WG3 N247
2015-10-13

Title: Japan MB's Report on 29h meeting of WG3

Date Assigned: 2015-10-13

Source: Japan MB

Backward Pointer: None

Document Type: Report

Status: This document is distributed to all the participants according to WG3 Directives..

FUJITSU

shaping tomorrow with you

Japan MB's Reports

2015/10/13

Fujitsu, Ltd. Masatoshi Yoshida
Fujitsu, Ltd. Kotaro Noyama
Fujitsu Social Science Laboratory, Ltd. Chieko Hiramatsu

Outline

FUJITSU

- The next WG3 face-to-face meeting in Tokyo
- RepOSS Collector's status
- Japan MB's comment of "A Study for Design and Implementation of OSMAAM for OSS Proliferation"

FUJITSU

The next face-to-face meeting in Tokyo

Introduction

FUJITSU

- The next WG and chairman’s meeting are held at Ministry of Economy, Trade and Industry (METI) in Japan.
- Therefore, there are some security notices different from the usual. Japan MB kindly ask for your cooperation.

Notices

FUJITSU

- To enter the building
  - All participants of WG must enter the room at the same time.
  - We must not be late. We must not leave and not re-enter the room during the meeting.
  - Participants must go through the check of baggage.
- In the building
  - Do not enter rooms except the meeting rooms and WC.
  - No smoking in the building.
- After the meeting
  - Participants who don’t attend the chairman’s meeting must get out of the building after WG.
  - Participants must go to the exit gate with Japan co-chairperson （Noyama）.

Notices

FUJITSU

- To attend the chairman’s meeting
  - Participants must register in advance. Only people with prior registration can enter the room for chairman’s meeting.
  - Participants must go to the room for the chairman’s meeting with Japan c co-chairperson（Noyama）.
- After the Chairman’s meeting
  - Participants must go to the exit gate with Japan co-chairperson （Noyama）.

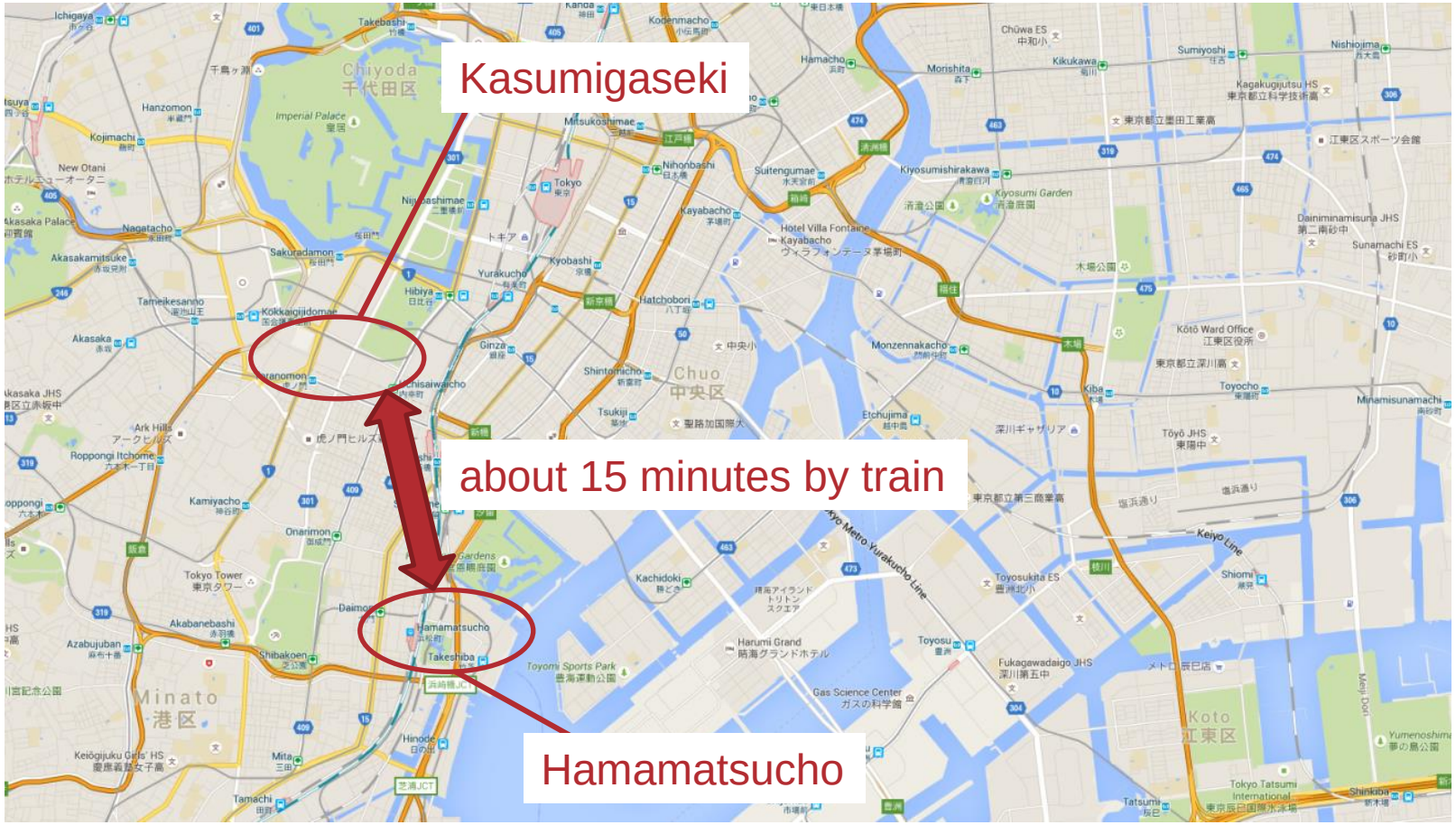
Time schedule plan of WG

- We can use meeting rooms of METI from 13:00.
- We should discuss the following 2 plans.
  - Plan A : We start WG at 13:00.
  - Plan B : We start WG at other meeting rooms in the morning. After having a lunch, we change the location to METI and continue the meeting.

	Plan A	Plan B
9:00		
		WG3
11:00		(lunch & move)
13:00	WG3	WG3
15:00		
15:30	Chairman's Meeting	Chairman's Meeting
17:00		

In the case of plan B

- We will take a train and change the location to Kasumigaseki (the seat of METI) from Hamamatsucho (the seat of Fujitsu).



