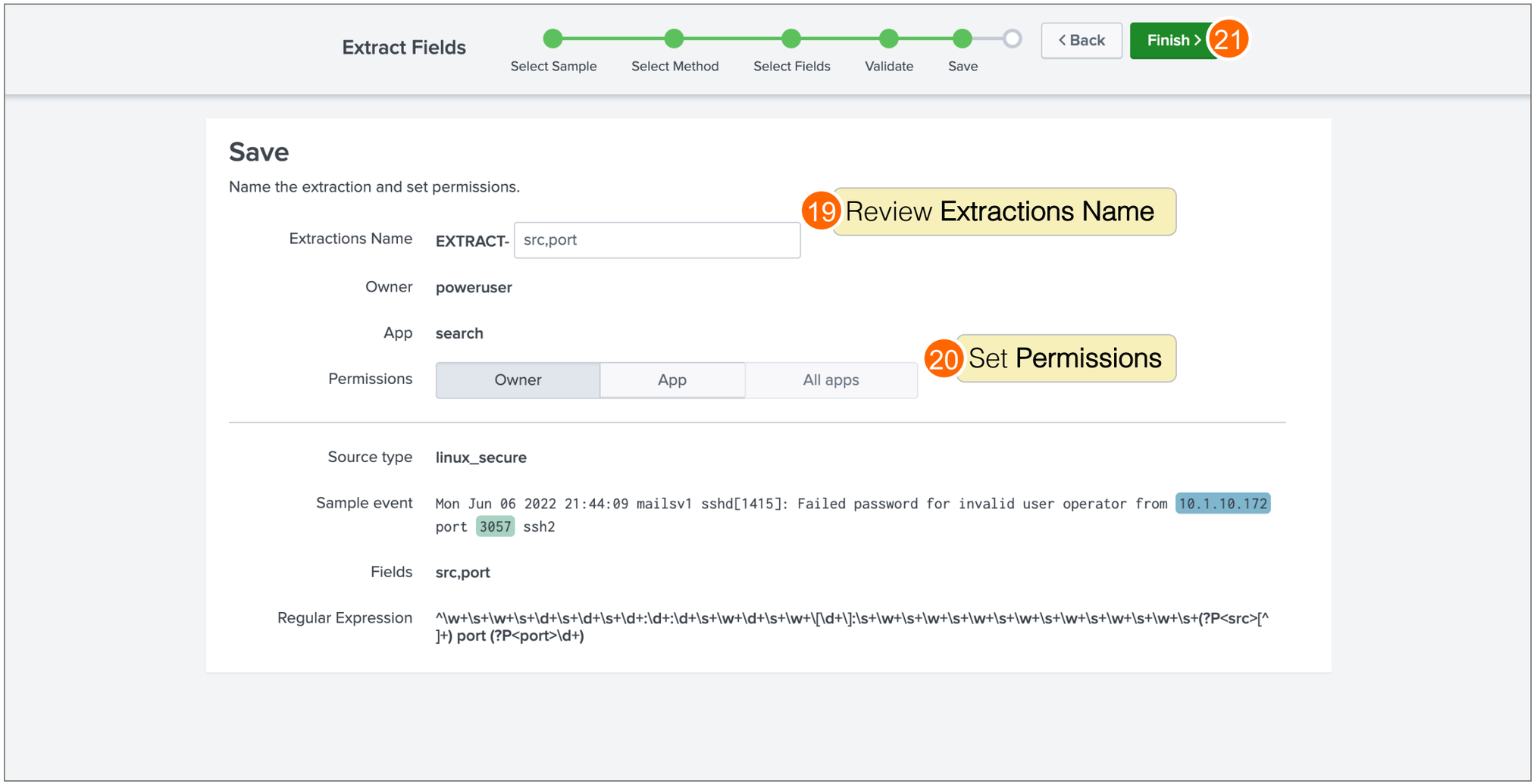
Regex Field Extraction Workflow (cont.)



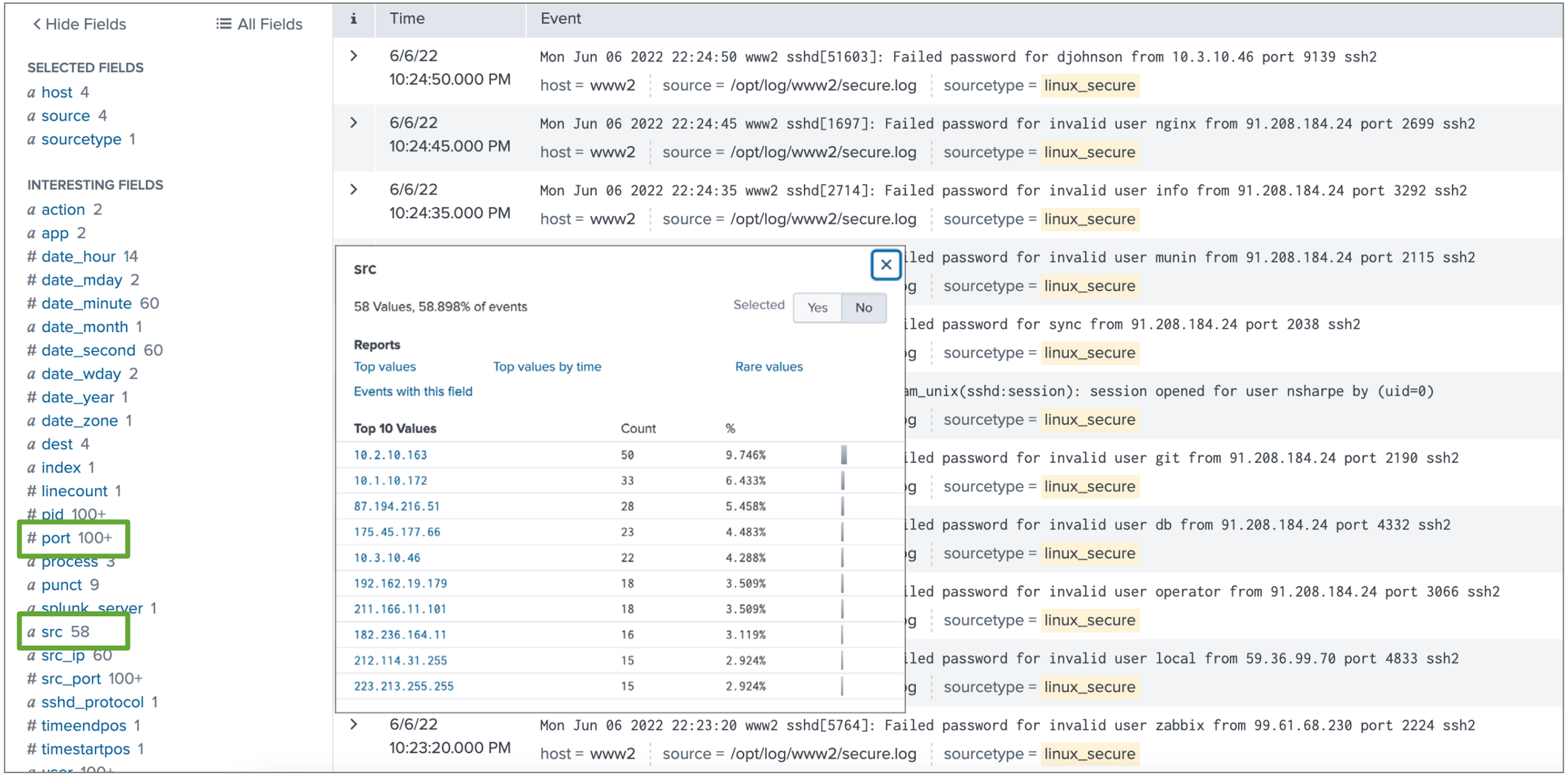
Regex Field Extraction Workflow (cont.)



