

# \* Design of CQA Systems for Flexible and Scalable Deployment and Evaluation

*Ivan Srba*, Maria Bielikova

Faculty of Informatics and Information Technologies  
Slovak University of Technology in Bratislava

# \* Community Question Answering (CQA)

- Communities of millions of users share their knowledge
  - by providing answers
  - on questions asked by the rest of the community



- CQA systems have been adapted into additional contexts and environments
  - Educational domain
  - Crowd-based customer services, integrated development environments (IDEs)
- Initial research how specifics of these new environments affect
  - Essential features (e.g. core question answering functions)
  - Collaboration support (e.g. question recommendation)

# \* Open Problems

## 1. Low adaptability of essential features to various settings

- Possibility to be deployed in several different instances at the same time (e.g. in several educational or enterprise organizations)
- How to make design of essential features flexible to handle various settings?

## 2. Ineffective integration and evaluation of collaboration support methods

- Possibility to perform live experiments (very rare in standard CQA systems)
- How to achieve loosely coupled integration of collaboration support methods with CQA systems?
- How to make combination of offline and online experiments as effective as possible?

# \* Case Study on Our CQA System Askalot

- A novel concept of an organization-wide educational CQA system

The screenshot shows the Askalot CQA system interface. At the top, there is a navigation bar with the following items: ASKALOT, Questions, Groups, Categories, Tags, Users, Activity, Statistics, Help, a help icon, a notification icon with '9', the user name 'Andrew', and a settings icon. Below the navigation bar is a green button labeled '+ Ask a question'. To the right of this button is a search bar containing the text 'e.g. midterm exam, ...' and a 'Search' button. Further right is a dropdown menu showing 'bachelor-1st'. Below the search bar are several tabs: 'Recent (2)', 'Unanswered (0)', 'Answered (1)', 'Solved (1)', and 'Favored (2)'. To the right of these tabs is a refresh icon and the text 'Refreshed less than a minute ago'. The main content area displays two questions. The first question is titled '? Anonymous class in always true condition' by user 'Ruby' (profile picture of a red cross) and was asked on '31. March 2016'. It has 0 votes, 1 answer, and 8 views. The question text is: 'There is an anonymous class creation in else branch of condition which will always return true. `'''java if (true) {...} else return new.....; '''` How many instanced of this class can be created? How can I implement a singleton in Java...'. The second question is titled '? Singleton in Java' by user 'Ivan' (profile picture of a man) and was asked on '23. March 2016'. It has 1 vote, 1 answer, and 8 views. The question text is: 'How can I implement a singleton in Java language? I need to secure that I will work with only one instance of a object from various threads. In addition, I need to serialize this object. What is the best possibility how to achieve it? \*[An...'. Both questions have tags: 'Object-oriented programming - Exercises' and 'anonymous-class' for the first; 'oop' for the first; 'Object-oriented programming - Lectures', 'bachelor-1st', 'java', 'lectures', 'oop', 'serialization', and 'singleton' for the second.

1150 users  
430 questions  
560 answers  
410 comments

Ruby on Rails  
Open source

# \* Case Study on Our CQA System Askalot

- The version of Askalot was proposed specifically for our university
  - ... and thus it lacked sufficient flexibility and scalability
  
- We started a cooperation with:
  1. **Harvard University** in order to transform Askalot into a plugin to MOOC system edX
  2. **University of Lugano** in order to deploy at additional universities as a part of cooperation project in the SCOPES program