FUJITSU

RepOSS Collector’s status

Overview

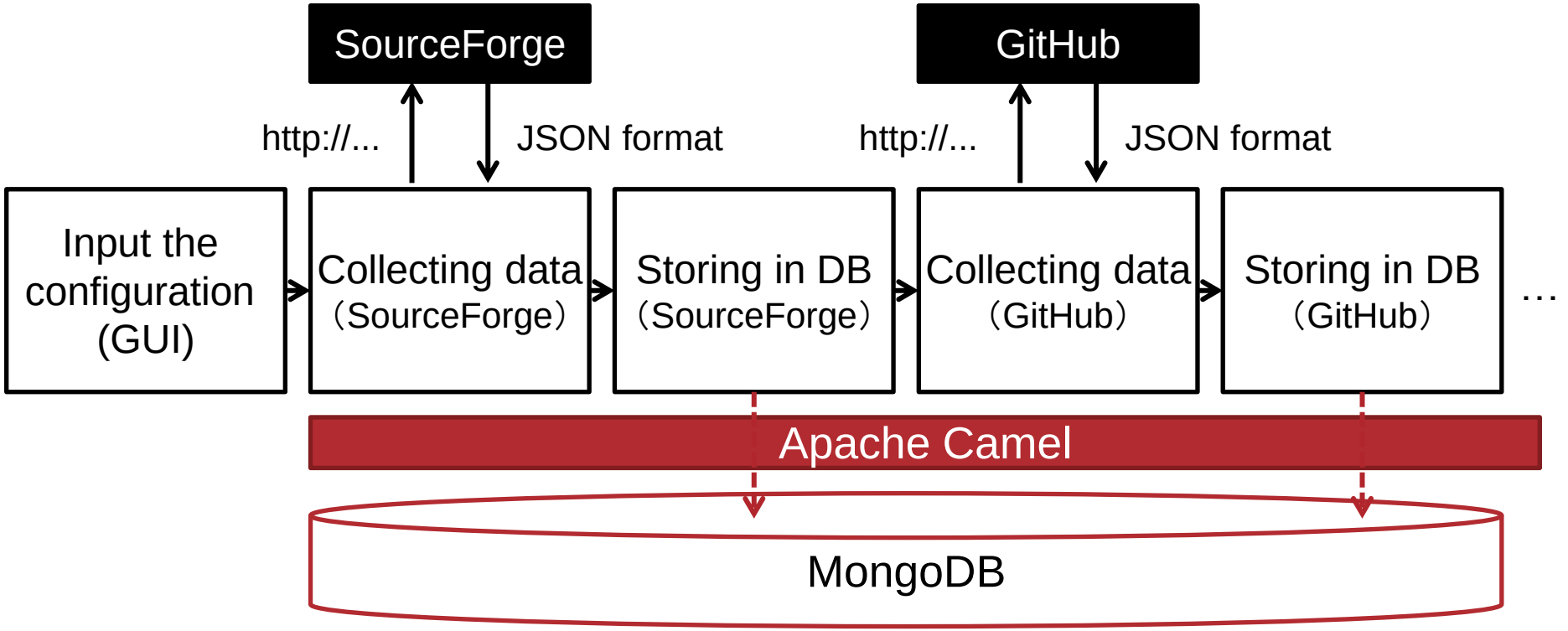
FUJITSU

- Data collecting tool for evaluation of OSS
  - Collecting information of projects from Web services using Web APIs
  - Storing the data into a database (MongoDB)
  - GUI for setting the configuration to collect data
    - Import from the configuration from Excel file
    - Registration / Revision of settings
    - Monitoring the status of progress with GUI

Image of RepOSSCollector

FUJITSU

- Collecting and Storing data by crawling Web services
  - Processing each of data by using Apache Camel
  - Storing the data in MongoBD with JSON format



Target data and implementation status

FUJITSU

Target Web services and information (in beta version)

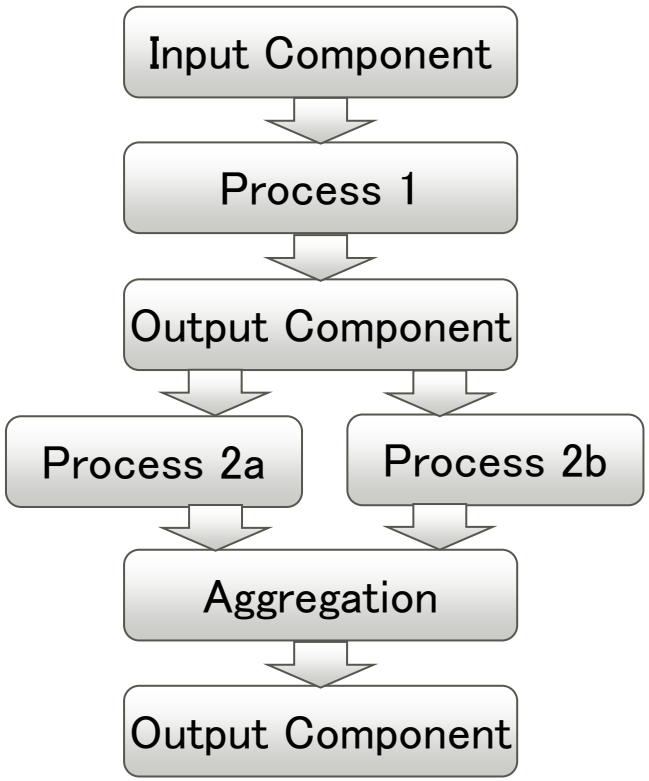
target	collecting information (example)	Impl.
SourceForge	Basic information of the project	Done
GitHub	Basic information of the project	Done
Bugzilla	Bug information	Done
JIRA	Bag information	Done
SlideShare	Slides about the target OSS	Done
Google Custom Search	The number of search results	Done
Google Trend	Transition of the number of search results	Done
Amazon	Books about the target OSS	Not yet
Wikipedia	Information of the project on Wikipedia	Done
CVE	Vulnerability information	Done
Twitter	Tweets with the keywords about the target OSS	Done
Mail archive site	Flow in the mailing list	Not yet
Package information	Version of the package in major Linux distributions	Not yet

(Appendix) Apache Camel

FUJITSU

- Rule-based routing engine
  - Defining the routing rule, how to input / process / output the data for automating data processing

Routing rule



Receive / import data (file, HTTP, DB, etc…)
Data conversion
Send / Restore data (file, HTTP, DB, etc…)
⋮

