Introduction

Organization

- **Data Stream Management Systems and Complex Event Processing**
  - João Moura Pires
  - 1/2 of the lectures (6 weeks)

- **Distributed Stream Processing Systems**
  - Nuno Preguiça
  - 1/2 of the lectures (6 weeks)

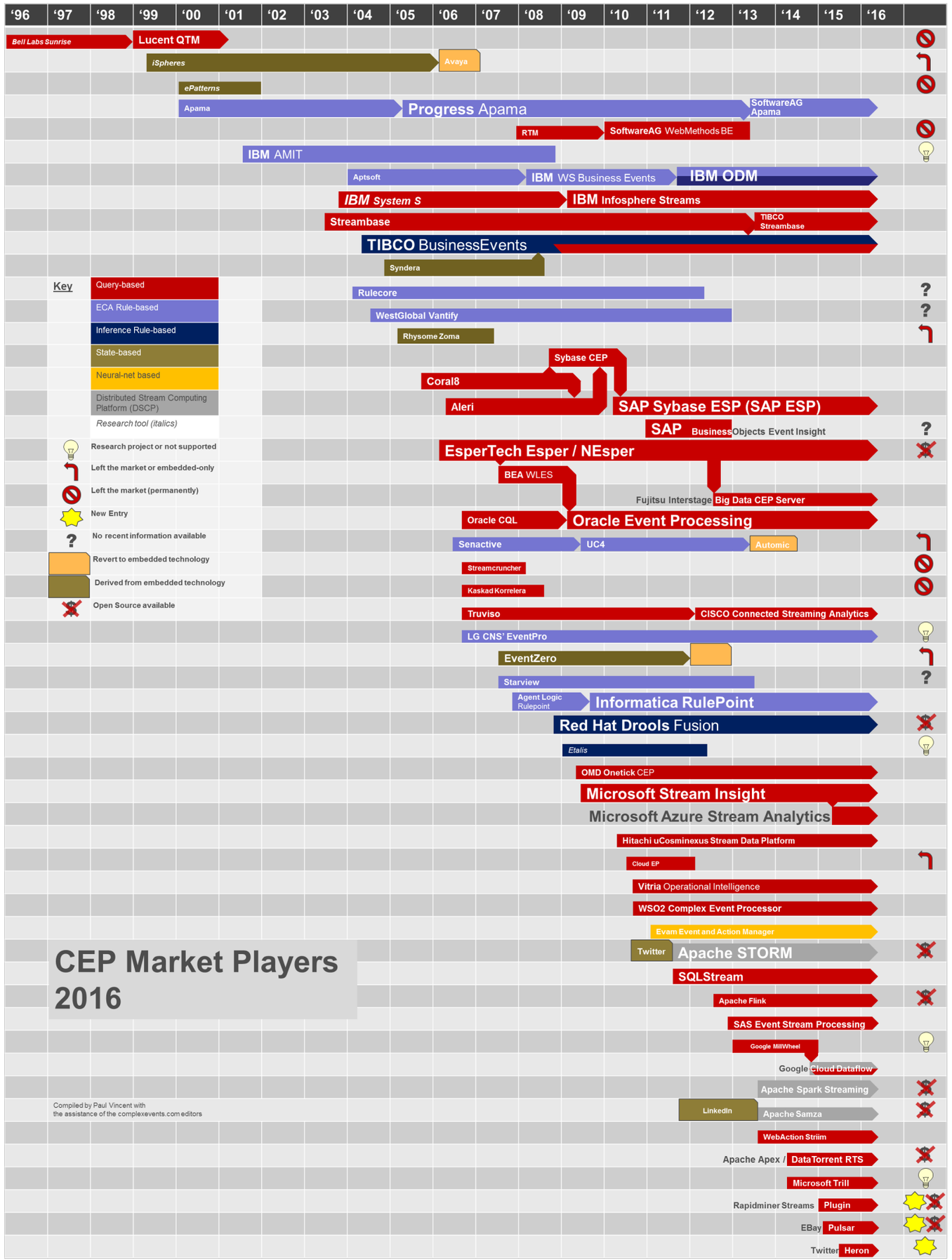
Goal

