



# Primo real-time analytics/metrics



Mehmet Celik

Senior Business  
Consultant

IGeLU 2015, Budapest, 3 September 2015

Session: 10.22 Room Q1: 15:30 – 16:00

# Who am I?

Mehmet Celik

Senior Business Consultant @ KULeuven/LIBIS

- Github:
  - <http://github.com/mehmetc>
  - <http://github.com/libis>
- Email: mehmet.celik@libis.be
- Twitter: @celikbe

# LIBIS

- Business Unit @ KULeuven since the early '70
- We provide library, museum and archives related services and consultancy see <http://www.libis.be>
- We support Aleph, Primo, Alma, Rosetta, Omeka, Collective Access, ...
- We host Aleph, Rosetta, Omeka, CA, ...
- We are a Service Provider are hosted for Alma and Primo but for now we still have shell access

# Service Provider

- We work closely with ExLibris
  - Need to know what is going on with the hosted environments. Status pages are not enough.
  - Need to provide quick solution to our customers
- ExLibris does extensive monitoring of all hosted environments.
  - Getting detailed information takes time. Tickets need to go up the chain.

# We need our own monitoring

- Beginning this year we had a problem
  - Primo was slow, really slow
  - Finding the cause took a long time
  - It is not in our nature to sit back and relax
- Monitoring does not need a lot of data
  - ExLibris helped a lot

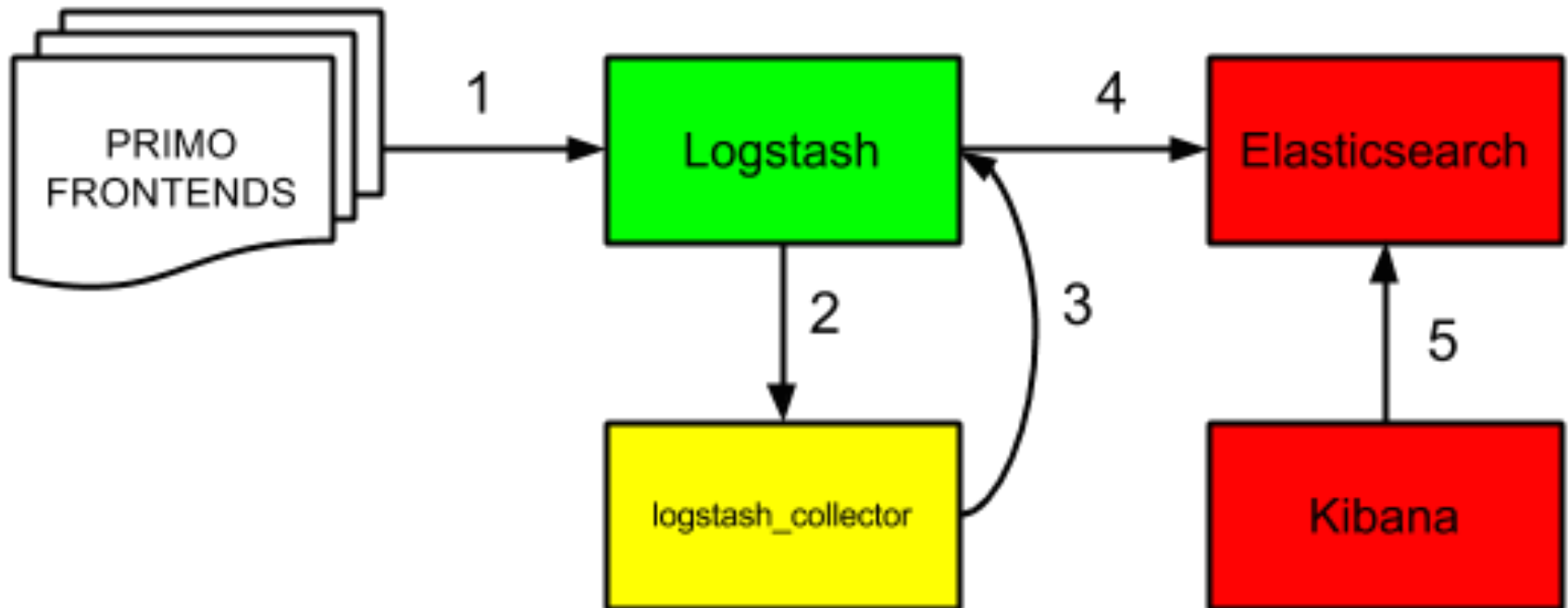
## The idea

- Primo is written in Java and uses a lot of standard libraries
- Logging is taken care by -> Log4J
- Log4J can write to files but also to a network socket
  - We redirect parts of the logs to a locally hosted analytics environment.