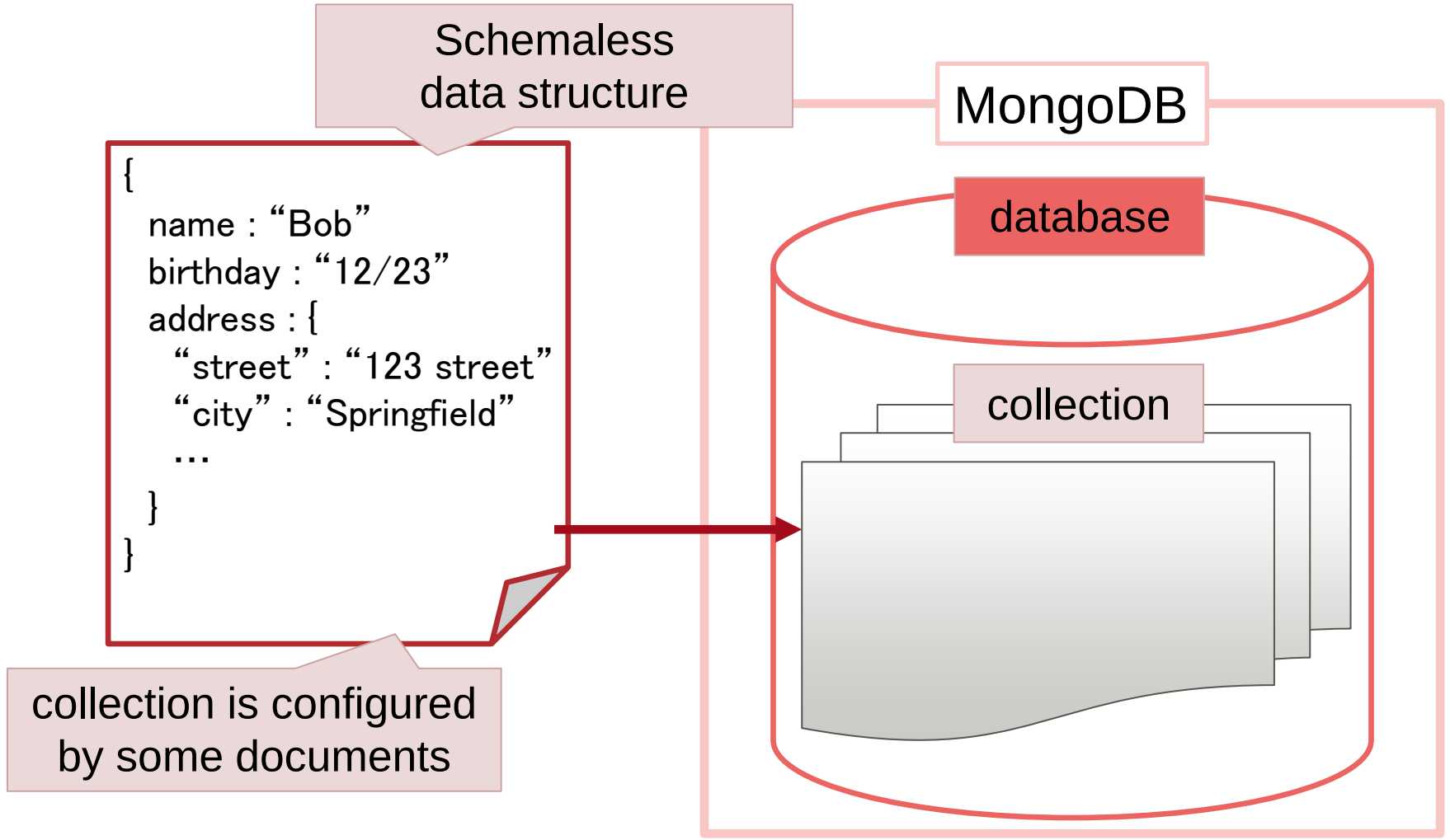
External library / Original implementation are available on the processes

(Appendix) MongoDB

FUJITSU

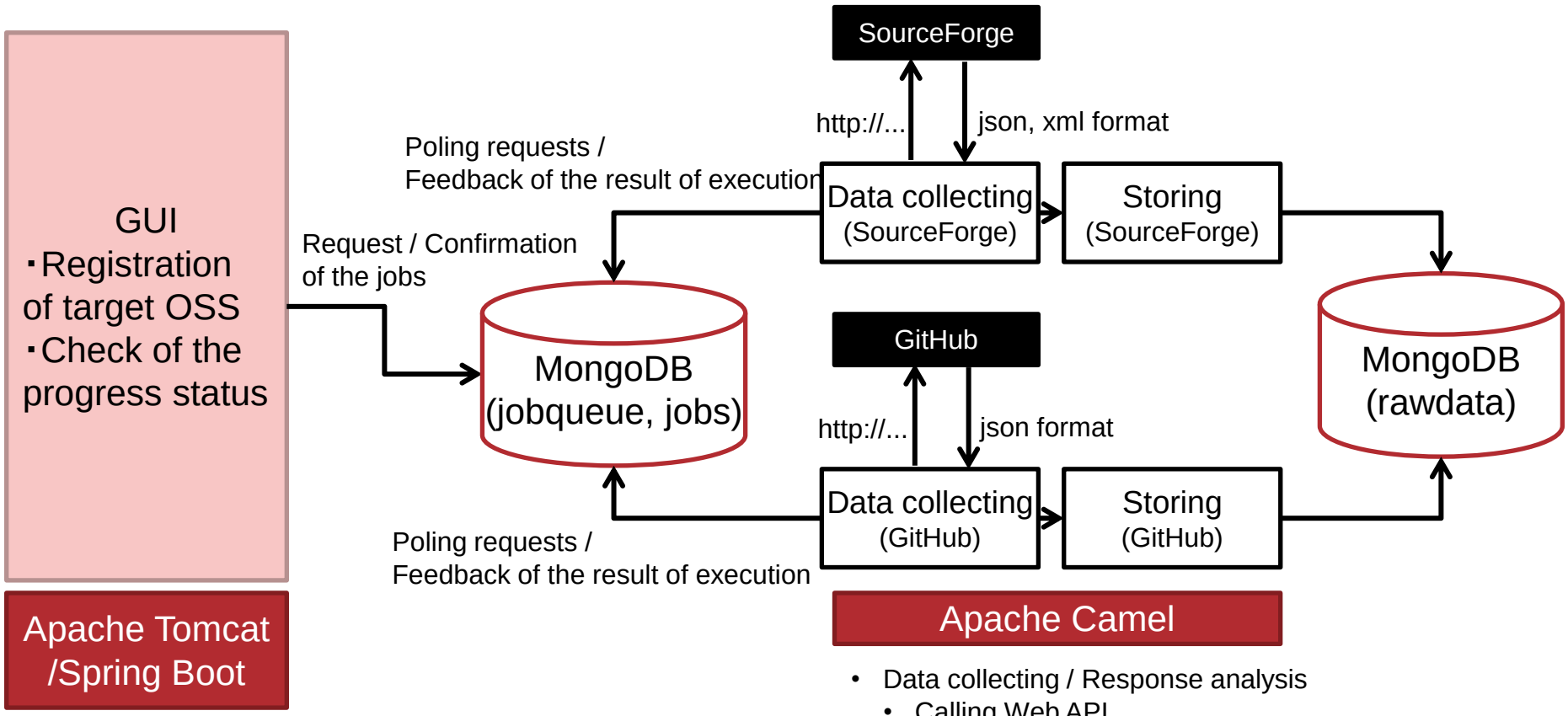
- No SQL, document-oriented database
- Using JSON format to store the data