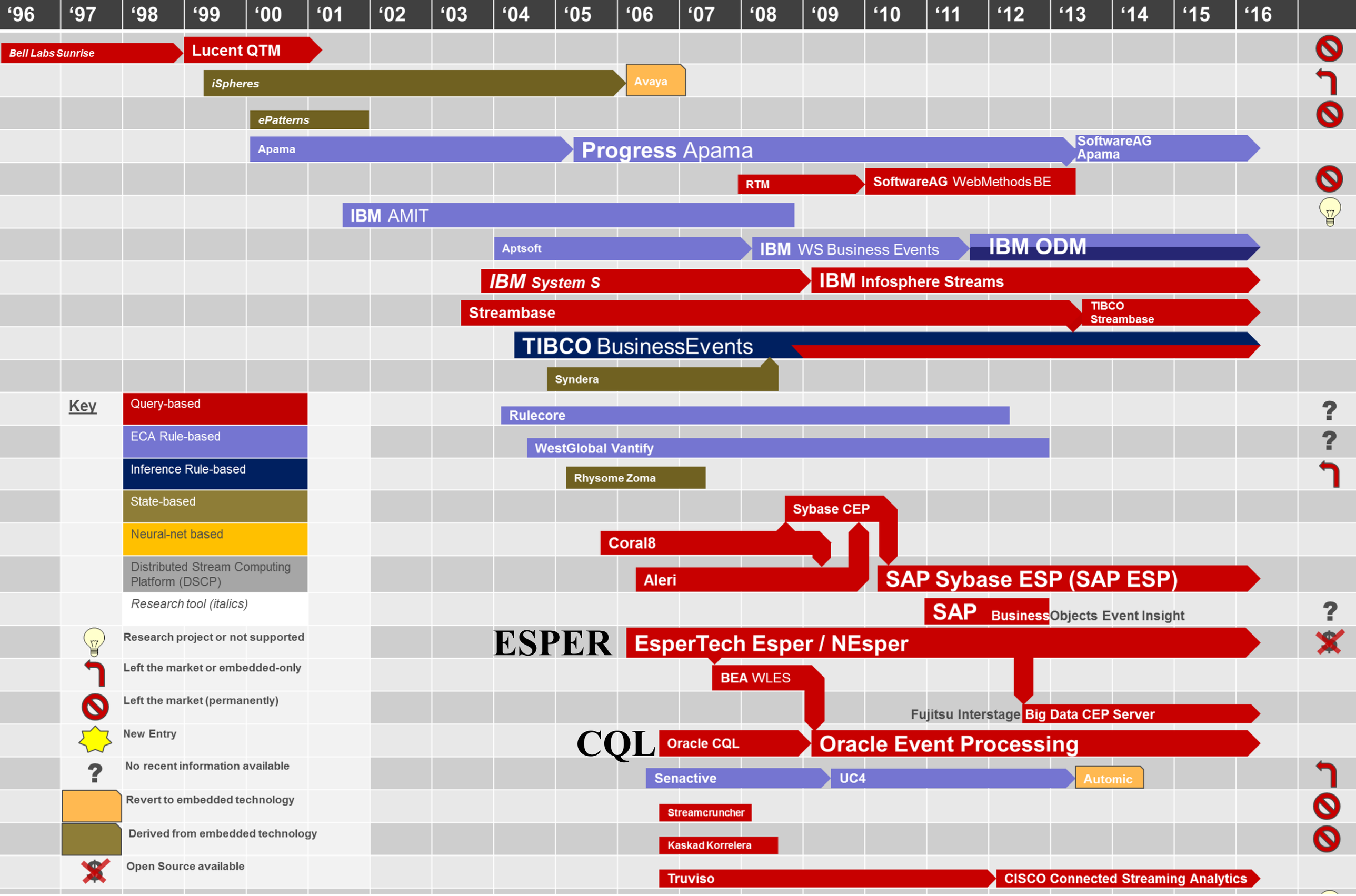
- ***Learn the fundamentals, languages and systems for building applications that process streams of data***
  - *ranging from general purpose distributed realtime stream processing systems*
  - *…*
  - *structured data models for dealing with streams*.