# Tools

- Primo jaguar.log
  - [E]lastic search
  - [L]ogstash
  - [K]ibana
  - Collector (CEP)
- 
- ExLibris uses the same ELK environment for their daily monitoring

# Data Flow



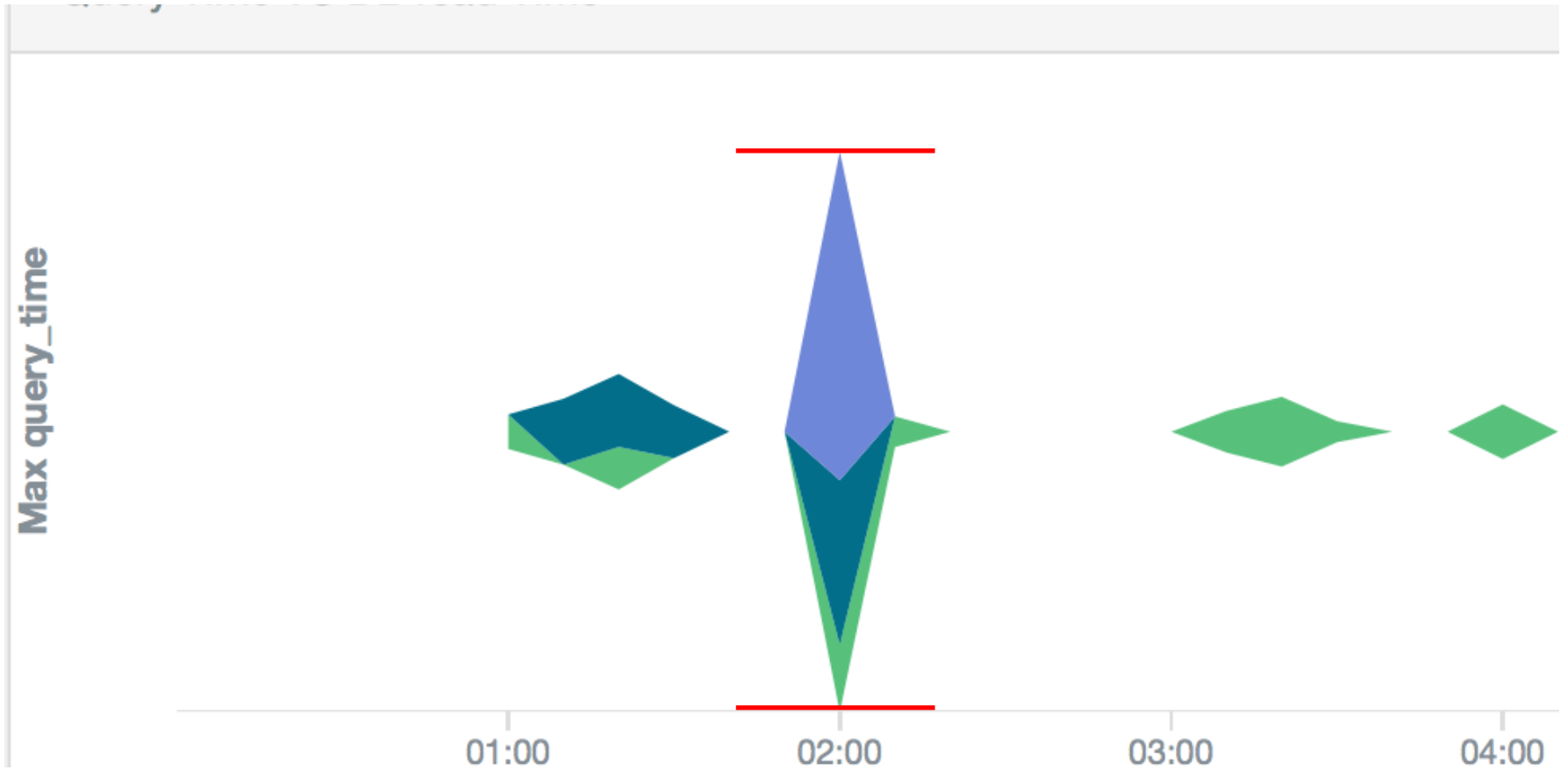


# jaguar.log

- 2015-06-09 11:05:51,990 INFO [t-RMI TCP Connection(20)-31.186.254.54] [c-JaguarParallelMSearcher] - JaguarParallelMSearcher.search() start, for query: Q: ((wood)) AND (scope:(LUCA LEMM LIB)):KUL:false:400:L::dut
- 2015-06-09 11:05:52,025 INFO [t-RMI TCP Connection(20)-31.186.254.54] [c-JaguarParallelMSearcher] - JaguarParallelMSearcher.search() start, for query: SQ: ((wood)) AND (scope:(LUCA LEMM LIB)):KUL:false:400:L::dut
- 2015-06-09 11:05:52,099 INFO [t-RMI TCP Connection(20)-31.186.254.54] [c-Searcher] - Net network search times: S: 23, S2: 51, D: 4, D2: 0, D3: 0 ,F: 2 ,DB: 5. Time in milliseconds
- 2015-06-09 11:05:52,099 INFO [t-RMI TCP Connection(20)-31.186.254.54] [c-JaguarSearchBean] - Elapsed Jaguar Search time: **116 milliseconds**