mongoDB



Detail of RepOSSCollector

Detailed image of processing in RepOSSCollector



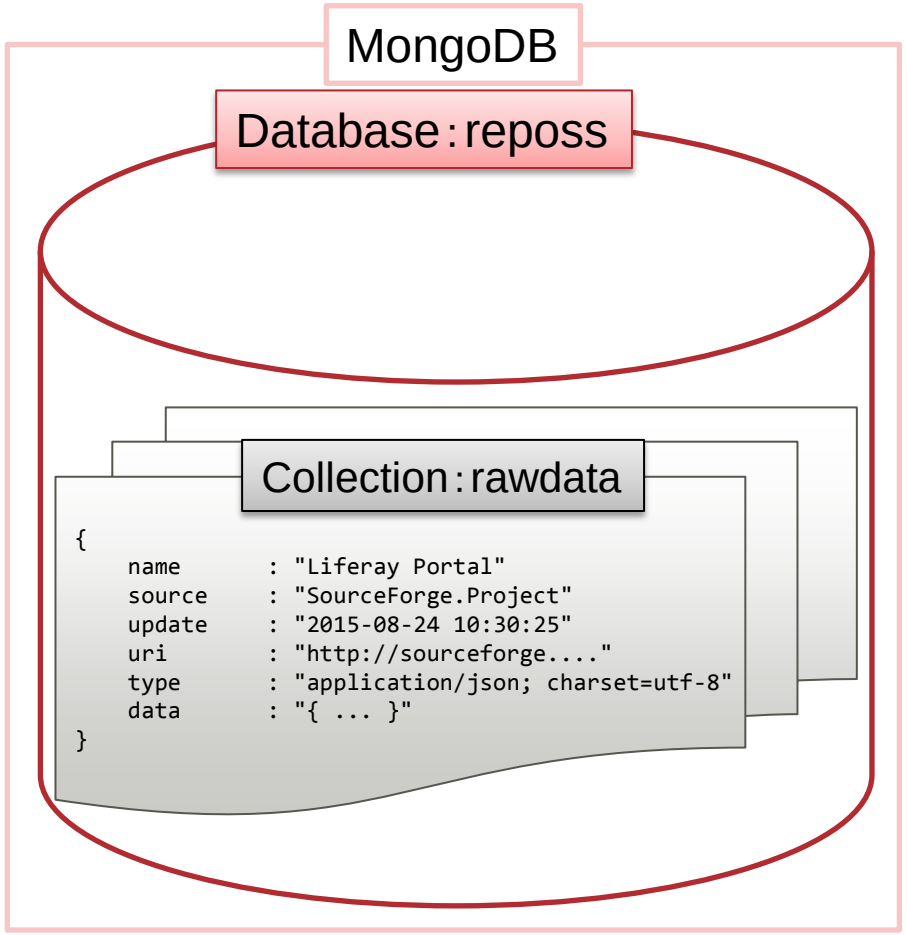
- Registration of target OSS
- Import of the list of target OSS and shelf registration (from Excel Book manipulated by Apache POI)
- Request of the jobs (to jobqueue)
- Check of the progress status

- Data collecting / Response analysis
  - Calling Web API
  - Scraping
- Storing collected data
- Feedback of the result of execution
- Throttling (for flow limitation)

Example of the data collection in MongoDB

FUJITSU

- Collection
  - rawdata
- Items (data type)
  - name (string)
    - name of software
  - source (string)
    - source of information
  - update (string)
    - date of request from Console
      - yyyy-MM-dd HH:mm:ss
  - type (string)
    - value of the Content-Type header*
  - data (string)
    - collected information
      - JSON with response body, or made by partial information on body



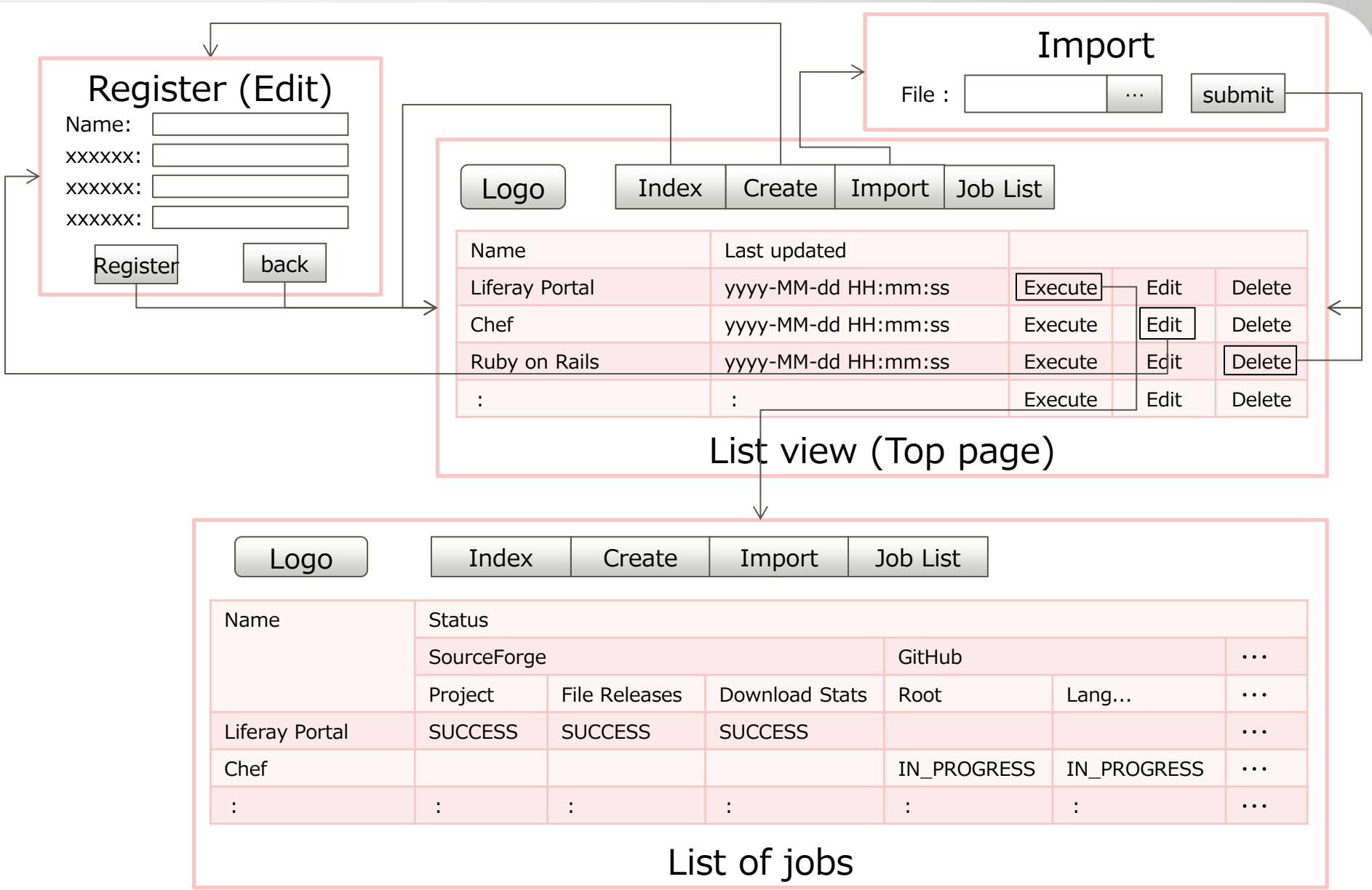
* necessary only if 'data' is raw response body

Functions of GUI

FUJITSU

- RepOSSCollector has the following GUI functions.
  - List view
    - Showing the all registered software information
  - Input / Edit of the configuration
    - Input / edit the configuration to collect data from Web services
  - Import of the configuration
    - Import of the configuration data from Excel Book (Excel2007/2010/2013)
  - Request for collecting data
    - Requesting the collecting tool to start collecting data
  - Monitoring the status
    - Confirmation of the progress (success / processing / error)

Views and transitions (sample)



