

CREATING THE KNOWLEDGE ABOUT IT EVENTS

Gilad Barash, Ira Cohen, Eli Mordechai, Carl Staelin, Rafael Dakar
HP-Labs Israel

©2009 HP



EXAMPLE: DEBUGGING PROBLEM USING

The screenshot shows a Mozilla Firefox browser window displaying a search result for a forum post. The search results are for the log entry: "Failed processing http request: report_ss_samples: Failed to acquire lock for publishing sample". The forum post is titled "BAC not collecting data" and is from the "IT Resource Center Forums". The post content describes a problem with BAC installation not displaying data from BPIs or Silescope. A pink callout box highlights the log entry and the forum post title, with a "Get Info" button below it. The forum post content includes a description of the problem, a note about assigning points, and a list of answers from other users.

Search results for log: "Failed processing http request: report_ss_samples: Failed to acquire lock for publishing sample"

IT Resource Center forums - BAC not collecting data
<http://forums11.itrc.hp.com/service/forums/questionanswer.do?threadId=1341011>

03/15/2009 02:27 "Failed processing http request: report_ss_samples, from remoteHost :3.49.40.25 : Failed to acquire lock for publishing sample at..."

Get Info

draft:fm-meng-smp-lock-A<...
 Common Reason Codes
 GI_ASS 101: Fire <(gem)Stone Soup
 LISTSERV 15.5 - DB2-1 - Archives
 Nibble - SVNKit - Users - REPORT requ...
 Df2?1: fail: to open stream: ...
 P: Fw: Processing of multip...

java:431) at
 servlet.java:55) at

Message Information:
 • Message
 • Message
 • Host:
 • Application
 • Log type:
 Failed processing
 com.mercury.am...
 com.mercury.am...
 javax.servlet.http...

Member icons
 HP moderator
 Expert in this area
 Member status

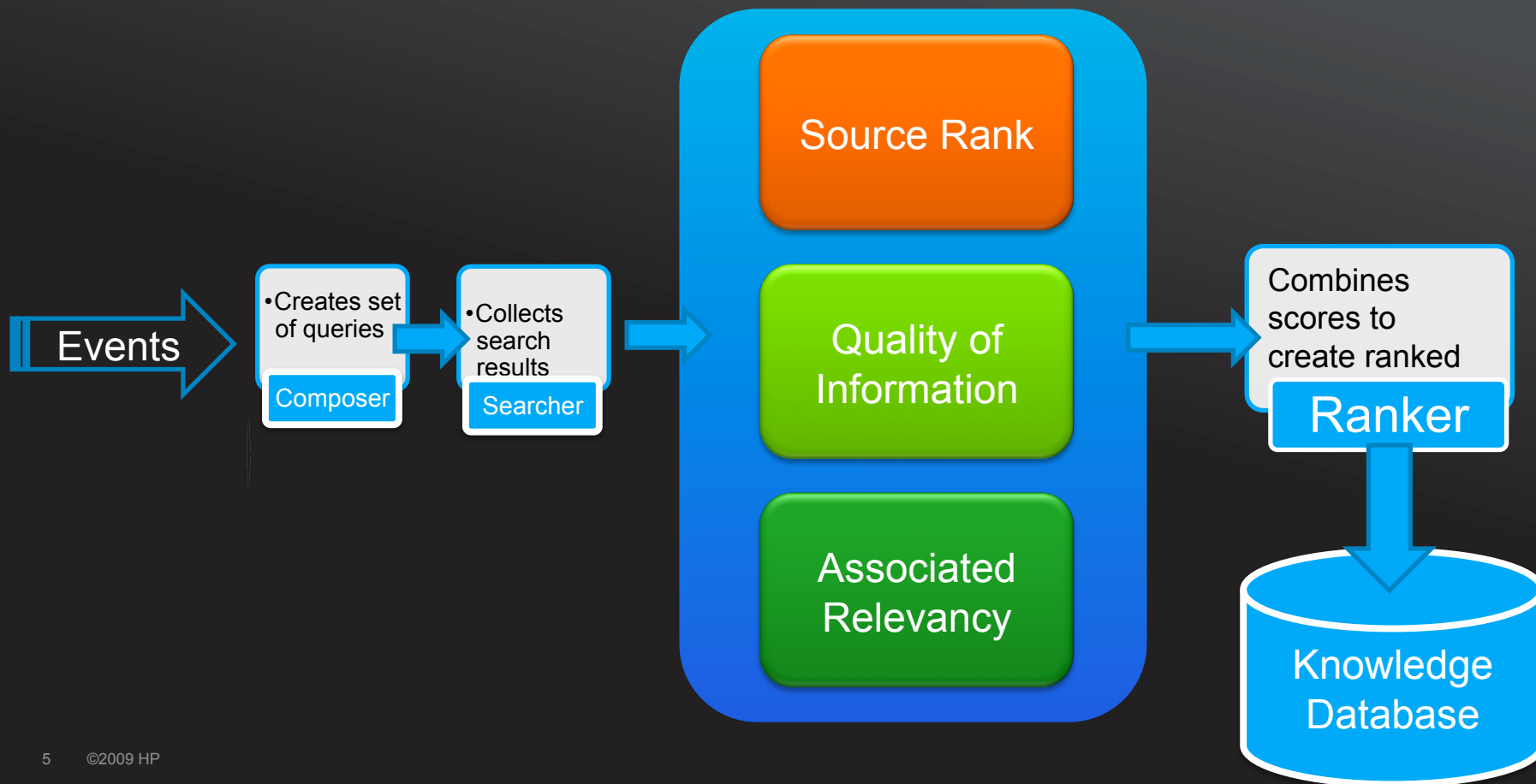
Author Subject: BAC not collecting data
 bisda May 19, 2009 00:09:50 GMT
 I have a problem with my BAC installation not displaying data from BPIs or Silescope. All monitors are appearing grey in the top view.
 I have checked EUM and all profiles are running.
 WDE is listed in System health as being down, and checking http://<BAC_GW>/extmod_mdriv_wrap.diltype=test results in a HTTP status 400 - The request sent by the client was syntactically incorrect ().
 As WDE seems to be the problem here, how do I kick it back up again?
 Note: If you are the author of this question and wish to assign points to any of the answers, please login first. For more information on assigning points, click [here](#).