Regex Field Extraction Workflow (cont.)

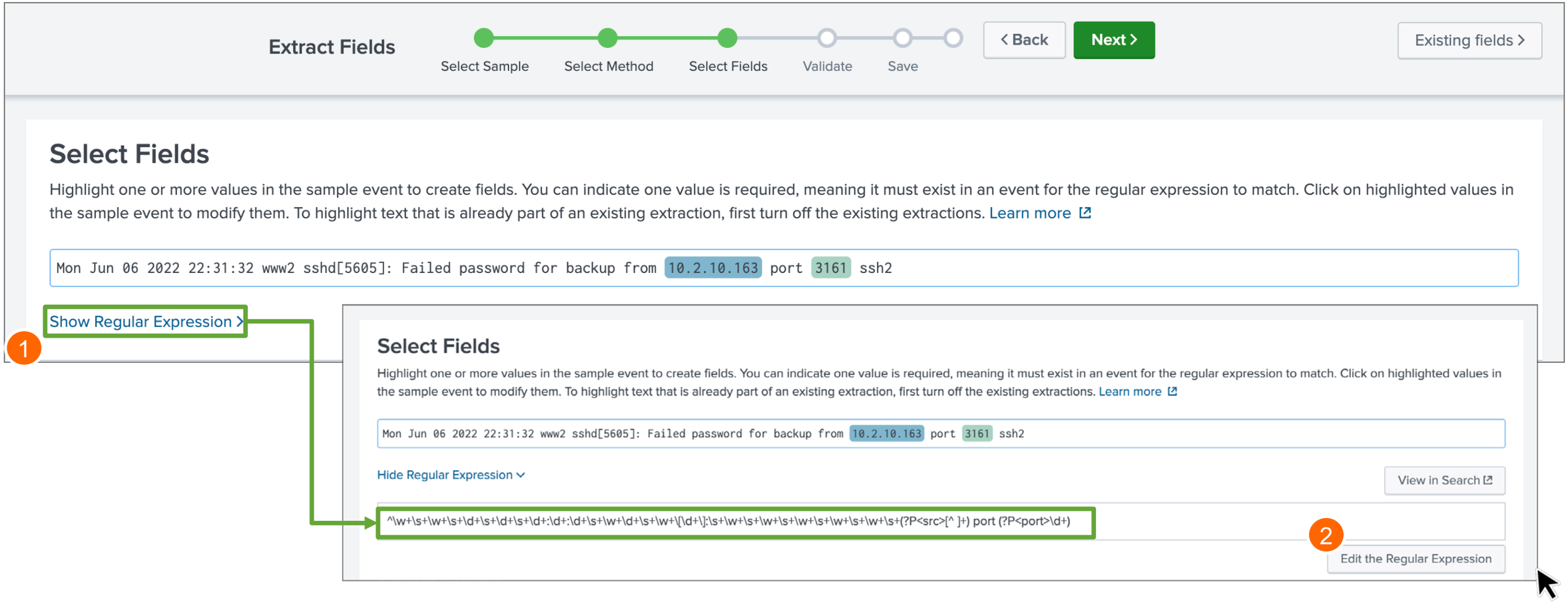


Use the Extracted Fields



Edit Regex for Field Extractions

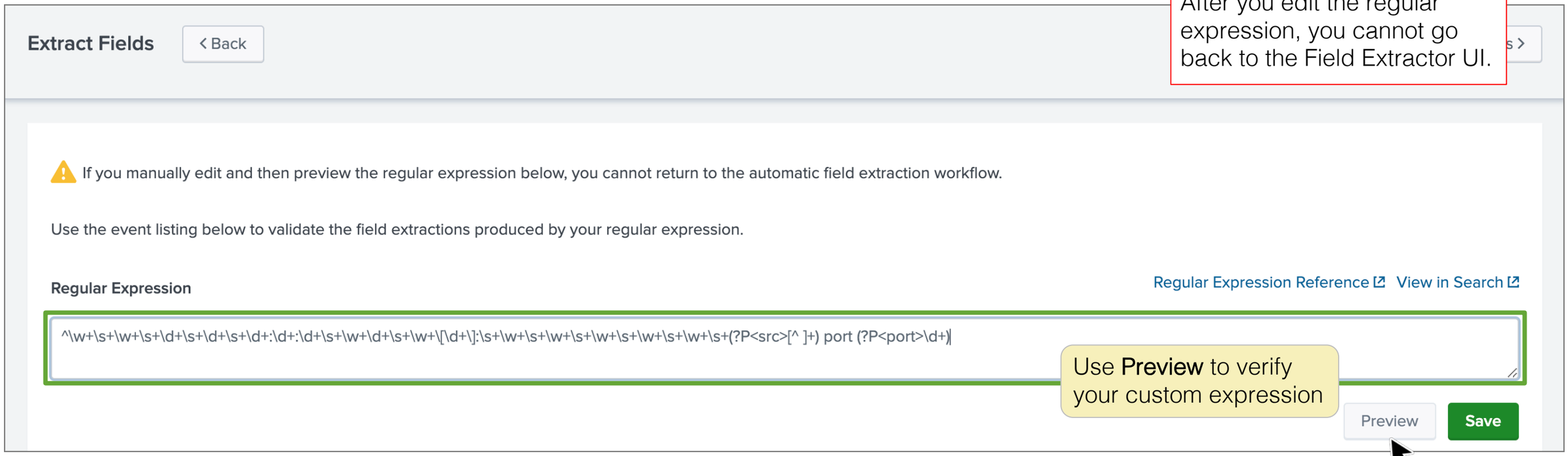
Regular Expression can be edited during the **Select Fields** step



Edit Regex for Field Extractions (cont.)

Warning

After you edit the regular expression, you cannot go back to the Field Extractor UI.