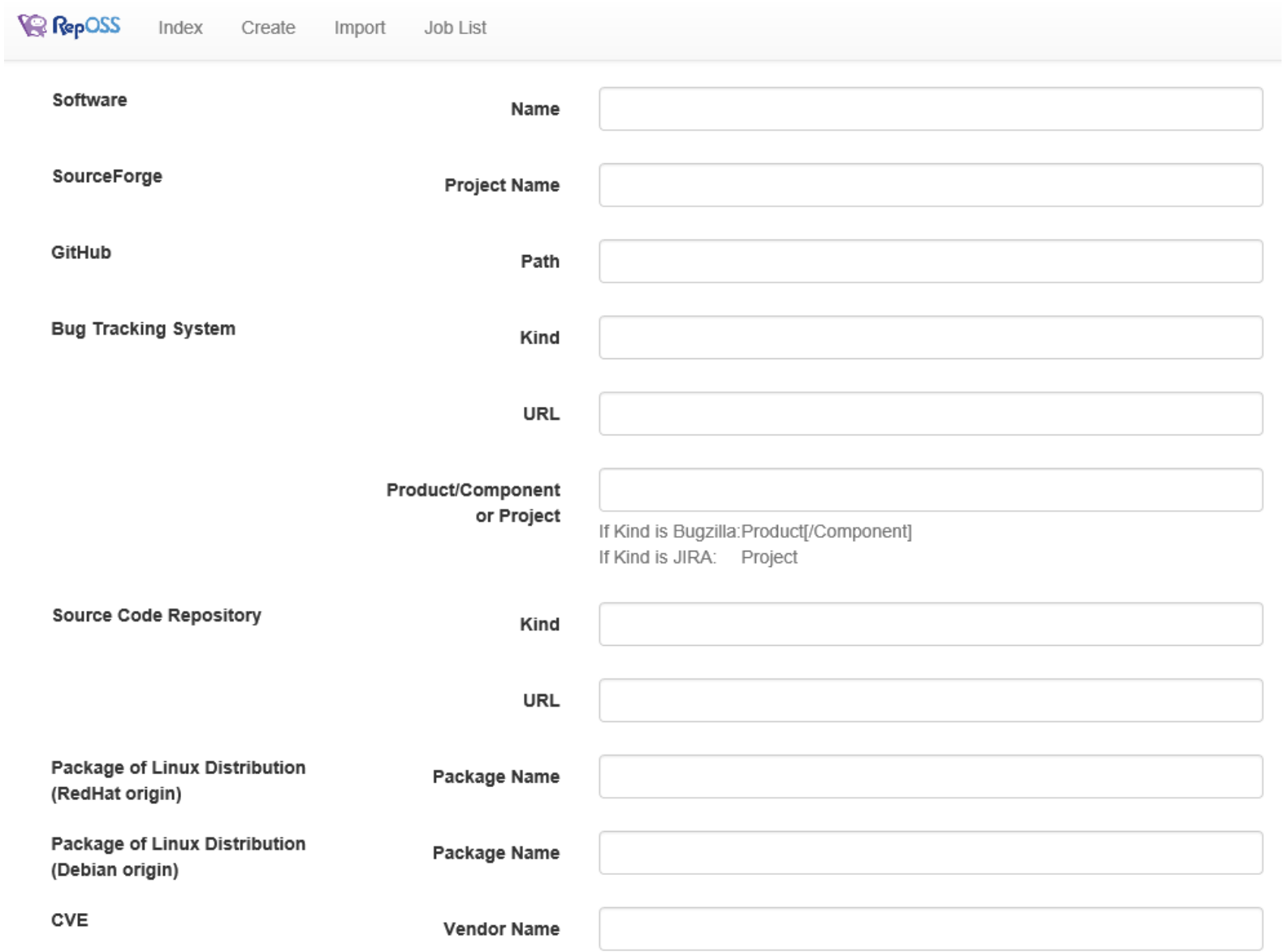
Screenshots (Top Page)

FUJITSU

RepOSS Index Create Import Job List

ID	Name	Last Updated	Operation		
1	Action Mailer	2015-10-06 17:31:59	Execute	Edit	Delete
2	Action Pack	2015-10-06 18:21:53	Execute	Edit	Delete
3	Active Record	2015-10-06 18:30:37	Execute	Edit	Delete
4	Active Support	2015-10-06 18:32:13	Execute	Edit	Delete
5	ActiveMQ	2015-10-06 19:32:17	Execute	Edit	Delete
6	Android	2015-10-06 19:35:45	Execute	Edit	Delete
7	ANTLR	2015-10-06 19:49:52	Execute	Edit	Delete
8	Apache Ant	2015-10-06 20:14:00	Execute	Edit	Delete
9	Apache Axis2	2015-10-06 20:17:11	Execute	Edit	Delete
10	Apache Commons Codec	2015-10-06 20:22:09	Execute	Edit	Delete
11	Apache Commons DBCP	2015-10-06 23:25:52	Execute	Edit	Delete
12	Apache Commons EL	2015-10-07 00:50:41	Execute	Edit	Delete
13	Apache Commons FileUpload	2015-10-07 00:50:50	Execute	Edit	Delete
14	Apache Commons IO	2015-10-07 00:51:03	Execute	Edit	Delete
15	Apache Commons Lang	2015-10-07 00:51:11	Execute	Edit	Delete
16	Apache Commons Logging	2015-10-07 08:38:29	Execute	Edit	Delete

Screenshots (Register)



Screenshots (Job List)

FUJITSU

RepOSS Index Create Import Job List

ID	Name	Date	Status						
			Sourceforge			GitHub		Bug Tracking System	
			Project	File Releases	Download Stats	Root	Languages	Bugzilla	JIRA
48	Cloud Foundry	2015-10-09 09:42:49				SUCCESS	SUCCESS		
47	Clam AntiVirus	2015-10-09 09:42:21	SUCCESS	SUCCESS	SUCCESS	SUCCESS	SUCCESS	SUCCESS	
46	Chef	2015-10-09 09:42:14	SUCCESS	SUCCESS	SUCCESS	SUCCESS	SUCCESS		IN_PROGRESS
45	CGLib	2015-10-09 09:41:36	SUCCESS	SUCCESS	SUCCESS	SUCCESS	SUCCESS		
44	CentOS 7	2015-10-09 09:41:26							
43	CentOS 6	2015-10-09 09:41:17							
42	CentOS 5	2015-10-09 09:41:12							
41	BIND	2015-10-09 09:41:07				SUCCESS	SUCCESS		
40	ASM	2015-10-09 09:40:58							
39	as3corelib	2015-10-09 09:40:43				SUCCESS	SUCCESS		
38	Apache Xerces	2015-10-09 09:40:27				SUCCESS	SUCCESS		SUCCESS
37	Apache Xalan Java	2015-10-09 09:40:15				SUCCESS	SUCCESS		SUCCESS
36	Apache Xalan C++	2015-10-09 09:40:05				SUCCESS	SUCCESS		SUCCESS
35	Apache Velocity	2015-10-09 09:39:47				SUCCESS	SUCCESS		SUCCESS
34	Apache Tomcat 8	2015-10-09 09:39:40				SUCCESS	SUCCESS	SUCCESS	
33	Apache Tomcat 7	2015-10-09 09:39:34				SUCCESS	SUCCESS	SUCCESS	
32	Apache Tomcat 6	2015-10-09 09:39:27				SUCCESS	SUCCESS	SUCCESS	

Document

- Preparing two types of documents
  - Installation Manual
  - Introduction of RepOSSCollector

FUJITSU CONFIDENTIAL

FUJITSU

RepOSS Collector
Installation Manual

Ver. 0.1
2015/10

FUJISU LIMITED
Software Technologies Unit
Software Development Technologies Division
Center of OSS Technology Excellence