Sort Answers By: Date or Points

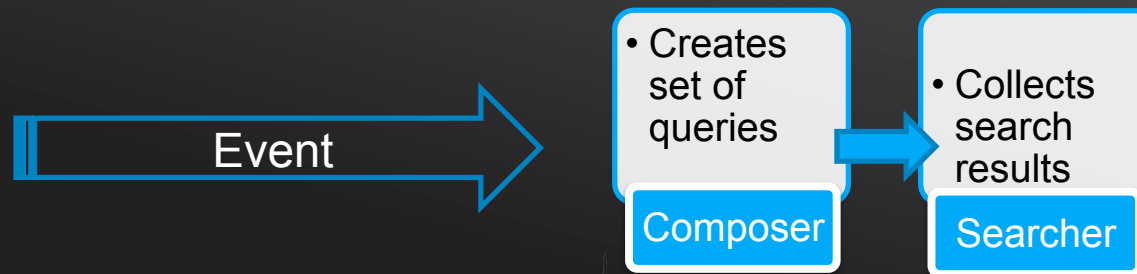
Amit Bandekar May 19, 2009 00:17:58 GMT 1 pts
 Did you try restarting the BAC services?
 --Amit

bisda May 19, 2009 00:24:01 GMT N/A: Question Author
 yes, on both the Gateway and Data Processor, for good measure.

KNOWLEDGE CREATION SYSTEM



COMPOSER /SEARCHER



“EJB spec violation Bean Section 7.10.2 Warning A Session bean must implement directly”

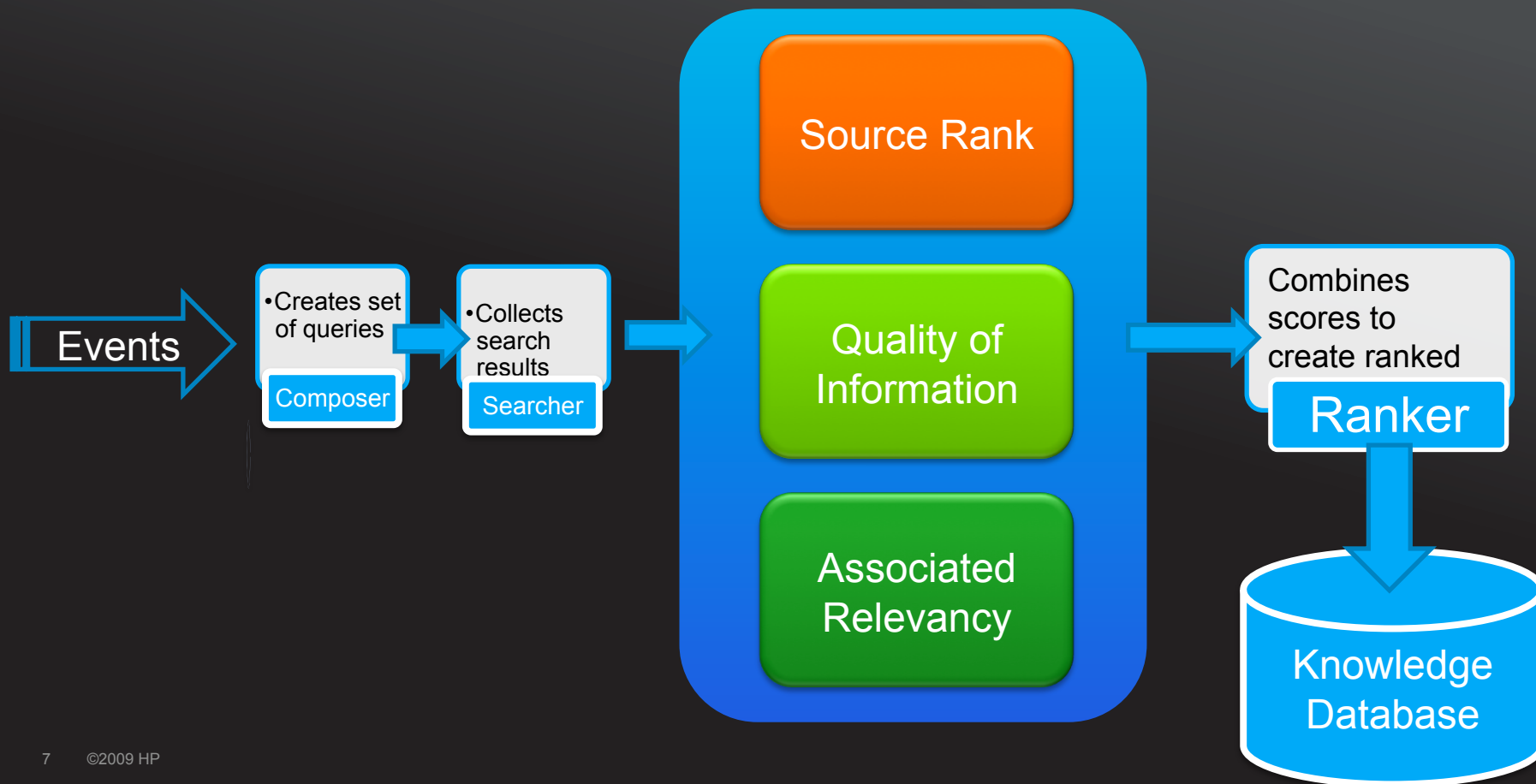
EJB spec violation Bean Section 7.10.2 Warning A Session bean must implement directly

EJB spec violation Bean Section Warning A Session bean must implement directly

EJB spec violation Bean Section Warning Session bean must implement directly

EJB spec violation Bean Section Warning Session bean must implement

KNOWLEDGE CREATION SYSTEM



SOURCE RANKING

- To rank any source, s , we must solve the following set of equations to obtain the ranking u

$$\frac{\partial u}{\partial t} - f_v + \vec{v} \cdot \nabla u = -\frac{\partial \phi}{\partial x} - \left(\frac{g\rho}{\rho_0} \right) \frac{\partial z}{\partial x} - g \frac{\partial \zeta}{\partial x} + F_u + D_u$$

$$\frac{\partial v}{\partial t} - f_u + \vec{v} \cdot \nabla v = -\frac{\partial \phi}{\partial y} - \left(\frac{g\rho}{\rho_0} \right) \frac{\partial z}{\partial y} - g \frac{\partial \zeta}{\partial y} + F_v + D_v$$