Create Regex Field Extractions Lab Exercise

Time: 15 minutes

Tasks:

- Use the Field Extractor (FX) to extract fields using the Regular Expression method

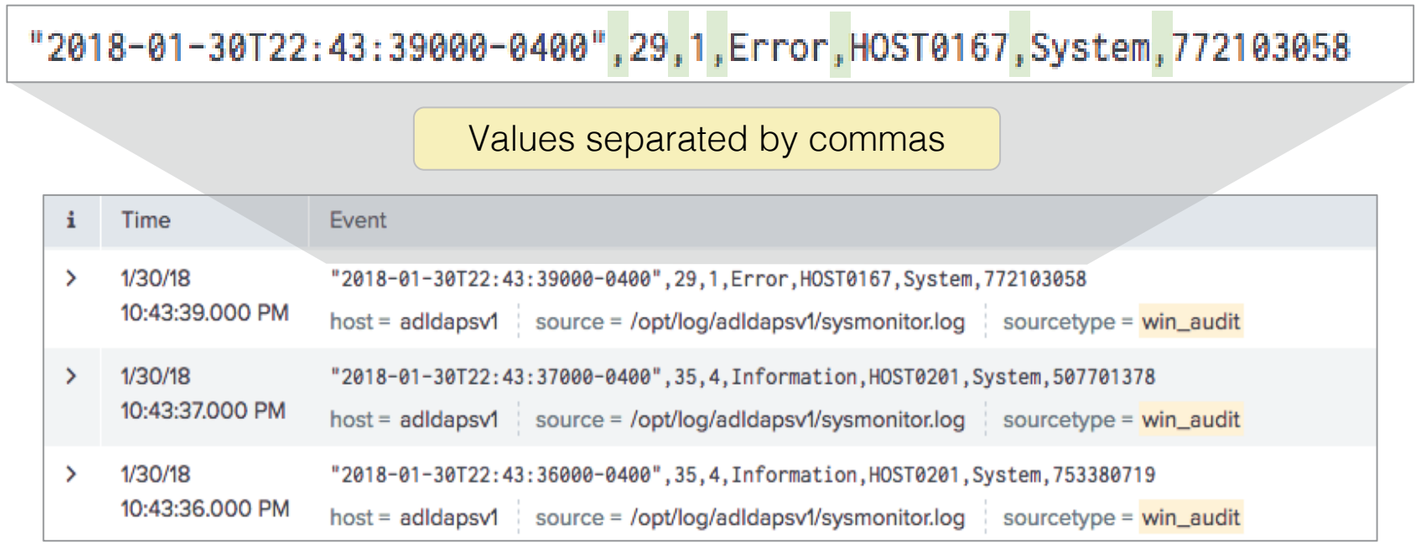
Create Delimited Field Extractions

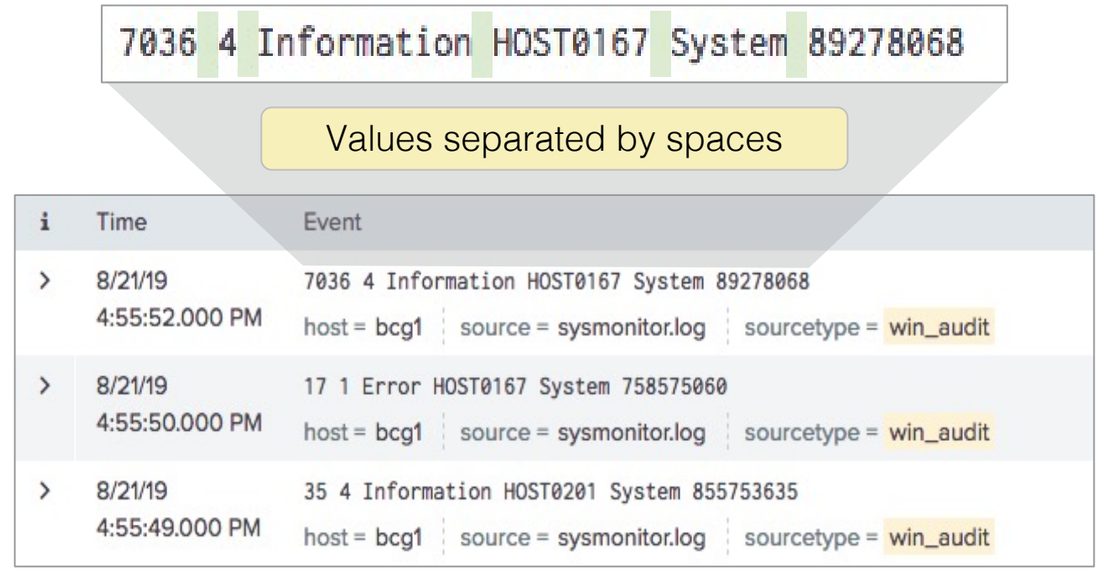
Topic Objectives

- Identify delimited field values in event data
- Explore the delimited field extraction workflow