FUJITSU
shaping tomorrow with you

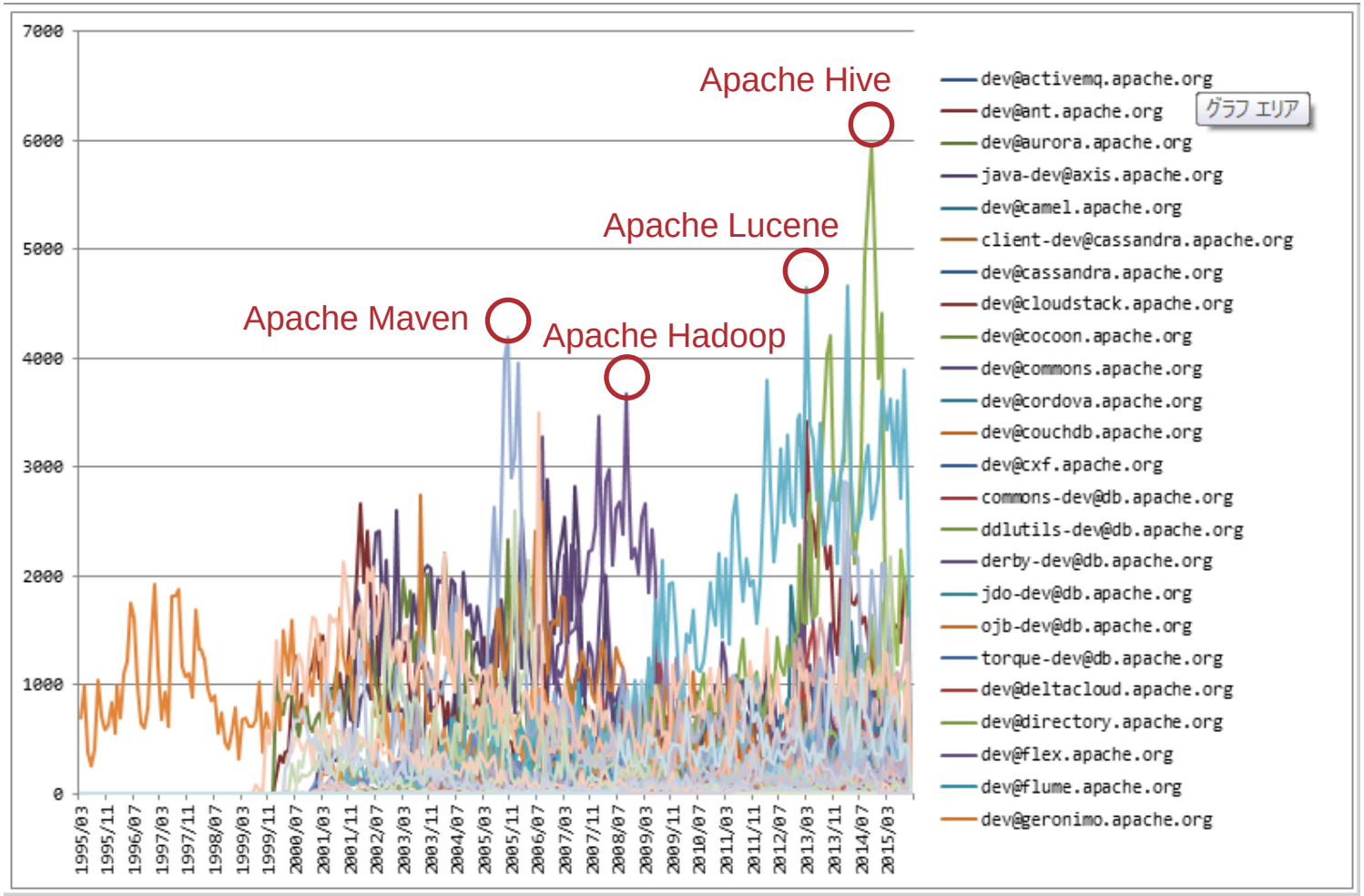
Automated data collecting tool for maturity evaluation of OSS

2015/10/8
FUJITSU LIMITED
Center of OSS Technology Excellence

Example results of analysis (1)

FUJITSU

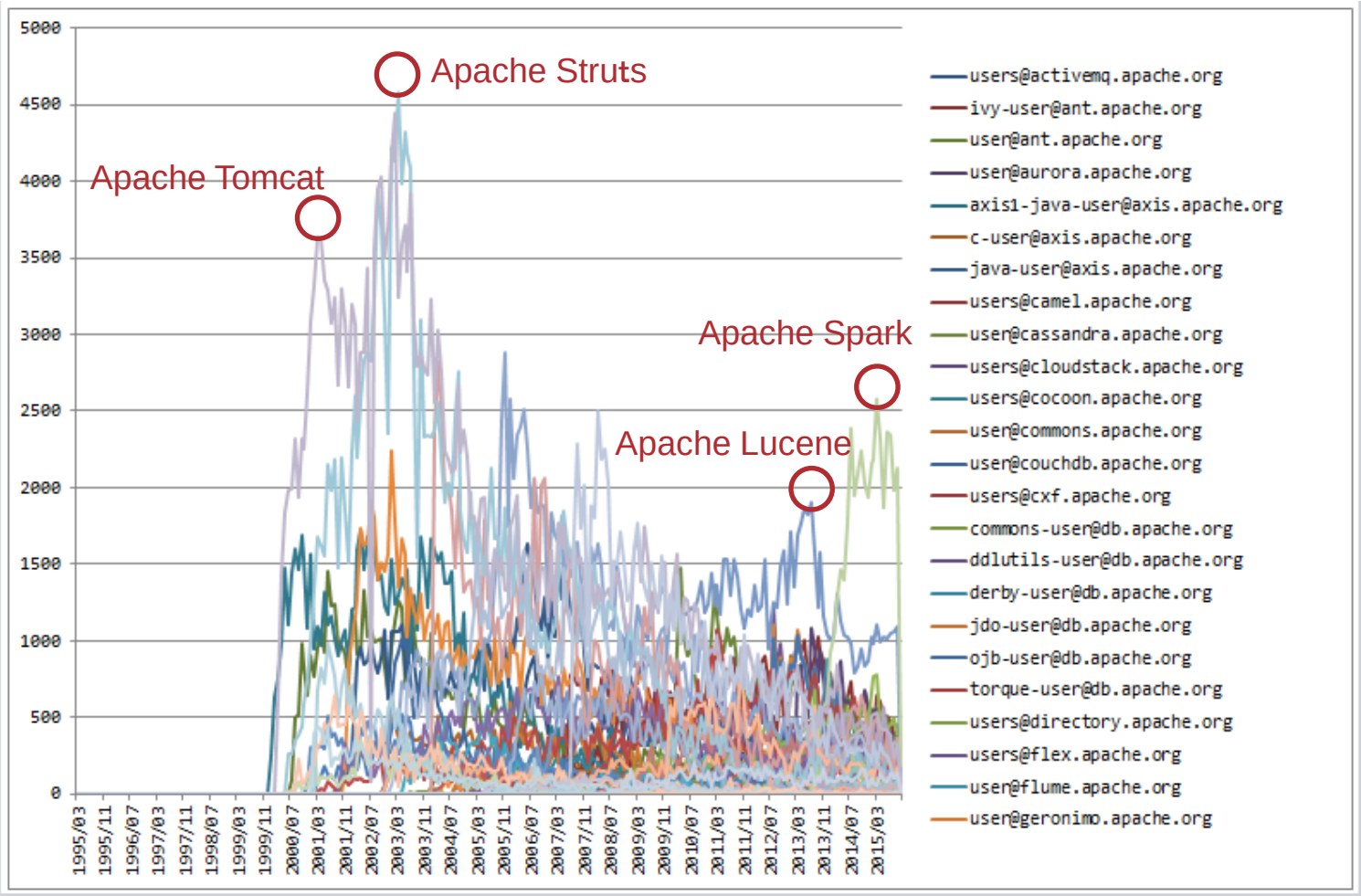
- Transition of the number of mails on mailing list for developers per month（address: dev@xxx）
  - Sample: major Apache Projects