$$\frac{\partial T}{\partial t} + \vec{v} \cdot \nabla T = F_T + D_T$$

$$\frac{\partial S}{\partial t} + \vec{v} \cdot \nabla S = F_S + D_S$$

$$\rho = \rho(T, S, P)$$

$$\frac{\partial \phi}{\partial s} = \left(\frac{-gH_z \rho}{\rho_0} \right)$$

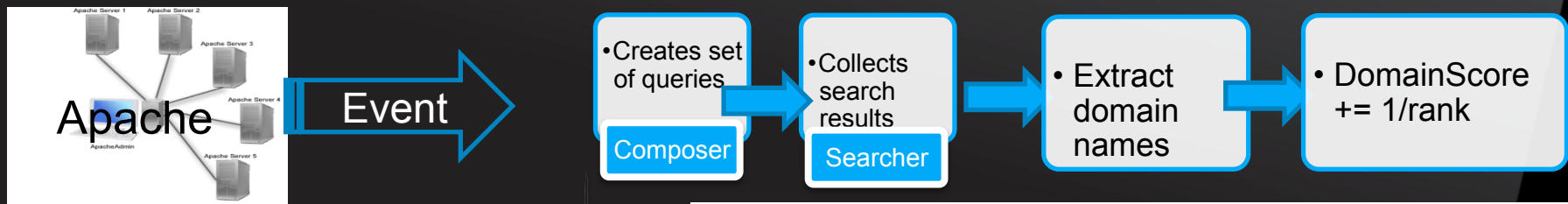
$$\frac{\partial H_z}{\partial t} + \frac{\partial(H_z u)}{\partial x} + \frac{\partial(H_z v)}{\partial y} + \frac{\partial(H_z \Omega)}{\partial s} = 0$$



SOURCE RANKING

Question: Which sources of documents (e.g., domains in www) are most relevant to the system I'm working on?

Method:



“EJB spec violation Bean Section 7.10.2 Warning bean must implement directly”

hp.com	1	+= 1/1
hp.com	2	---
hp.com	3	+= 1/3
epfl.ch	4	+= 1/4
epfl.ch	5	+= 1/5
jboss.org	6	+= 1/6
oracle.com	7	+= 1/7
jboss.org	8	+= 1/8

- Repeat above for each event
- Rank sources based on DomainScore

SOURCE RANKING: EXAMPLE RESULTS

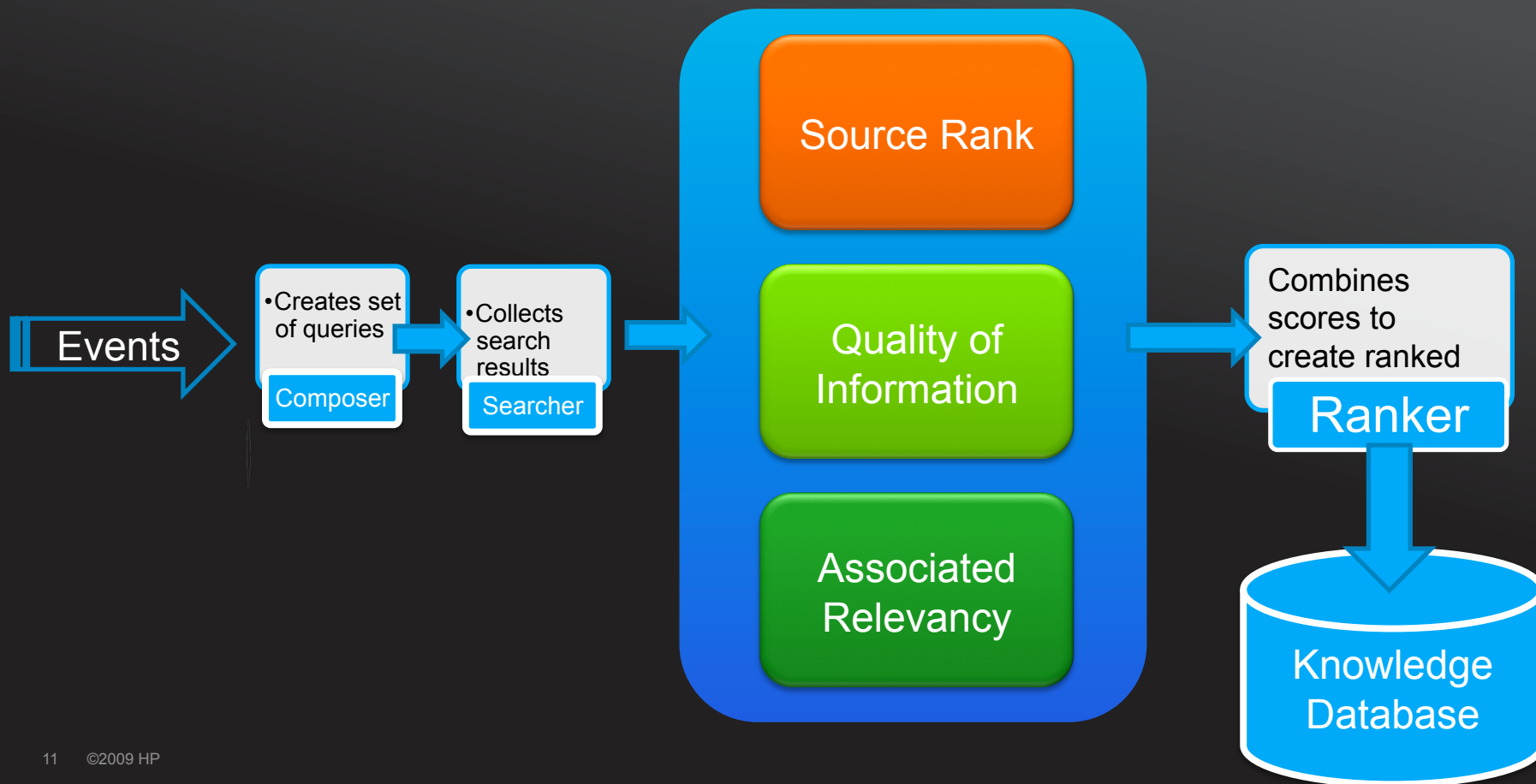
Networked printer logs

Multiple office HP laserjets. Logs collected from Microsoft event log.

Rank	Domain	Domain Score
1	hp.com	59.36915
2	microsoft.com	51.9311
3	eggheadcafe.com	36.79568
4	experts-exchange.com	33.69552
5	forums.techarena.in	25.29344
6	pcreview.co.uk	14.20567
7	tech-archive.net	14.05515
8	soft32.com	13.24757



KNOWLEDGE CREATION SYSTEM



QUALITY OF INFORMATION

- A measure of how fit the information is for a purpose
- Research Challenges:
 - Identifying important measures
 - Providing mechanisms to quantify and predict them



QUALITY OF INFORMATION FOR FORUMS

- Extract generic quality related measures for forums and incidents:
 - Ranking of users
 - Number of replies
 - Duration
 - ...