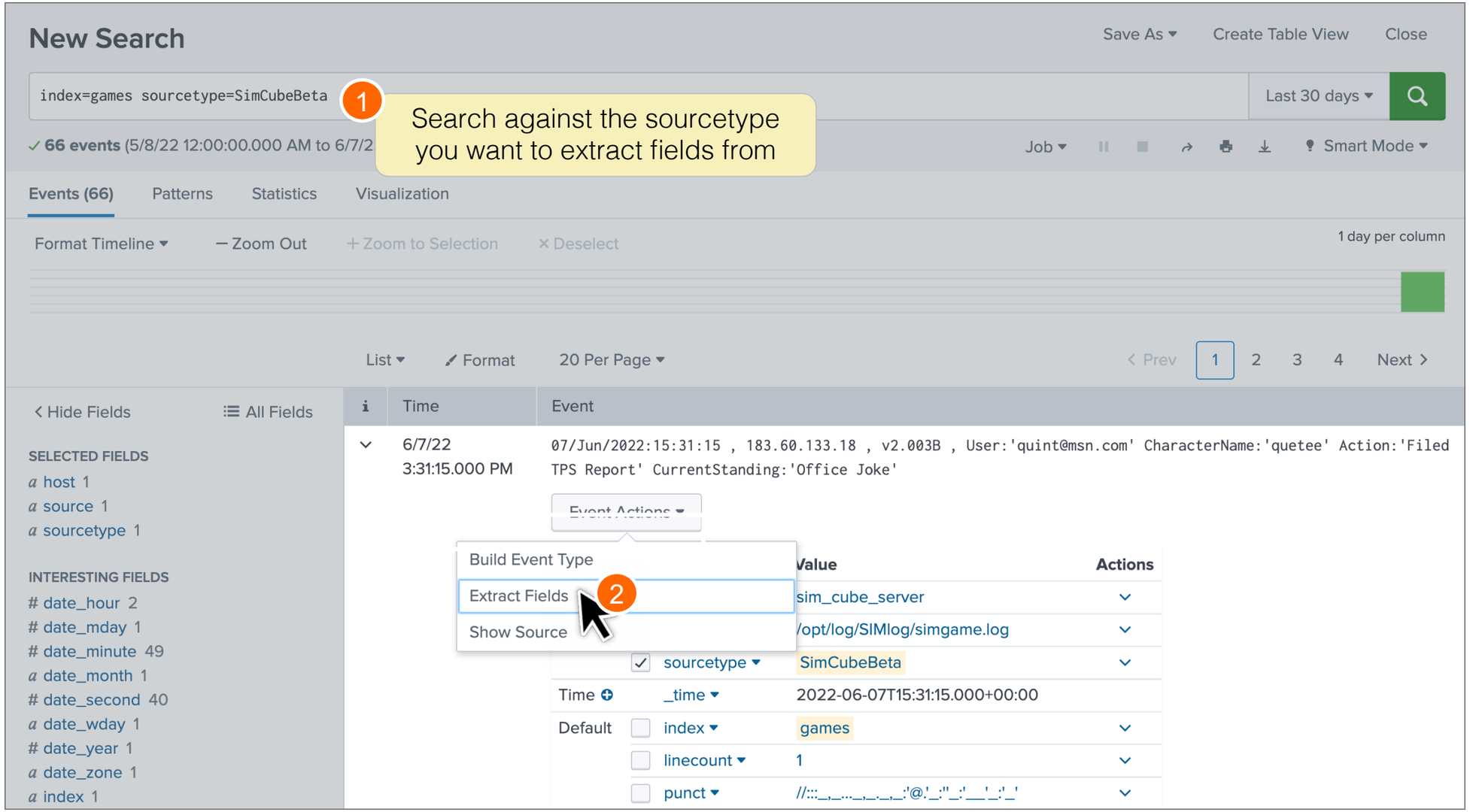
Delimited Field Values in Event Data

Delimited field values exist in consistently structured events and are separated by spaces, commas, or characters (i.e. delimiters)