Syllabus

- **Data Stream Management Systems (DSMS).**
  - Structured Data Models for Streams. Algebraic operators on stream and relations
  - Continuous query languages (extensions to SQL and database management systems to deal with data streams).
- **Complex Event Processing.**
  - Streams as sequences of events.
  - Production rules, reactive rules, and event-driven computing.
  - Event processing networks, agents and channels.
  - Complex and derived events. Detection of event patterns.
  - Event-processing languages and systems.
- **Distributed Stream Processing Systems.**
  - Large Scale Distributed Data Processing: MapReduce/Hadoop, Apache Pig, Apache Spark
  - Distributed Stream Processing: Spark Streaming, Apache Storm, …
  - Iterative Distributed Data Processing: Microsoft Naiad, ...
  - Structured Data Distributed Processing: Hive, SparkSQL, ...

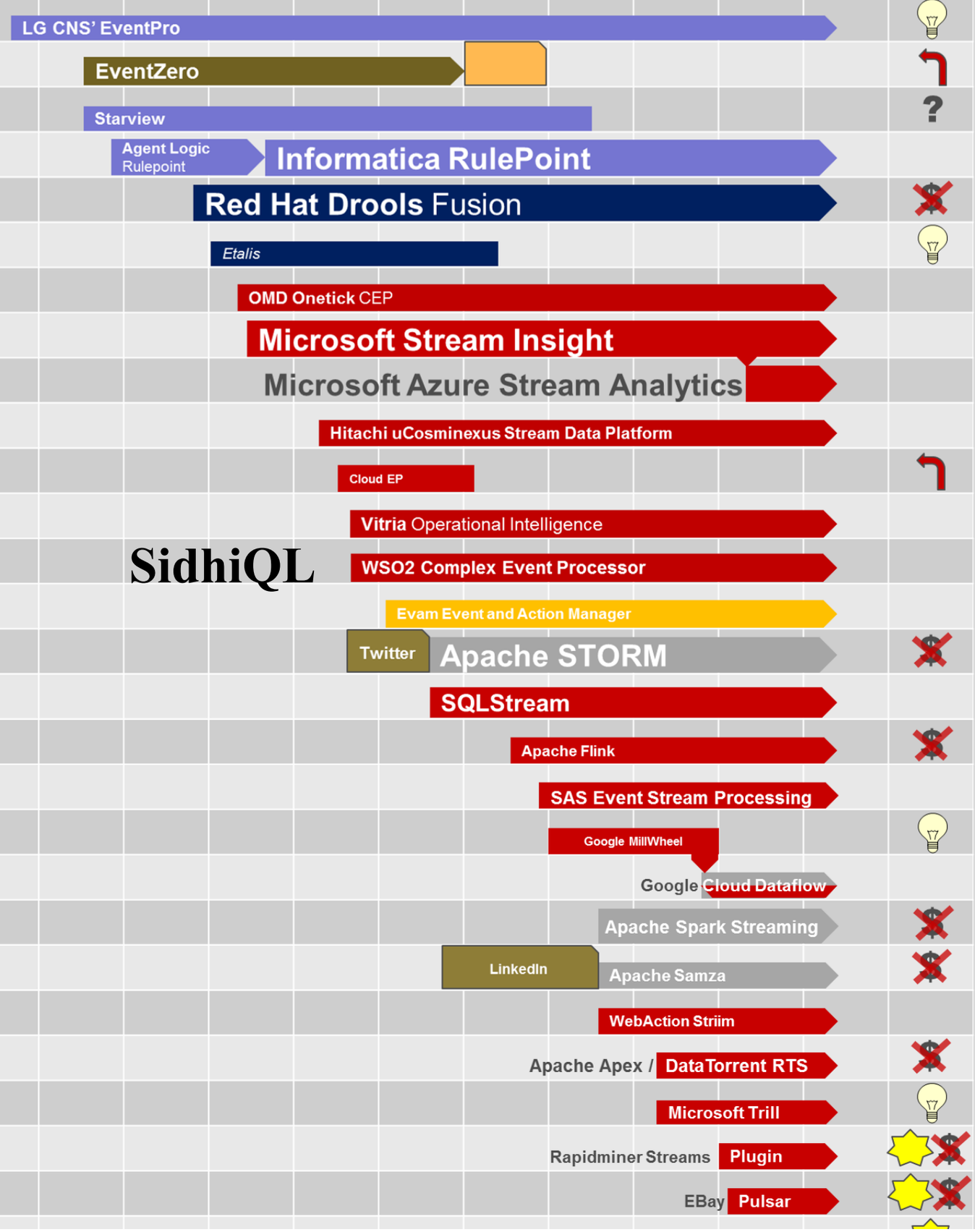
http://www.complexevents.com/2016/05/12/cep-tooling-market-survey-2016/