Challenge: Automatic methods for extraction from any forum type.

- Infer quality measures:
 - Was the question answered?
 - Which post(s) are answers / which are not
 - Difficulty of solution
 - ...

Challenges:

- How to infer them? Can they be learned from other QOI measures?



PROCESS: INFER “ANSWERED/NOT ANSWERED”

Extract

- Collect forum threads
- Extract and compute generic features

Train

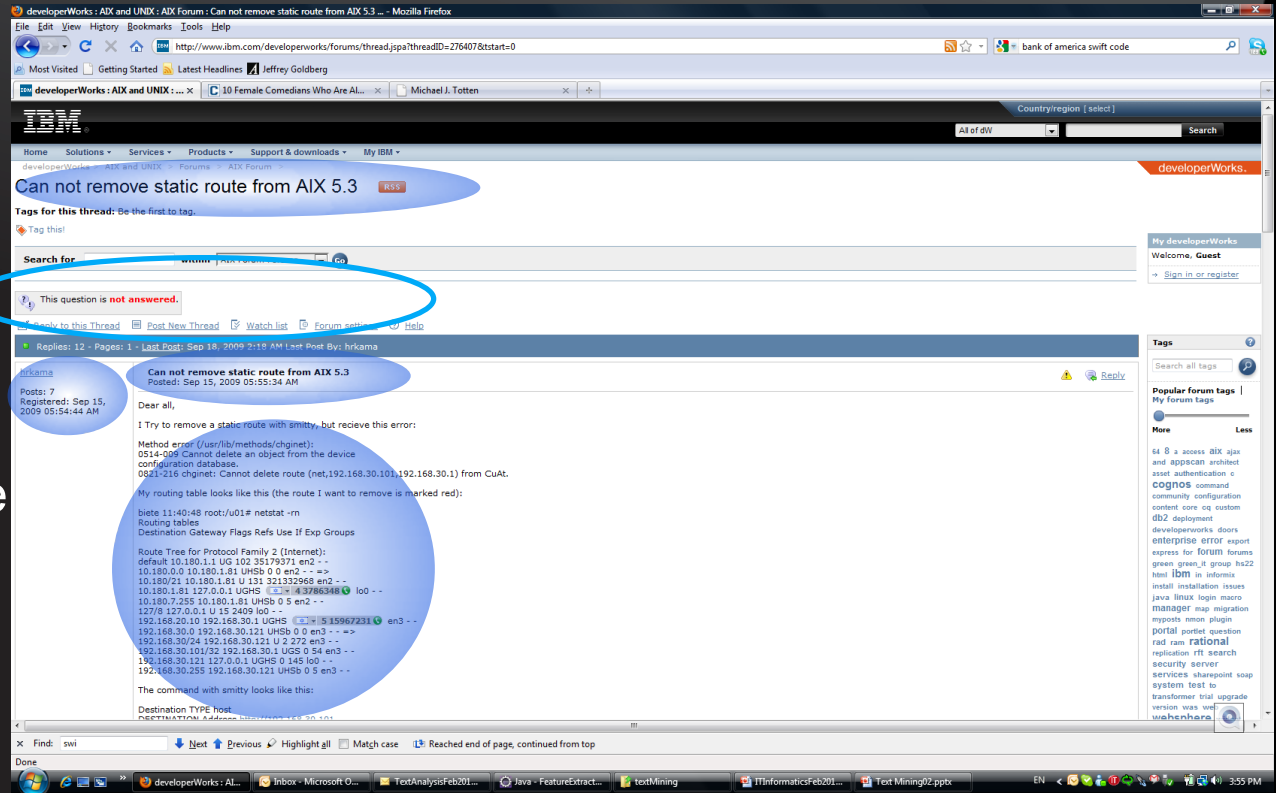
- Obtain labeled examples
- Train classifiers