Example results of analysis (2)

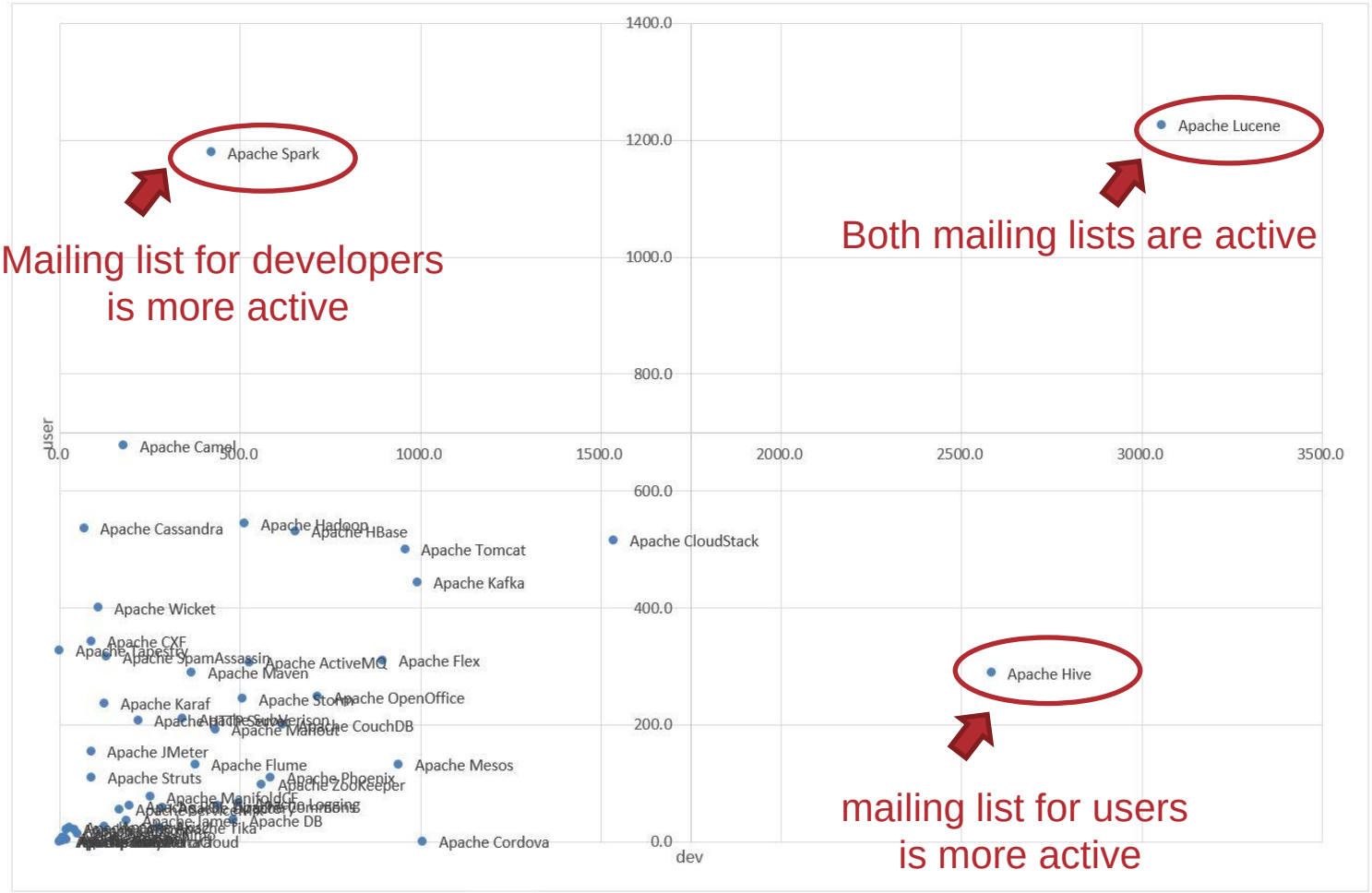
FUJITSU

- Transition of the number of mails on mailing list for users per month（address: user@xxx、users@xxx）
  - Sample: major Apache Projects



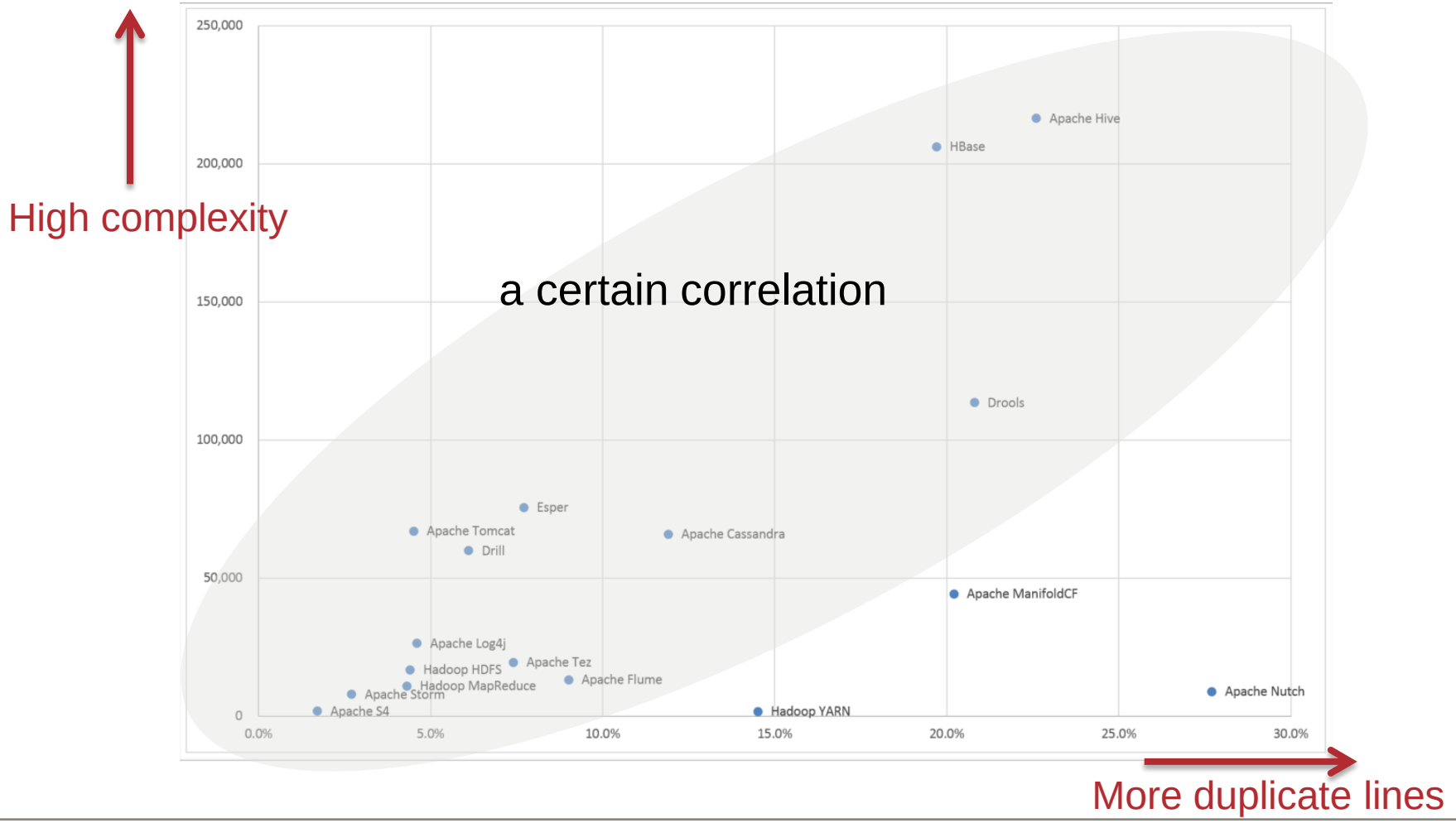
Example results of analysis (3)