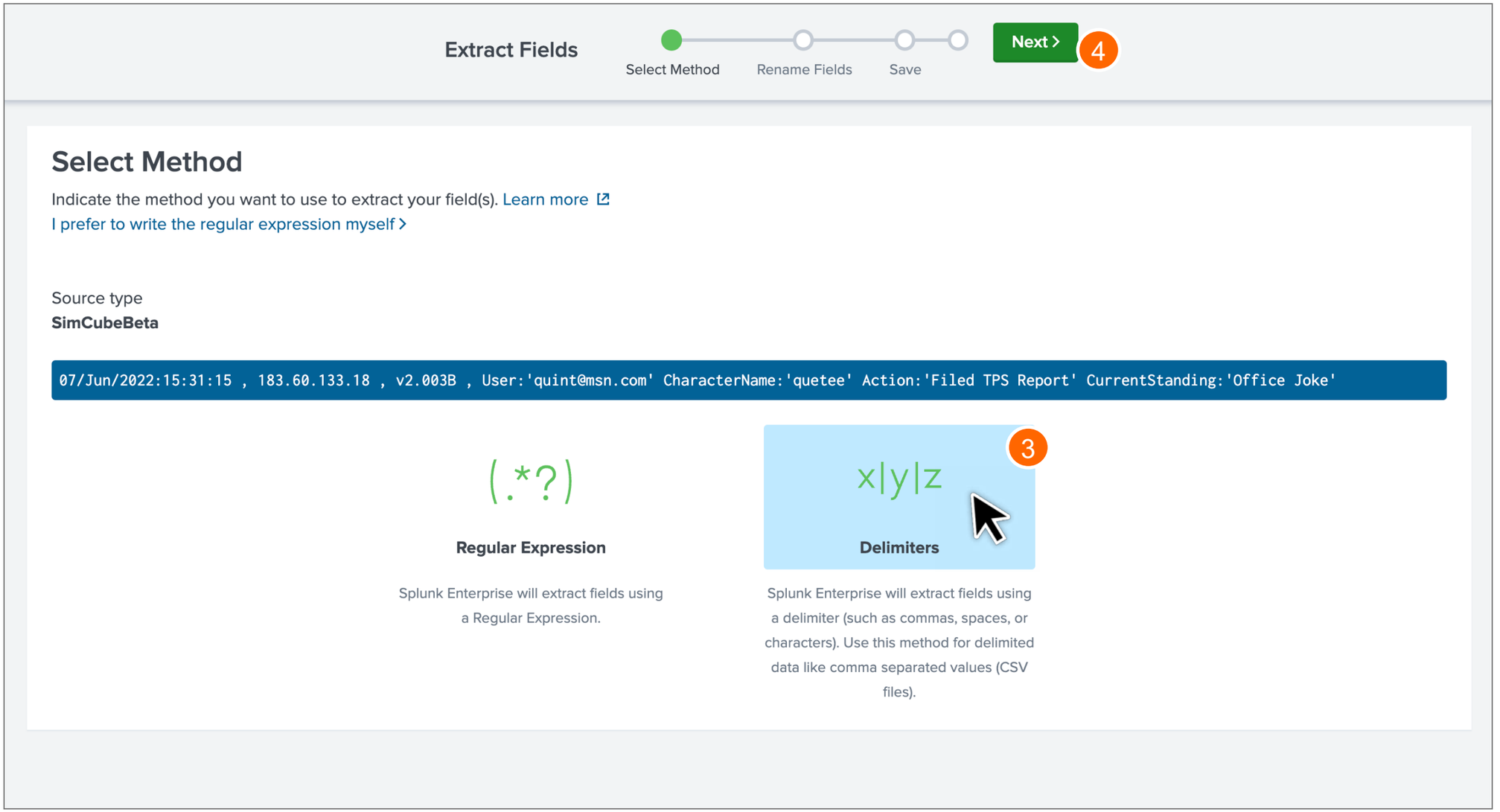




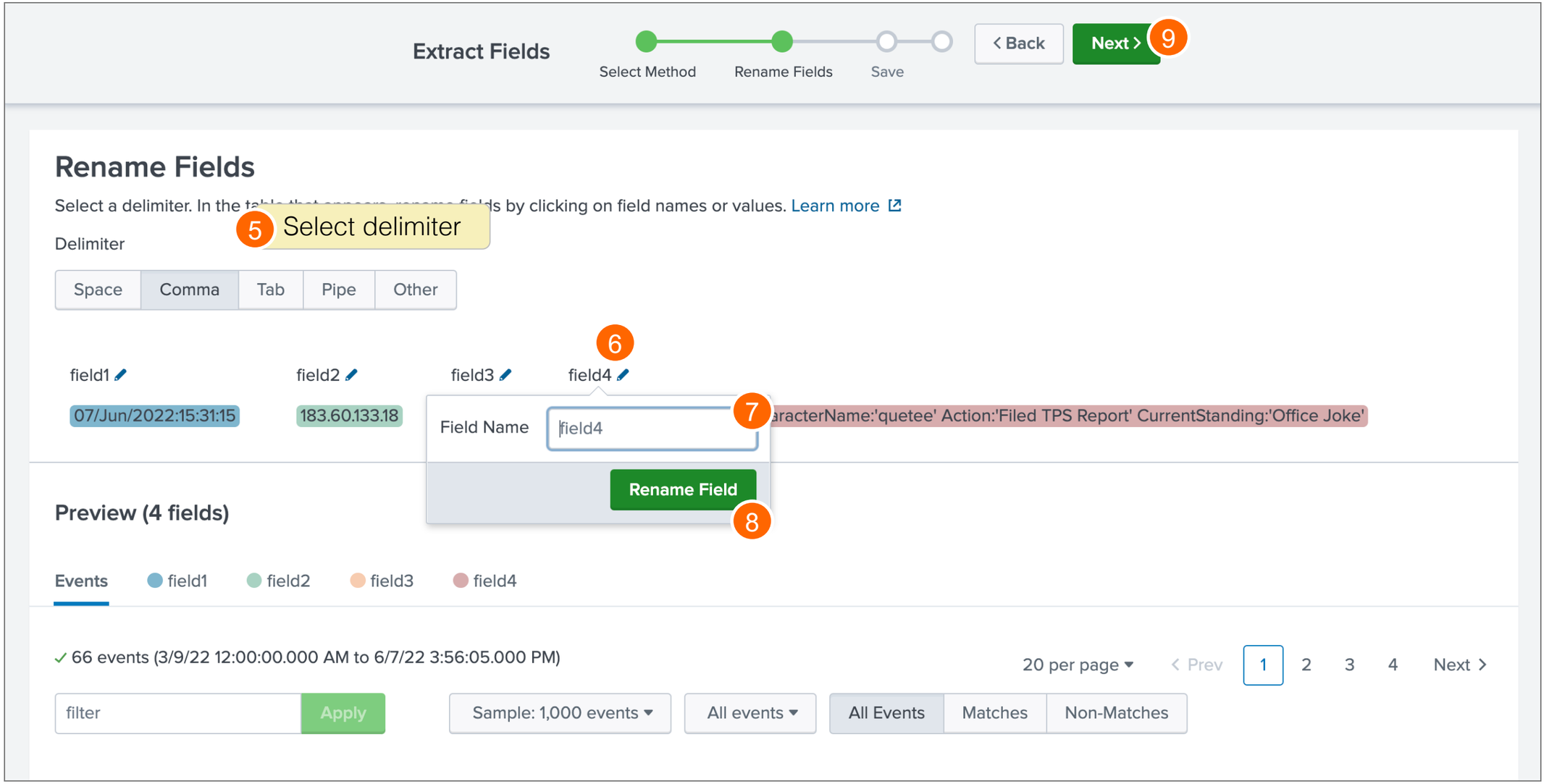
Delimited Field