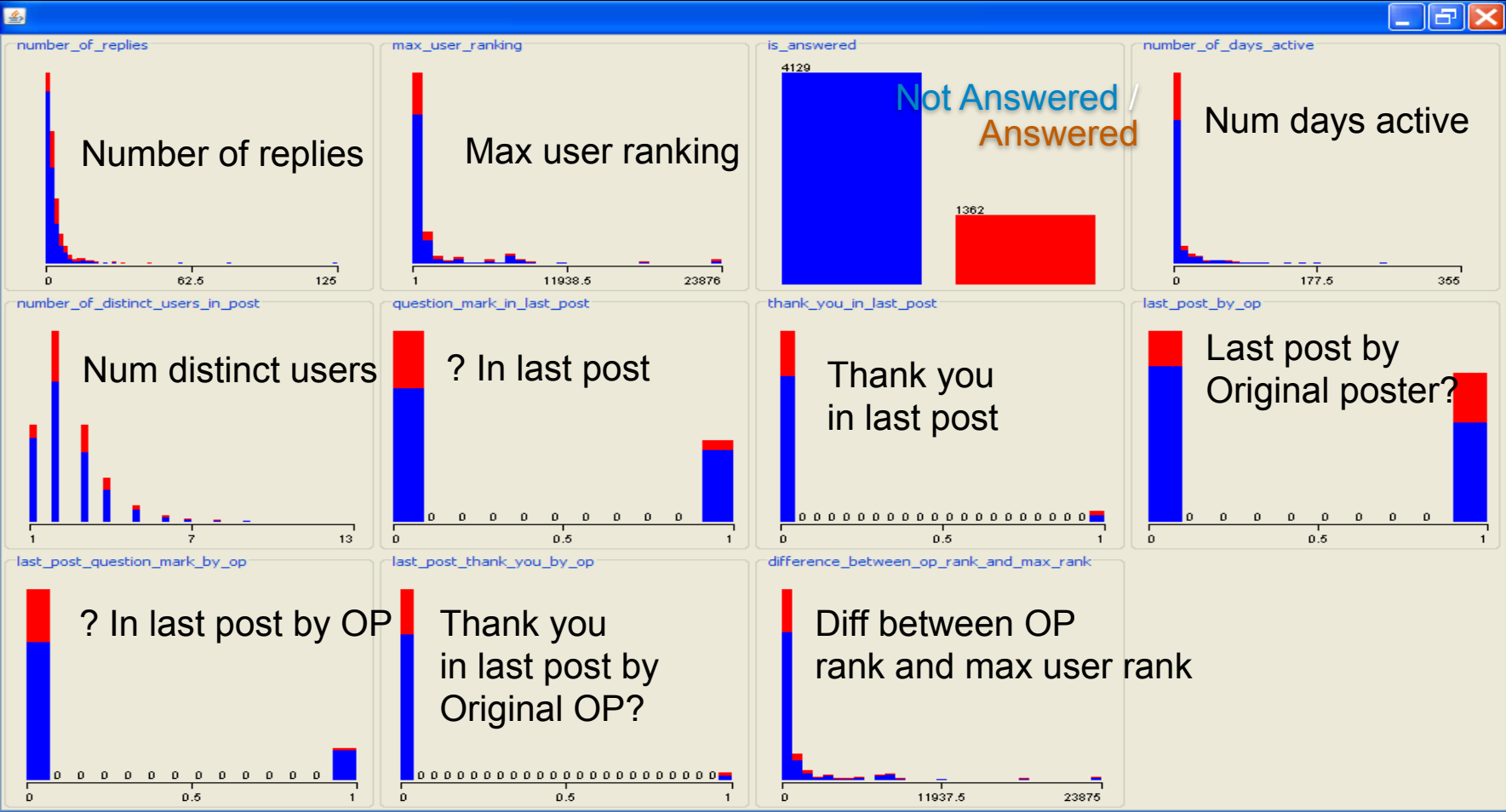
Classify

- Use classifiers to label any forum thread

EXTRACT

- Java utility to download user forums and screen-scrape content elements
- Analyze and aggregate structured and unstructured features





PROCESS: INFER “ANSWERED/NOT ANSWERED”

Extract

- Collect forum threads
- Extract and compute generic features

Train

- Obtain labeled examples
- Train classifiers

Challenge: Label Noise

Users are responsible to change question from “not answered” to “answered”

LABEL NOISE – EXAMPLE

IBM
Country/region [select]
All of dW

Home Solutions Services Products Support & downloads My IBM
developerWorks > WebSphere > Forums > WebSphere Application Server

Using Log4j in WebSphere v6.1

Tags for this thread: Be the first to tag.

This question is not answered.

Replies: 11 - Pages: 1 - Last Post: Sep 17, 2009 2:39 AM Last Post By: Webspherev61

zs4352
Posts: 6
Registered: Jul 26, 2004 04:40:57 PM

Using Log4j in WebSphere v6.1
Posted: Sep 22, 2008 12:51:21 PM

Anyone here can help me on using Log4j in Websphere?

I use RAD7 (run time environment: WebSphere v6.1) to develop a Enterprise Application myProject. myProject includes a web project myWeb, an EJB project myEJB. I also created a java project commons-logging that imported commons-logging-1.1.1.jar and common-logging.properties file.

this is my commons-logging.properties content:

[zs4352](#)

Posts: 6
Registered: Jul 26, 2004 04:40:57 PM

Re: Using Log4j in WebSphere v6.1
Posted: Sep 24, 2008 04:31:22 PM ↑ in response to: [BSteinbach's post](#)

BSteinbach,

Thank you very much. That works well.

Thanks again.

Permission:
getClassLoader : Access denied (java.lang.RuntimePermission getClassLoader)

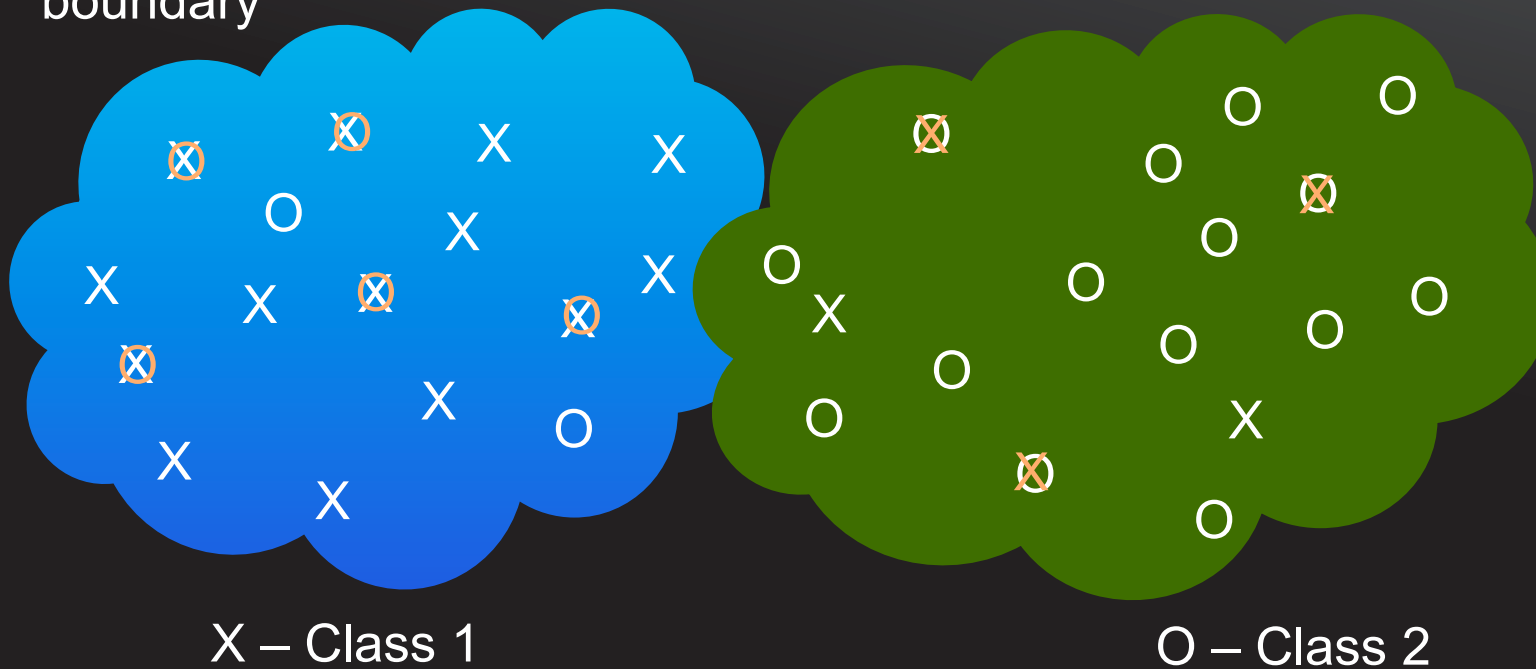
Code:
org.apache.log4j.helpers.Loader in (file:/C:/Documents and Settings/zs4352/IBM/rational/sdp7.0/workspace/myProject/log4j-1.2.15.jar)