'96
'97
'98
'99
'00
'01
'02
'03
'04
'05
'06
'07
'08
'09
'10
'11
'12
'13
'14
'15
'16
Bell Labs Sunrise
Lucent QTM
iSpheres
Avaya
ePatterns
Apama
Progress Apama
SoftwareAG Apama
RTM
SoftwareAG WebMethods BE
IBM AMIT
Aptsoft
IBM WS Business Events
IBM ODM
IBM System S
IBM Infosphere Streams
Streambase
TIBCO Streambase
TIBCO BusinessEvents
Syndera
Rulecore
WestGlobal Vantify
Rhysome Zoma
Sybase CEP
Coral8
Aleri
SAP Sybase ESP (SAP ESP)
SAP BusinessObjects Event Insight
ESPER
EsperTech Esper / NEsper
BEA WLES
Fujitsu Interstage Big Data CEP Server
CQL
Oracle CQL
Oracle Event Processing
Senactive
UC4
Automic
Streamcruncher
Kaskad Korrelera
Truviso
CISCO Connected Streaming Analytics
Key
Query-based
ECA Rule-based
Inference Rule-based
State-based
Neural-net based
Distributed Stream Computing Platform (DSCP)
Research tool (italics)
Research project or not supported
Left the market or embedded-only
Left the market (permanently)
New Entry
No recent information available
Revert to embedded technology
Derived from embedded technology
Open Source available

CEP Market Players 2016

Compiled by Paul Vincent with
the assistance of the complexevents.com editors



Assessment

- **2 midterms [50%+50%]*66%**
  - Test 1: April, 27th (Saturday at 9:00)
  - Test 2: June, 12th (Wednesday at 16:00)
- **2 projects [50%+50%]*34%**
  - Mainly groups of 2 students
  - Project 1: up to May 3th.
  - Project 2: up to May 30th.
  - Both projects are based on the same data
- **Requirements to succeed**
  - Each test ≥ 8.5; Average of tests ≥ 9.5; Average of projects ≥ 9.5

Planning

			Lectures	Lab	Other information
1	4-Mar	8-Mar			
2	11-Mar	15-Mar	Introduction to PStr. IFP Overview		
3	18-Mar	22-Mar	IFP Framework	Setting the WS02	
4	25-Mar	29-Mar	CEP	SidhiQL	Team definition
5	1-Apr	5-Apr	CQL	SidhiQL	
6	8-Apr	12-Apr	Project presentation and discussion	Student Support	
7	15-Apr	19-Apr	Stream Reasoning	Student Support	
8	22-Apr	26-Apr			T1: April 27th
9	29-Apr	3-May	Intro to big data frameworks	Spark streaming	TP1: May 1th
10	6-May	10-May	Non-structured programming	Spark streaming SQL	
11	13-May	17-May	Structured programming and SQL	Kafka	
12	20-May	24-May	Continuous streaming	Student support	
13	27-May	31-May	Stream processing ecosystem	Student support	TP2: May 30th
14	3-Jun	7-Jun	Storage for streamable data		
15	10-Jun	14-Jun	Revision		T2: June 12th

Bibliography

- **Processing Flows of Information: From Data Stream to Complex Event Processing**, GIANPAOLO CUGOLA and ALESSANDRO MARGARA, Politecnico di Milano, ACM Computing Surveys, Vol. 44, No. 3, Article 15, Publication date: June 2012
- **Event Processing in Action**, OPHER ETZION PETER NIBLETT, 2011, Manning Publications Co
- **Data Stream Management,** Lukasz Golab and Tamer Özsu. Morgan and Claypool, 2010.
- **Papers that will be provide during the semester**