- Correlation between the number of mails in mailing list for developers and users
  - Sample: Mails last 3 years of major Apache Projects



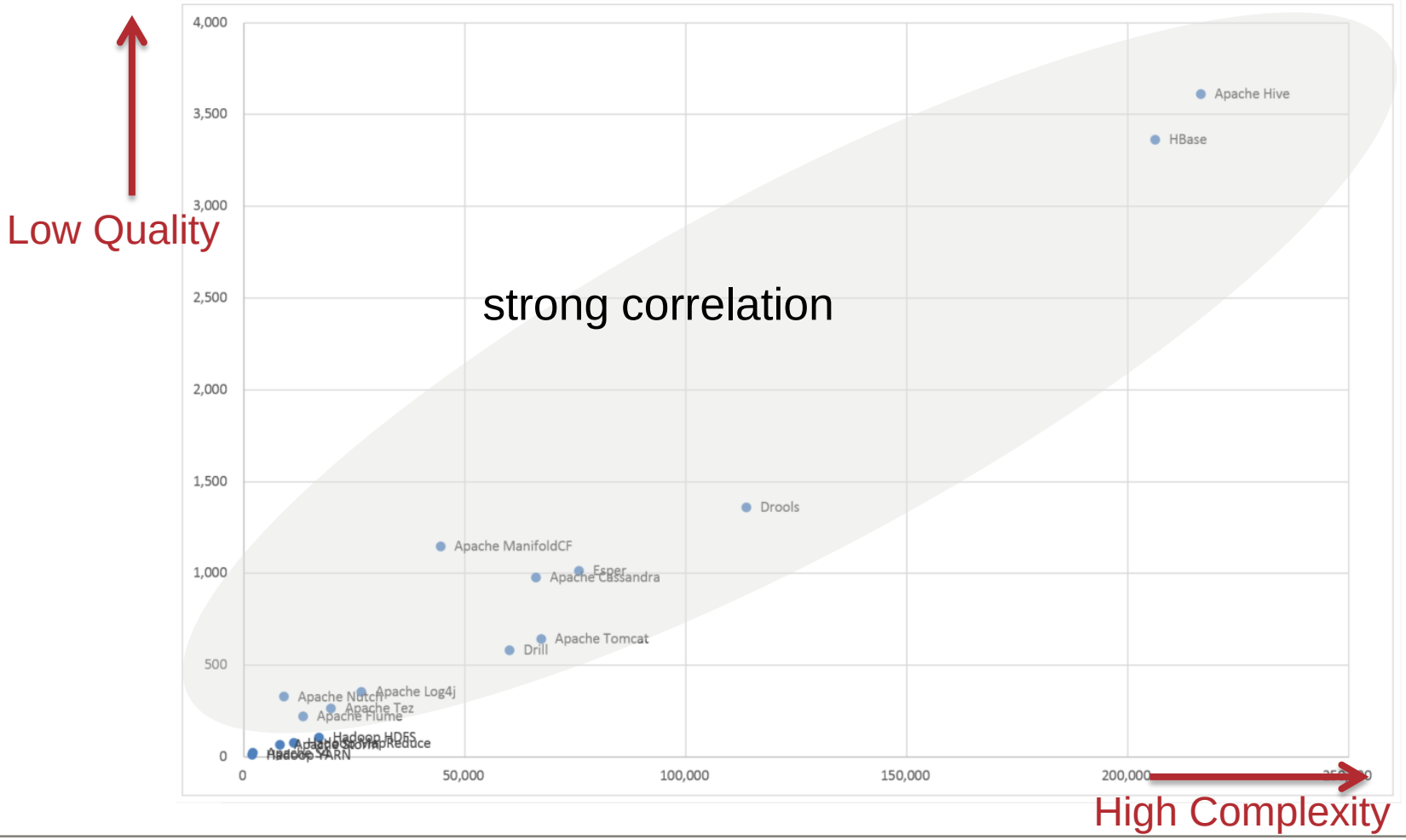
Example results of analysis (4)

- Correlation between the duplication of source code and the complexity of source code
  - Sample: major software developed in Java



Example results of analysis (5)

- Correlation between the complexity and the quality of that by static analysis
  - Sample : major software developed in Java



Repository on GitHub

FUJITSU

- https://github.com/neaosspf-wg3/reposs-collector
- We must decide the software's license.
- Japan MB think "GNU General Public License v2" is better
  - Do you have any opinions?

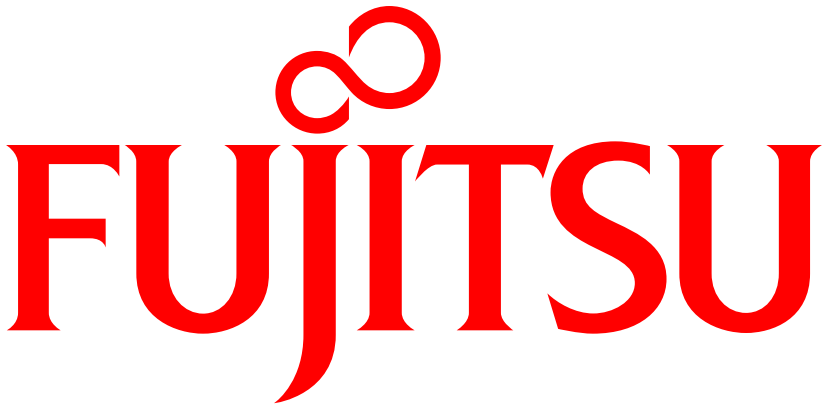
Japan MB's comment of "A Study for Design and Implementation of OSMAAM for OSS Proliferation"

FUJITSU

Our status

FUJITSU

- There is no critical comments from us.
- We find a typo in the English version's paper.
  - p1 Abstract's 15th line: scoree -> score
  - p2 2nd line: indispensible -> (maybe) indispensable
- We finished to translate into Japanese.



shaping tomorrow with you