Stack Trace:
java.security.AccessControlException: Access denied (java.lang.RuntimePermission getClassLoader)



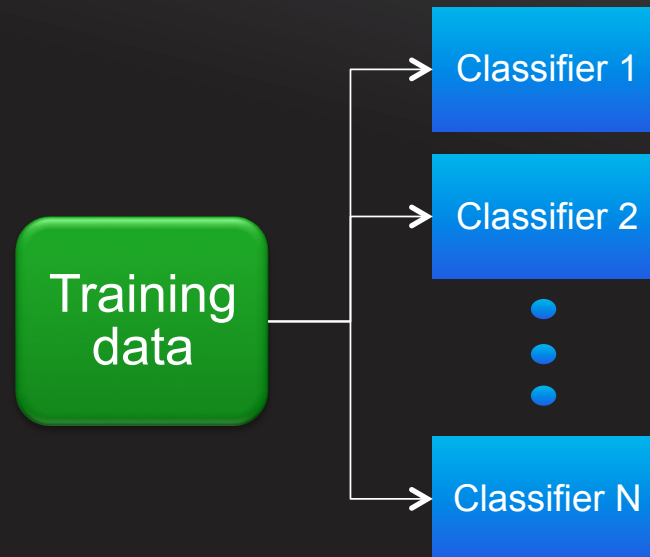
LABEL NOISE : THE PROBLEM

- Random label noise – does not occur around any class boundary



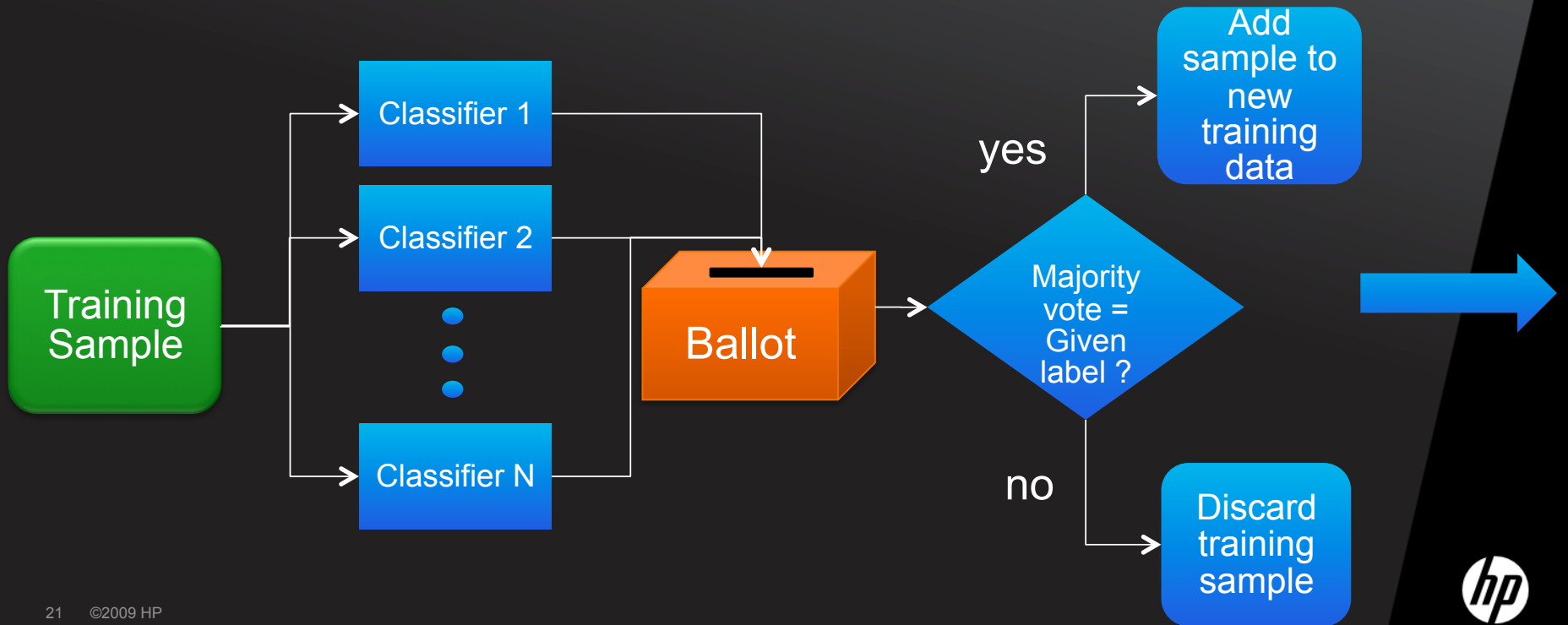
SOLUTION: ENSEMBLE METHOD*

- Train N Classifiers with all training data



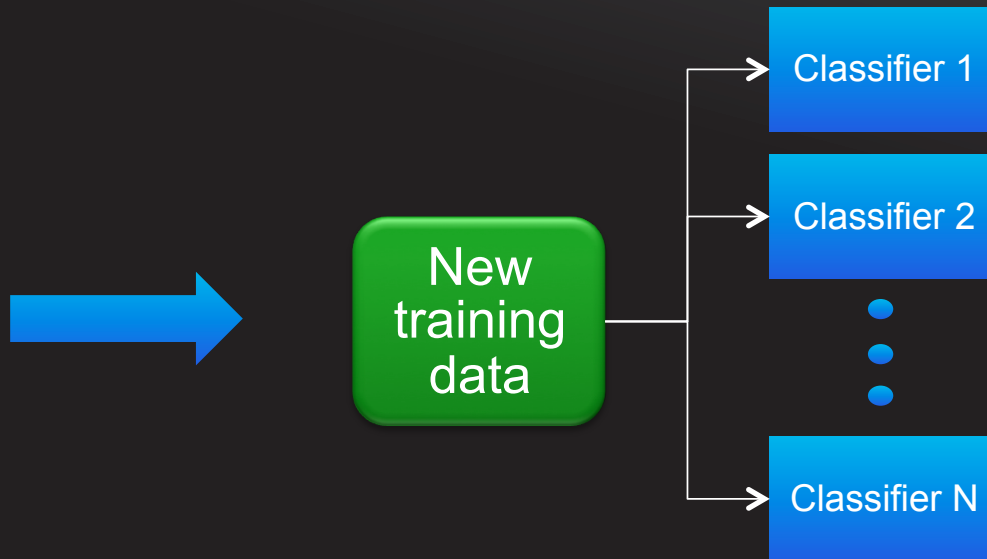
SOLUTION: ENSEMBLE METHOD

– Classify each sample with each classifier



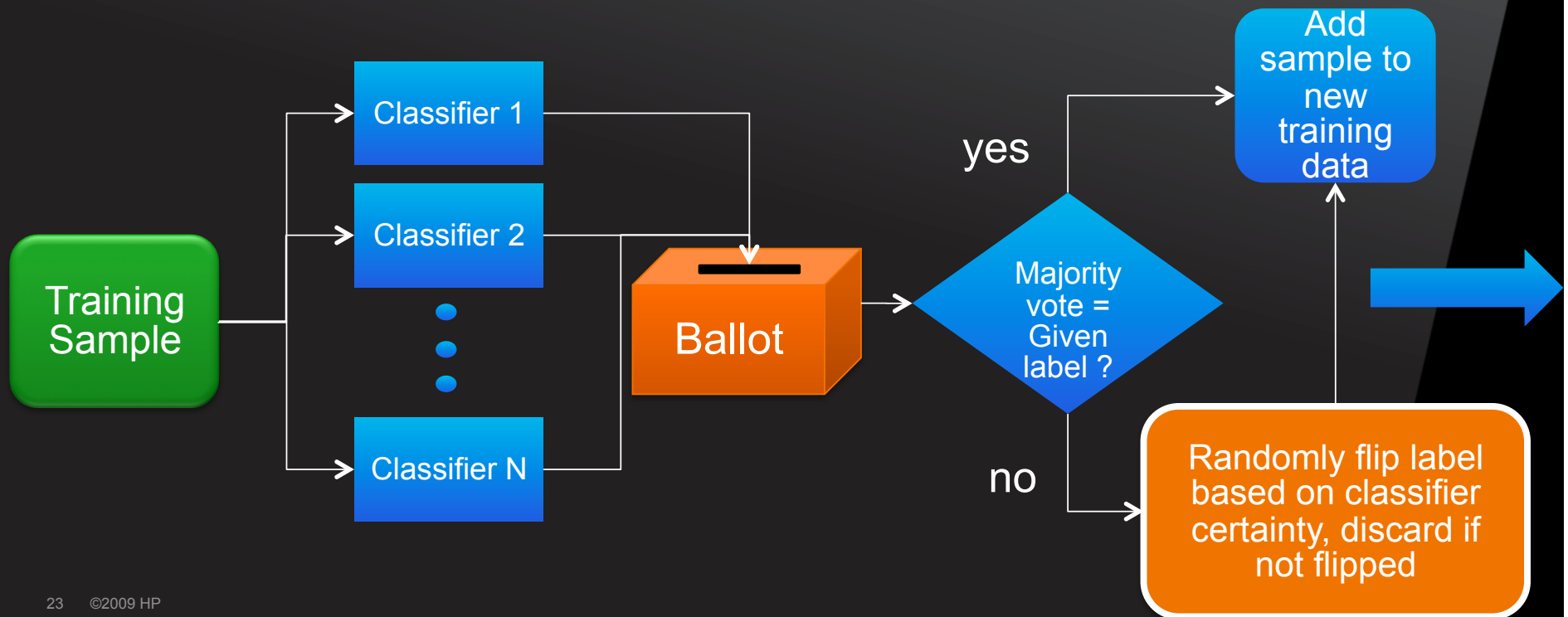
SOLUTION 1: ENSEMBLE METHOD

– Train Classifier(s) with new training data



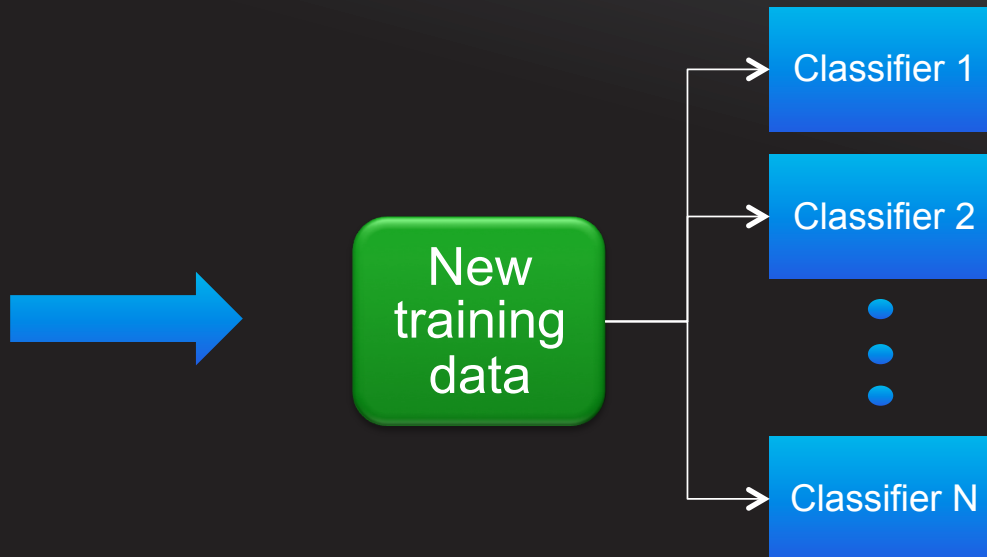
SOLUTION: ENSEMBLE METHOD + FLIP

– Classify each sample with each classifier



SOLUTION: ENSEMBLE METHOD + FLIP

– Train Classifier(s) with new training data



NOISY LABELS: ACCURACY RESULTS

Method\% Noise	0%	10%	20%	30%	40%
No Noise Filter	0.78	0.75	0.73	0.69	0.65
Ensemble filter	0.78	0.77	0.75	0.72	0.69
Ensemble flip filter	0.78	0.77	0.75	0.73	0.70

*Results on UCI machine learning repository data



PROCESS: INFER “ANSWERED/NOT ANSWERED”

Extract

- Collect forum threads
- Extract and compute generic features

Train

- Obtain labeled examples
- Train classifiers

Challenge: Transferability

Can a classifier trained on Forum A be used to classify threads on Forum B?