Extraction Workflow



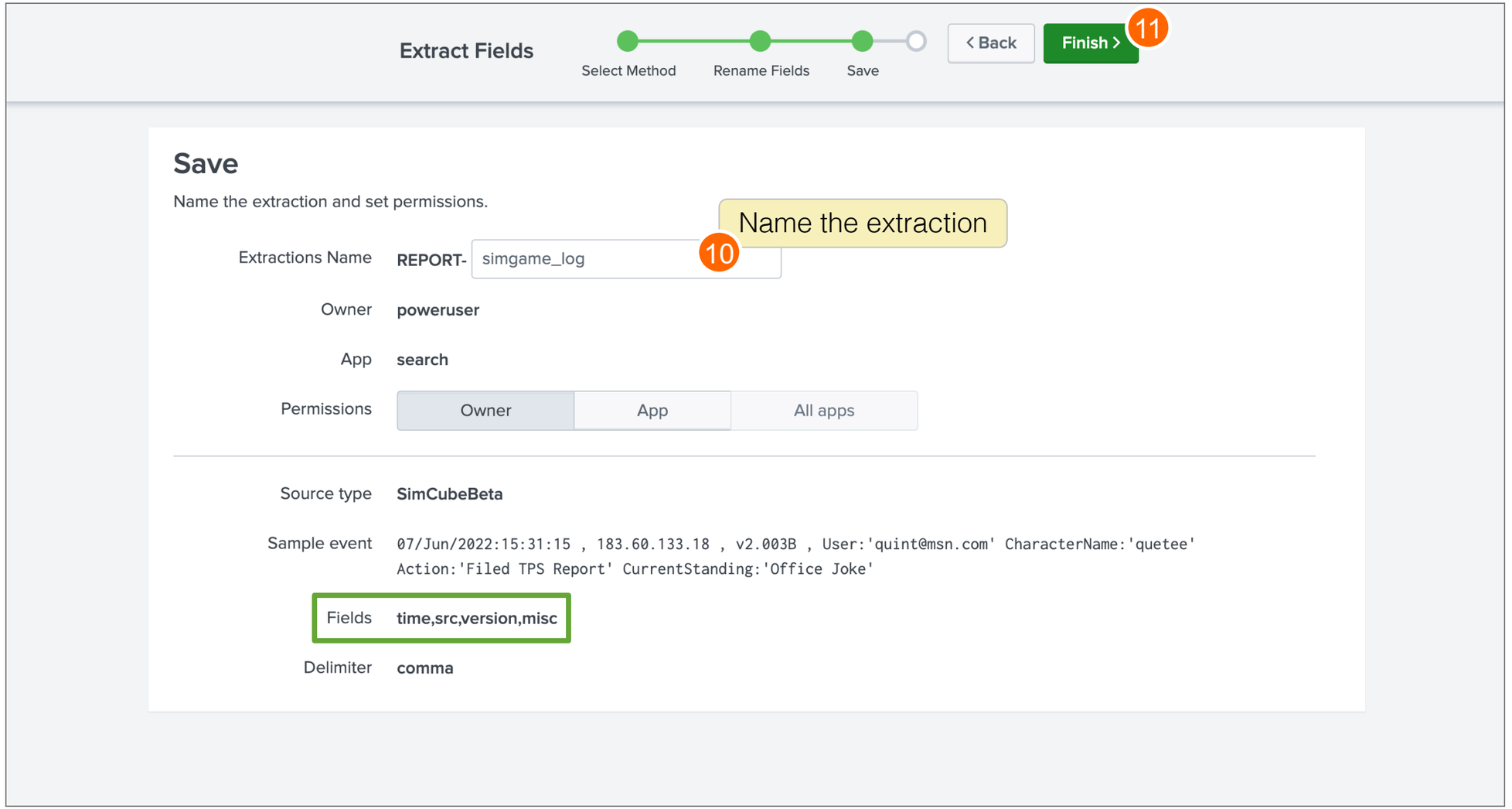
Delimited Field Extraction Workflow (cont.)