- 2015-06-09 11:05:52,077 INFO [t-RMI TCP Connection(20)-31.186.254.54] [c-Searcher] - Found 2 document(s) (in 93 milliseconds) that matched query '((wood)) AND (scope:(LUCA LEMM LIB)):KUL:false:400' [1,10]

```
{
  "_index": "logstash-2015.06.22",
  "_type": "log",
  "_id": "AU4Z0tXe10JOUfI5MV8j",
  "_score": null,
  "_source": {
    "@timestamp": "2015-06-22T05:52:00.584Z",
    "tid": "production-1-23326",
    "remote_ip": "10.24.0.103",
    "query": "( (wood) )",
    "frontend": 1,
    "system": "production",
    "query_count_local": 2167,
    "S": 48,
    "S2": 0,
    "D": 72,
    "D2": 0,
    "D3": 0,
    "F": 17,
    "DB": 28,
    "query_time": 227,
    "view": "PRIMO",
    "logged_in": "false",
    "language": "eng",
    "scope": "(scope:(32LIBIS_ALMA_DS))",
    "@version": "1",
    "type": "log"
  }
}
```

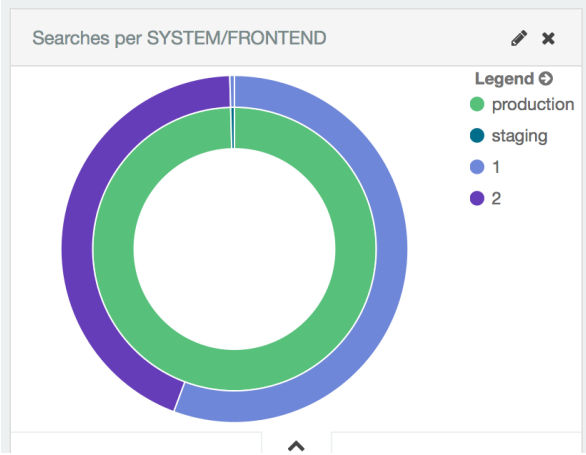
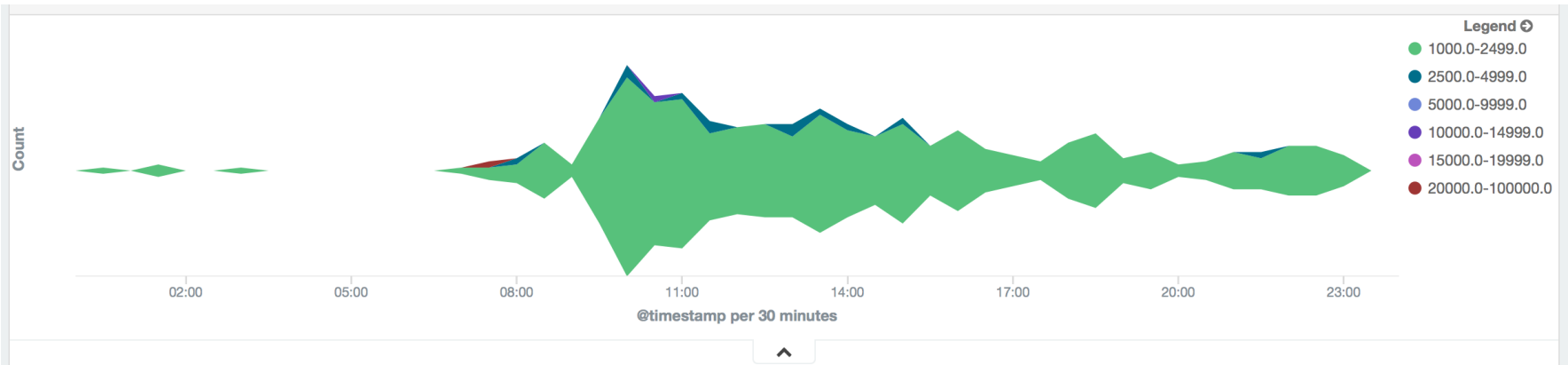
# Visualizing the