TRANSFERABILITY EXPERIMENT

Extract

- Collected 5500 Oracle forum threads, 1300 IBM forum threads
- Extracted 10 features

Train

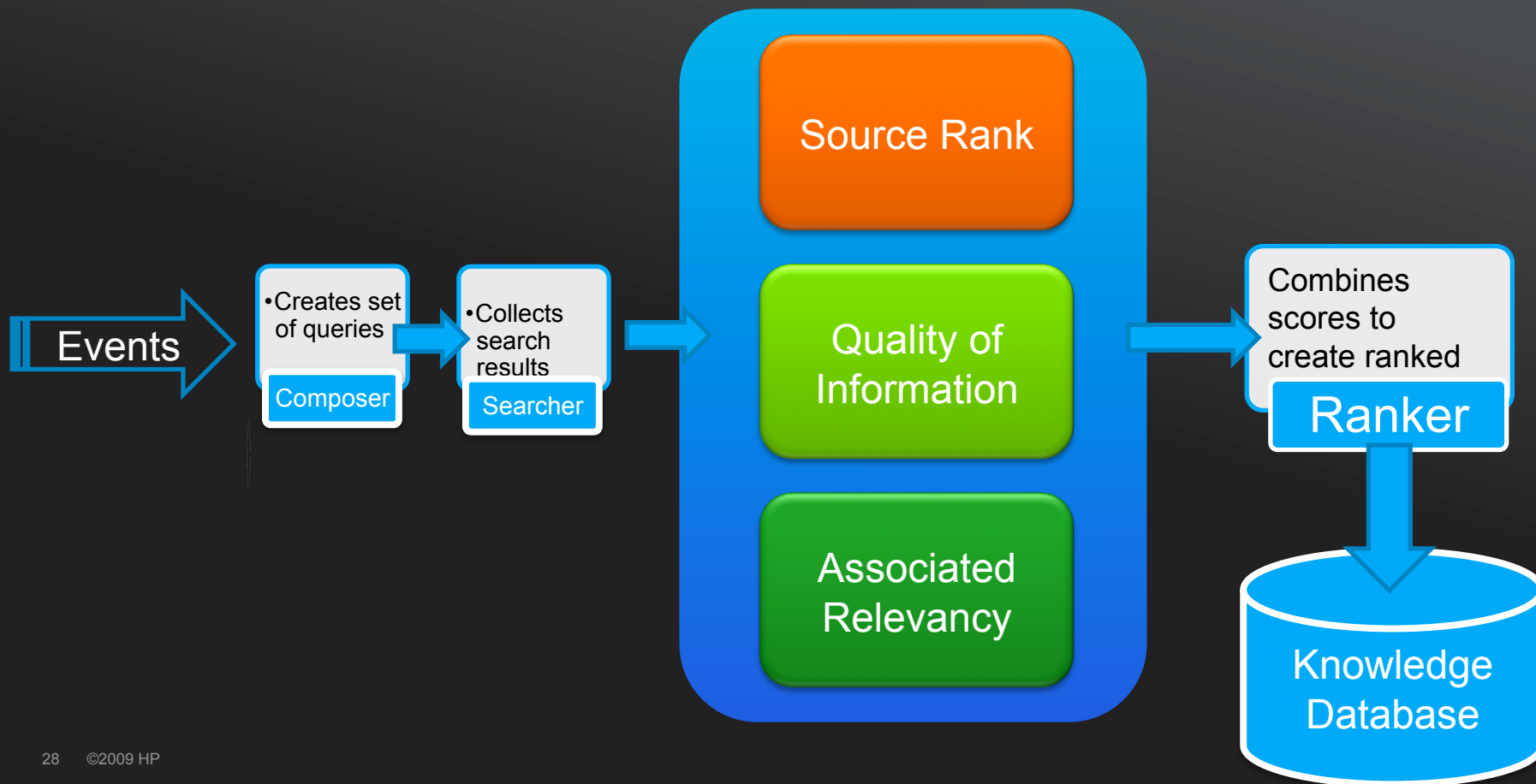
- Training on threads from one domain, testing on the other

Classif y

Train/Test	Oracle	IBM
Oracle	90%	85%
IBM	79%	97%



KNOWLEDGE CREATION SYSTEM



ASSOCIATED RELEVANCY

- Compute Levenshtein Distance between event and document
- Regular search engine may not have found the event but rather a collection of the words in the search string which are not related to each other



PARIS SAMPLE RESULTS: HP ITRC FORUM

Multifunction printers

Product,
print,
scan,
printer,
multifunct,
fax,
copier

NNM

nnm, agent,
insight,
network, node,
event, ov, trap,
monitor, alert,
snmp, sim

Databases

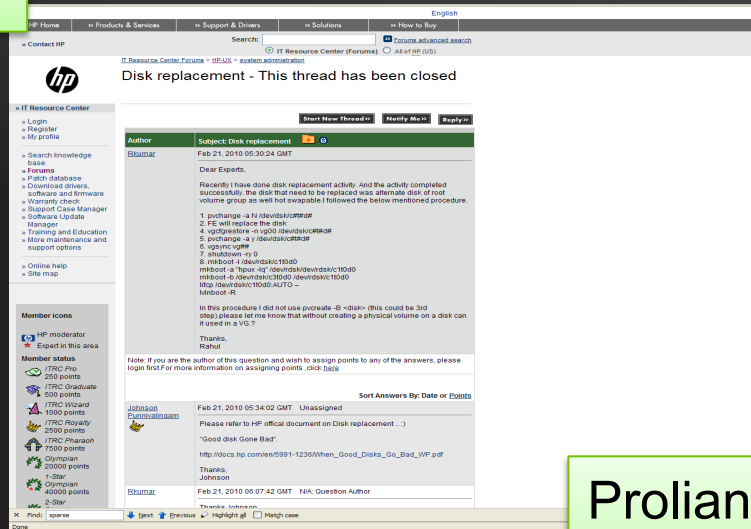
table, sql,
connect,
field,
name,
value,
record, db

hp,
hpux
ux,
unix

HPUX

Proliant
Servers

mgmt, out, remot, pack,
light, consol, pro, reset, dl,
380, liant, proliant, ilo,
firmwar, lightsout



KNOWLEDGE CREATION SYSTEM

Status & Summary

- Created a system that gathered and reranked pertinent knowledge from the web to aid in troubleshooting and understanding system events in logs.
- System slated for HP Software's BSM products
- Future work: Continue to refine feature selection and QOI measures

$$w_{QOI}QOI_i + w_{AR}AR_i$$

Combines
scores to
create ranked

Ranker

Knowledge
Database