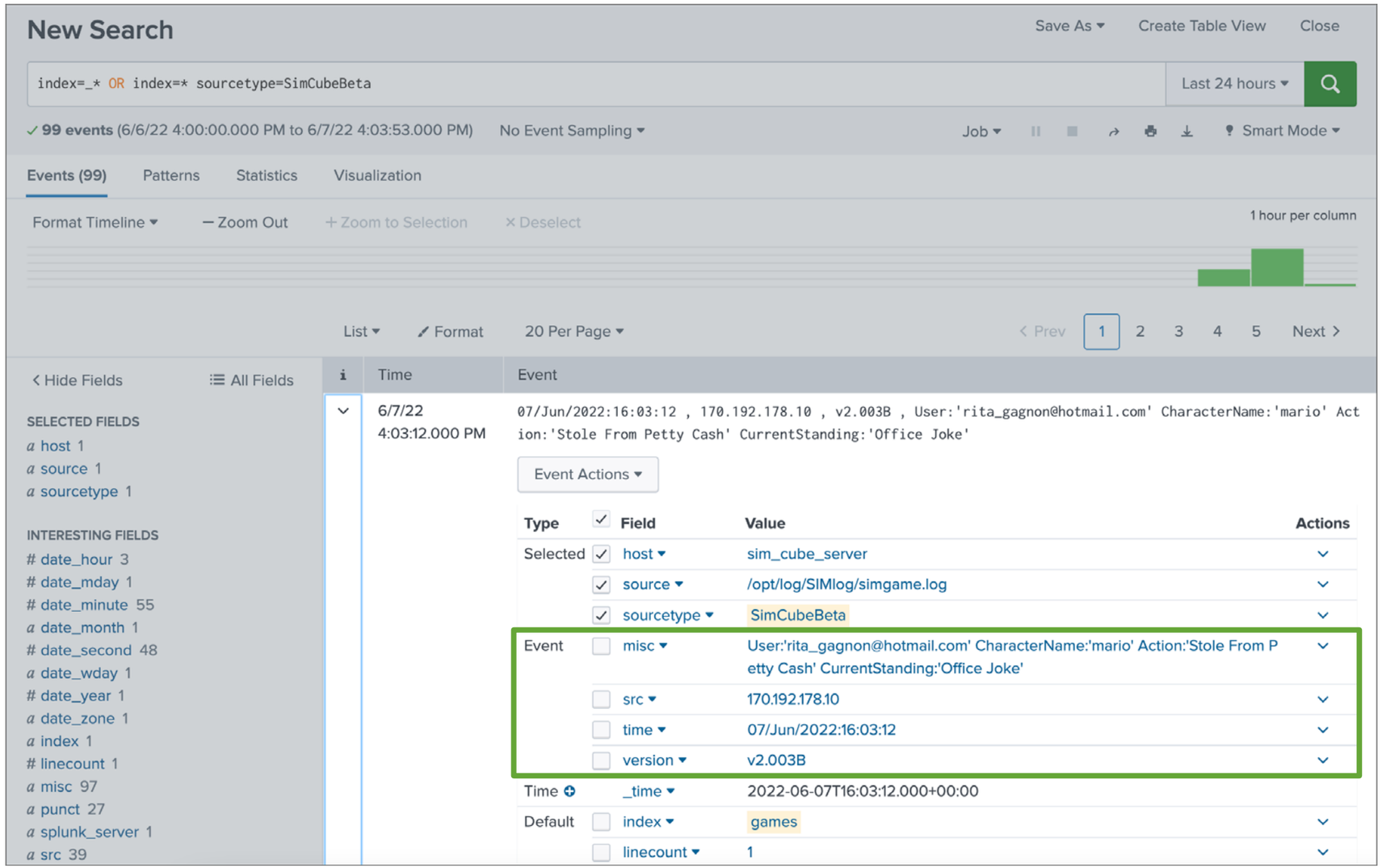
Delimited Field Extraction Workflow (cont.)



Delimited Field Extraction Workflow (cont.)



Use a Delimited Field Extraction



Create Delimited Field Extractions Lab Exercise

Time: 15 minutes

Tasks:

- Use the Field Extractor (FX) to extract fields using the delimiters method
- Use the regex method to extract additional fields that were not captured using the delimiters method

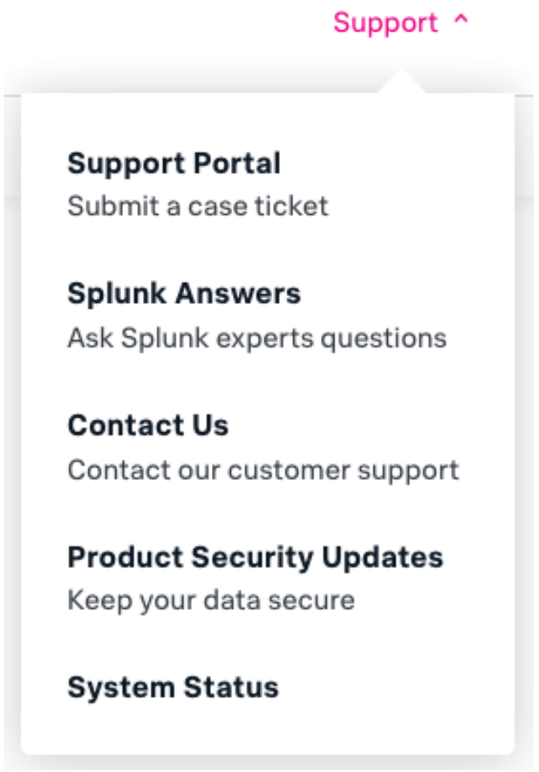
Wrap-up Slides

Community

- Splunk Community Portal
  community.splunk.com
  - Answers
  - Discussions
  - Splunk Trust
  - User Groups
  - Ideas
- Splunk Blogs
  splunk.com/blog/
- Splunk Apps
  splunkbase.com
- Splunk Dev Google Group
  groups.google.com/forum/#!forum/splunkdev
- Splunk Docs on Twitter
  twitter.com/splunkdocs
- Splunk Dev on Twitter
  twitter.com/splunkdev
- Splunk Live!
  splunklive.splunk.com
- .conf
  conf.splunk.com

Support Programs

- **Web**
  - Documentation: dev.splunk.com and docs.splunk.com
- **Splunk Lantern**
  Guidance from Splunk experts
  - lantern.splunk.com
- **Global Support**
  Support for critical issues, a dedicated resource to manage your account – 24 x 7 x 365
  - Web: splunk.com/index.php/submit_issue
- **Enterprise, Cloud, ITSI, Security Support**
  - Web: splunk.com/en_us/about-splunk/contact-us.html#tabs/customersupport
  - Phone: (855) SPLUNK-S or (855) 775-8657



Learning Paths (cont.)

Knowledge Manager - Recommended Courses

Free eLearning courses are in blue and courses with an * are present in both learning paths.