# What we noticed

Time ▼	query	query_time	DB	D	S
▶ March 8th 2015, 06:02:06.719	((trends in cognitive sciences))	12347	1876	15	9392
▶ March 8th 2015, 03:05:14.035	((women's emotional adjustment to IVF a systematic review of 25 years of research.))	16179	0	2098	2787
▶ March 8th 2015, 02:08:03.864	((lev manovich))	10509	2640	245	7115
▶ March 8th 2015, 02:06:23.129	((new media a critical introduction))	20998	10405	19	10000

# Demo



Top slow queries

Q	Top 5 query.raw	Max query_time
KUL	((coffee) AND facet_tlevel:(\"online_resources\") AND facet_lang:(\"eng\") AND facet_title:(\"American Ethnologist\"))	23607
KUL	((CollectionB39))	13875
KUL	((is there a difference in performance by the legal status of))	3676
KUL	((Stereotact Funct Neurosurg))	2995
KUL	((NJF))	2940
KADOC	((Het arrest Damseaux Hof van Justitie grijpt de herkansing niet))	3176
KADOC	((Vlaanderen))	1088

## But ...

- You need shell access to harvest and create these dashboards

## The alternative ... jQuery.PRIMO

- Client side convenience [library](#) for Primo
  - Free
  - MIT licenced
- Just a couple of lines to log searches as events in ELK
- [Gist](#) with configuration files

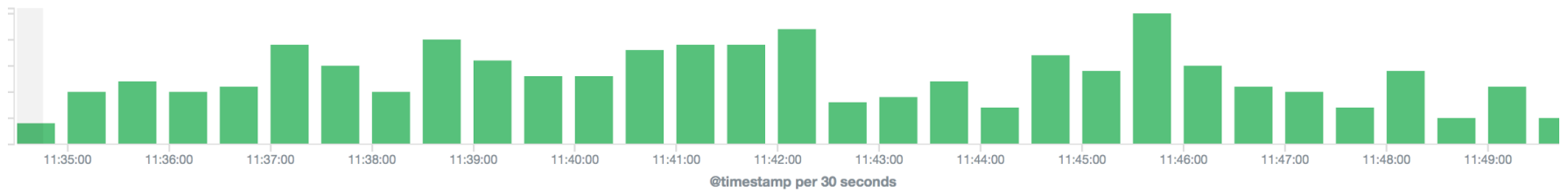
# What is logged?

```
{
  session_id:      jQuery.PRIMO.session.sessionId,
  ip:              jQuery.PRIMO.session.ip.address,
  fe:              jQuery.PRIMO.session.view.frontEndID,
  loggedin:        jQuery.PRIMO.session.user.isLoggedIn(),
  oncampus:        jQuery.PRIMO.session.user.isOnCampus(),
  count:           jQuery.PRIMO.query.count,
  scope:           jQuery.PRIMO.query.scope,
  sorted_by:       jQuery.PRIMO.query.sorted_by,
  step:            jQuery.PRIMO.query.step,
  page:            jQuery.PRIMO.query.page,
  tab:             jQuery.PRIMO.query.tab,
  search_type:     jQuery.PRIMO.query.type,
  text_query:      '',
  query:           {},
  facets:          {},
  view:            jQuery.PRIMO.session.view.code,
  ui_lang:         jQuery.PRIMO.session.view.interfaceLanguage,
  full_display:    jQuery.PRIMO.session.view.isFullDisplay(),
  ranking_categories: jQuery.PRIMO.session.user.ranking.categories,
  ranking_prefer_new: jQuery.PRIMO.session.user.ranking.prefer_new,
  deeplink:        jQuery.PRIMO.query.isDeeplinkSearch()
}
```



