

Log Normalisation Application

Telstra Project - Team Wombat





The Wombat Team



Mustafa Awni



Sothea-Roth Bak



Ben Nguyen



Tina Tang



Michael Thomas



Sayyaf Waseem




► Problem Statement

The Data Normalisation Problem:

- ▶ Process Complexity
- ▶ Redundant work
- ▶ Lack of information
- ▶ Reduced Collaboration



The Product

 Home Upload Search [Sign out](#)

Regex

(?<linux_ftpd_server_date>w+ \d(2) \d(2) \d(2) \d(2)) * (?<linux_ftpd_server_code>w+ \d(5) \d(5)) * (?<ip_address>\d(1,3) \d(1,3) \d(1,3) \d(1,3))

Insert Save

Cisco v1

Compare another log Upload file

Captured Fields

Lines: 18/18

linux_ftpd_log_date	linux_ftpd_server_code	ip_address
06-25 09:20:24	ftpd[31475]	210.118.170.95
06-25 09:20:24	ftpd[31475]	210.118.170.95
06-25 09:20:24	ftpd[31475]	210.118.170.95
06-25 09:20:24	ftpd[31475]	210.118.170.95
06-25 09:20:24	ftpd[31475]	210.118.170.95
06-25 09:20:24	ftpd[31475]	210.118.170.95
06-25 09:20:24	ftpd[31475]	210.118.170.95
06-25 09:20:24	ftpd[31475]	210.118.170.95
06-25 09:20:24	ftpd[31475]	210.118.170.95
06-25 09:20:24	ftpd[31475]	210.118.170.95
06-25 09:20:24	ftpd[31475]	210.118.170.95
06-25 09:20:24	ftpd[31475]	210.118.170.95
06-25 09:20:24	ftpd[31475]	210.118.170.95
06-25 09:20:24	ftpd[31475]	210.118.170.95
06-25 09:20:24	ftpd[31475]	210.118.170.95
06-25 09:20:24	ftpd[31475]	210.118.170.95
06-25 09:20:24	ftpd[31475]	210.118.170.95
06-25 09:20:24	ftpd[31475]	210.118.170.95

A tool that simplifies the process of normalising log data, by providing a UI that assists in the creation of RegExes for unstructured files, and suggesting column types for structured files.

Scenarios

Structured Files

Uploading a structured file like a CSV and getting suggestions for the columns

New Unstructured File

Uploading an unstructured file that is new to the system and has not been encountered before, and building a RegEx for that file

Additional Features

Searching past RegExes

Generating a RegEx field by highlighting a section of a log

2

Demonstration

An overview of our Product



Business value

Log Normalisation Tool

Cost

- Reduces man hours spent on normalising both new and updated log files
- Suggestion feature mitigates errors associated with writing new regexes

Scalability

- An established process allows for easier scaling to service external clients
- Regex database allows for reusability

Security

- Reduces frictions associated with log files, allowing for faster analysis of security logs

What's Next? - 3 Sprints

Sprint 1

Themes

Miscellaneous

Building a New Regular Expression

Comparing With Existing Regular Expression

Epics

User Management

File Import

Building a Live Regular Expression Builder

Saving and Searching a Regular Expression

Compare Unstructured File

Compare Structured File

Stories

Sprint 1 | 12

Login to the System Using my Telstra Account

Upload an Unstructured File From My Computer

Regular Expression Building Bar

Save a Built Regular Expression

Allow Regular Expression Selection to Apply to Imported File

Register to the System

Show Regular Expression Captures

Saving of Individual Fields by Capture Groups

Show Fields Captured by Regular Expression From Imported File

Add Log Vendor Log Type and Documentation with Regular Expression

Allow Selected Regular Expression to be Edited in Live Builder

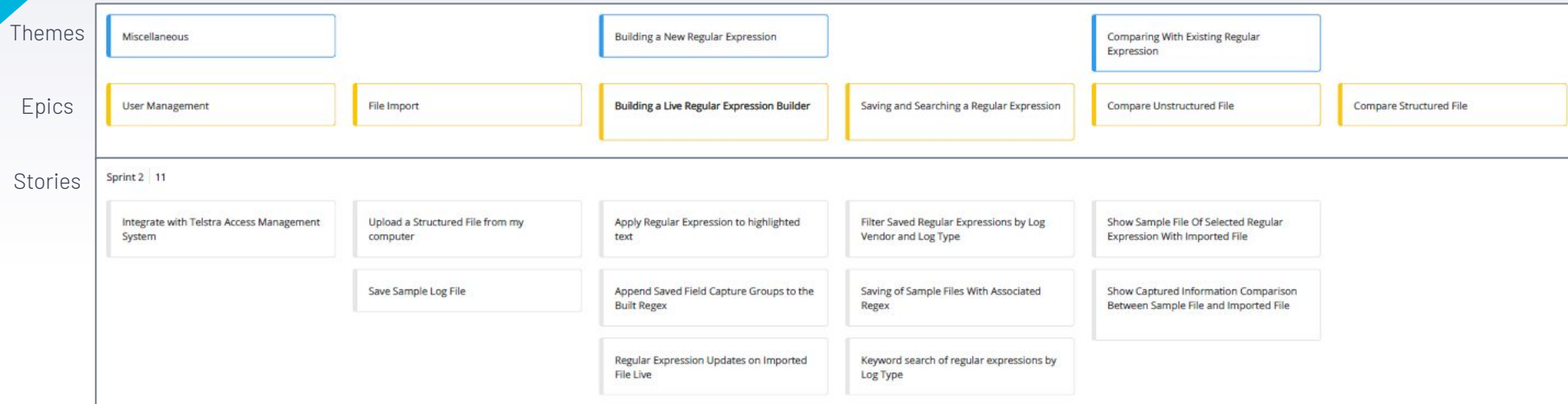
Access a List of Pre-Defined Capture Groups

'Must Have' User Stories

- User log-in
- Unstructured files
- Regex builder, execution, metadata
- Save capture groups

What's Next? - 3 Sprints

Sprint 2



'Should Have' User Stories

- Telstra SSO
- Structured files
- Appending to regex
- Search for saved items

What's Next? - 3 Sprints

Sprint 3

Themes

Miscellaneous

Building a New Regular Expression

Comparing With Existing Regular Expression

Epics

User Management

File Import

Building a Live Regular Expression Builder

Saving and Searching a Regular Expression

Compare Unstructured File

Compare Structured File

Stories

Sprint 3 10

Allow User to See History of Own Contributions on Platform

Provide field suggestions for highlighted text

Keyword search of regular expressions by Log Vendor

Display Documentation Link of Selected Regular Expression

Apply Selected Field Capture Group On Selected Structured Data Field

Allow User to See Contributions on Platform by all Users

Paste Sample Text Into Builder

Search Regular Expression Through Entering a Partial Regular Expression

Keyword search of regular expressions by Log Name

Provide field predictions for structured data

'Could Have' User Stories

- User contributions
- Field suggestions
- Capture groups on structured files
- Search by log name



Handover: Project Artefacts

Motivational Model

Do/Be/Feel Table

Goal Model

Personas

Major users

Expectations and concerns

Moodboard

What kind of impression does this software give people?

User stories

Provides context to the development team

Prototypes

Development team implement based on digital prototypes

Acceptance tests

Check if they meet the criteria of user stories

THANKS!

Any questions?