- What is Splunk *
- Introduction to Splunk *
- Using Fields *
- Introduction to Knowledge Objects
- Creating Knowledge Objects
- Creating Field Extractions
- Enriching Data with Lookups
- Data Models
- Introduction to Dashboards
- Dynamic Dashboards
- Using Choropleth
- Search Optimization *

Learning Paths

Search Expert - Recommended Courses

Free eLearning courses are in blue and courses with an * are present in both learning paths.

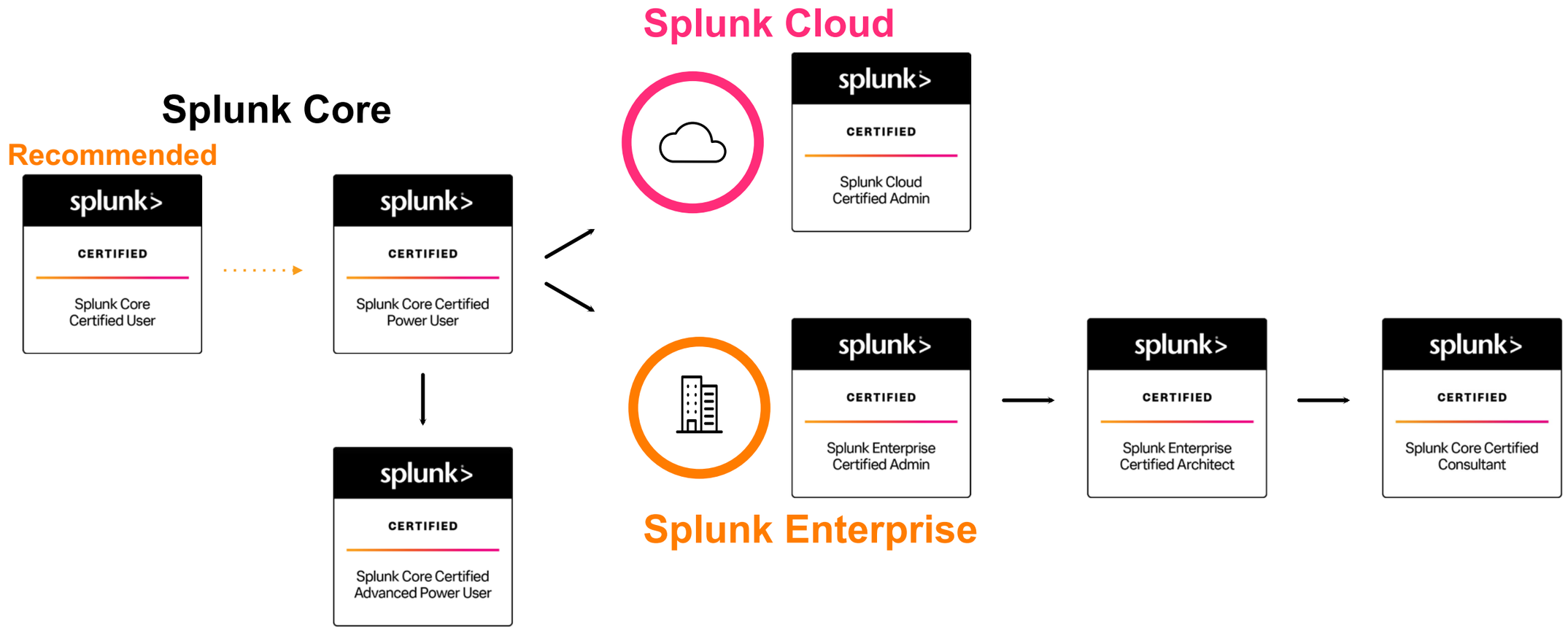
- What is Splunk *
- Introduction to Splunk *
- Using Fields *
- Scheduling Reports and Alerts
- Visualizations
- Statistical Processing
- Working with Time
- Comparing Values
- Result Modification
- Leveraging Lookups and Subsearches
- Correlation Analysis
- Search Under the Hood
- Multivalue Fields
- Search Optimization *

Splunk Certification

Offerings & Requirements

Splunk Core and Beyond

Regardless of which Splunk product you use, it all starts with Splunk Core



Splunk Core Certified Power User

This entry-level certification demonstrates an individual's foundational competence of Splunk's core software

Prerequisite Certification(s):

- None

Prerequisite Course(s):

- None

Splunk Core Certified Power User Exam

Time to study! We suggest candidates looking to prepare for this exam complete Fundamentals 2 **or** the following courses:

- Visualizations
- Statistical Processing
- Working with Time
- Comparing Values
- Result Modification
- Correlation Analysis
- Search Under the Hood
- Introduction to Knowledge Objects
- Creating Knowledge Objects
- Creating Field Extractions
- Data Models
- Using Choropleth

See here for registration assistance.

Congratulations! You are a...

Recommended Next Steps

- Splunk Core Certified Advanced Power User
- Splunk Enterprise Certified Admin
- Splunk Cloud Certified Admin

 | 10 August 2022

Splunk Core Certified Advanced Power User

This certification demonstrates an individual's ability to generate complex searches, reports, and dashboards with Splunk's core software to get the most out of their data

Prerequisite Certification(s):

- Splunk Core Certified Power User

Prerequisite Course(s):

- None

Splunk Core Certified Advanced Power User Exam

Time to study! We suggest candidates looking to prepare for this exam complete Fundamentals 3, Creating Dashboards, and Advanced Searching & Reporting **or** the following courses:

- Using Fields
- Working with Time
- Comparing Values
- Result Modification
- Leveraging Lookups and Subsearches
- Correlation Analysis
- Search Under the Hood
- Multivalue Fields
- Search Optimization
- Creating Field Extractions
- Enriching Data with Lookups
- Data Models
- Using Choropleth
- Introduction to Dashboards
- Dynamic Dashboards

See here for registration assistance.

Congratulations! You are a...

Recommended Next Steps

- Splunk Enterprise Certified Admin
- Splunk Cloud Certified Admin

Thank You