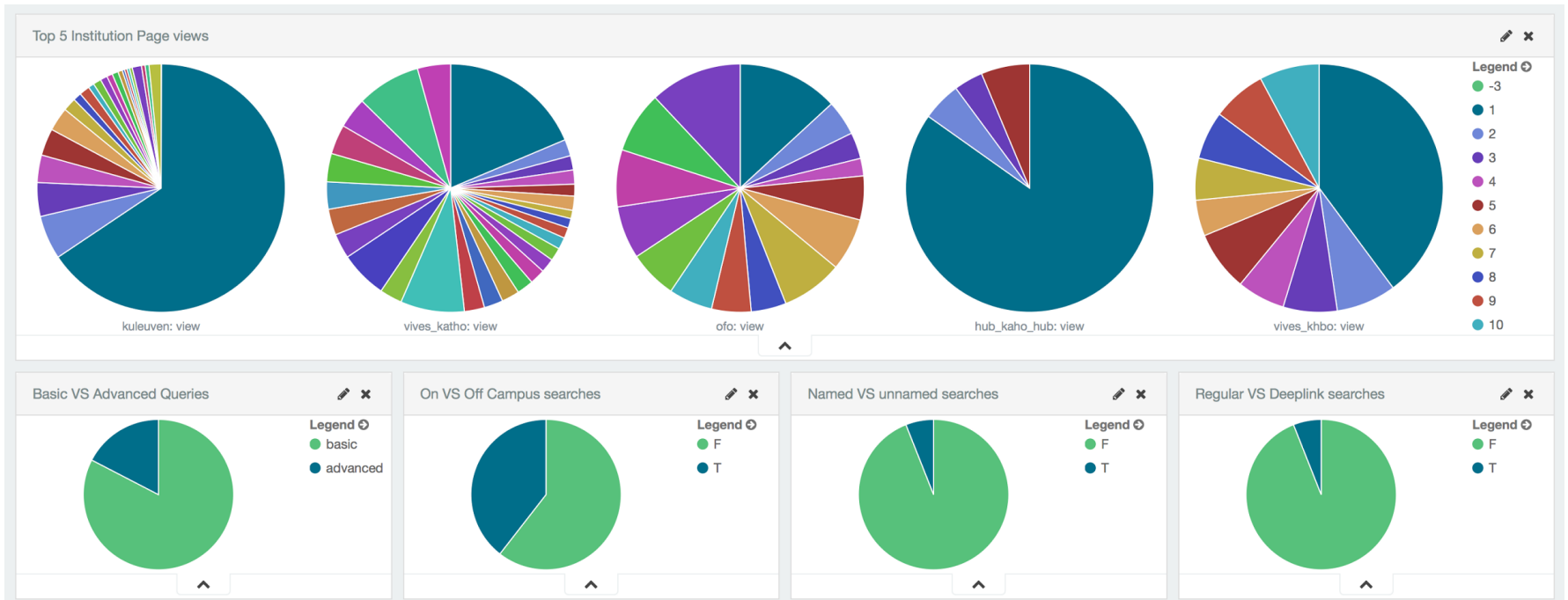
# Raw data in Kibana

September 2nd 2015, 11:34:45.424 - September 2nd 2015, 11:49:45.424 — [by 30 seconds](#)



	count	fe	full_display	deeplink	loggedin	oncampus	page	ui_lang	search_type	text_query
	1	1	false	true	false	false	1	n1_BE	-	any contains italo calvino
uven	68	1	false	false	false	false	1	en_US	Basic	(any contains Wit and Greece's economic crisis: I
uven	16	1	false	false	true	false	1	en_US	Basic	(any contains Anthony+van+Dyck:+An+Afterword+SUSA
S_KHBO	218	1	false	false	true	false	5	n1_BE	Advanced	(any contains Voedings- en Dieetkunde) AND (publi
uven	79,753	1	false	false	true	false	1	en_US	Basic	(any contains education trends) AND (facet_search
uven	79,380	1	false	false	false	true	1	en_US	Basic	(any contains comparative+genomics)
uven	100	1	false	false	false	false	1	en_US	Basic	(any contains blind+in+gaza)
uven	16	1	false	false	true	true	1	en_US	Advanced	(title contains activated+sludge) AND (creator co
uven	237,557	1	false	false	true	false	1	en_US	Basic	(any contains education+trends)
uven	3	1	false	true	false	true	1	en_US	-	any contains tijdschrift voor geneeskunde

# Demo



## Conclusion

- ExLibris helped us a lot.
- Data can be gather from multiple environments -> different context = different focus.
- jQuery.PRIMO lets you create any data structure to log.
- Setting up ELK and sending data to it is easy.



Thank you !  
Questions ?

**LIBIS**

into-info@libis.be – www.libis.be

+32 (0) 16 32 22 66

De Croylan 54 – PB 5592 – B-3001 Heverlee



Information.

Innovation.

Inspiration.