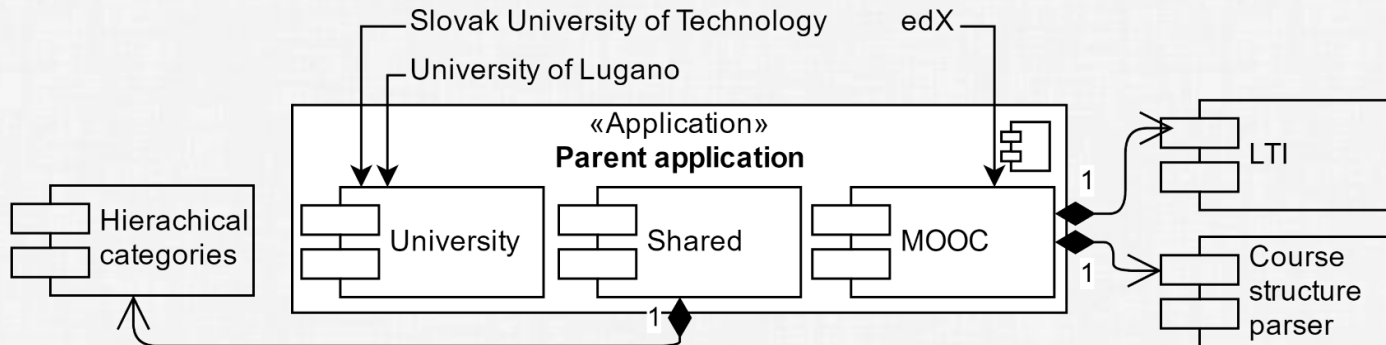
Università  
della  
Svizzera  
italiana

# \* Designing Essential Features for Various Settings

- **Modular system architecture**
- **Adaptable self-managed content organization**
- Flexible user management integration
- Ubiquitous activity awareness and notifications

# \* Modular System Architecture

- Necessity to develop two main configurations of our system
  - Askalot @university
  - Askalot @mooc
- Solution based on one application and three components (RoR engines)
  - **Shared**
    - core features that are common for both configurations
  - **University** and **MOOC**
    - inherit all features from the core component and add specialized features



# \* Adaptable Self-managed Content Organization

- Two-level topic structure which supports easy and flexible deployment
  - Category level
    - Reflects the formal structure of a university or a MOOC course
  - Tag level
    - Selected by asker to describe particular question topics
- Categories should
  - be hierarchical
  - reflect repeating sessions (i.e. academic years or course sessions)
- Solution based on nested set pattern
  - Hierarchical tree where each node has
    - **domain-specific ID** - to identify the same categories across all academic years or course sessions
    - **shareable flag** – whether questions from the previous sessions should be displayed also in the current session
    - **askable flag** – whether students can ask questions in this category



View this course as: Student

Home Course AskALot **A**

Bookmarks

Machine learning

Welcome Video

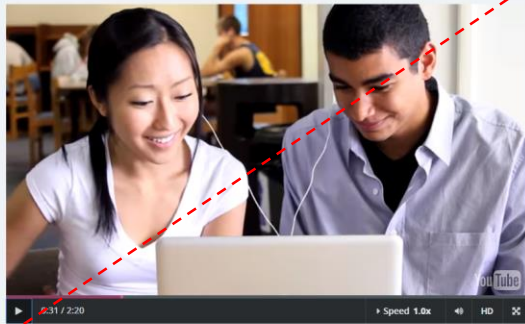
Introduction

Supervised Learning

Unsupervised Learning

Machine learning &gt; Welcome Video &gt; Welcome Video

Video



AskALot (External resource)

New question

1 favorite

0 votes

What's is the difference between clustering and classification?

[Welcome Video - Welcome Video](#)
[askalot-demo-2016](#)
[welcome-video](#)
[clustering](#)
[categorization](#)

1 answer

6 views

What is the difference between clustering and classification? \*[An example of general question]\*

Added by Sophia 25. February 2016

0 favorites

2 votes

Minimal requirements on project

[Welcome Video - Welcome Video](#)
[minimal-requirements](#)
[project](#)
[askalot-demo-2016](#)
[welcome-video](#)

1 answer

4 views

There is information in the minimal requirements on the project that it is necessary to use at least two datasets. Are these requirements valid for each algorithm or for overall project? \*[An example of question which tackles with...]

- A: Global view loaded from shared component
- B: Unit view loaded from MOOC component

## AskALot (External resource)

New question

1

favorite

0

votes

What's is the difference between clustering and classification?

[Welcome Video - Welcome Video](#)
[askalot-demo-2016](#)
[welcome-video](#)
[clustering](#)
[categorization](#)

1

answer

6

views

What is the difference between clustering and classification? \*[An example of general question]\*

Added by Sophia 25. February 2016

0

favorites

2

votes

Minimal requirements on project

[Welcome Video - Welcome Video](#)
[minimal-requirements](#)
[project](#)
[askalot-demo-2016](#)
[welcome-video](#)

1

answer

4

views

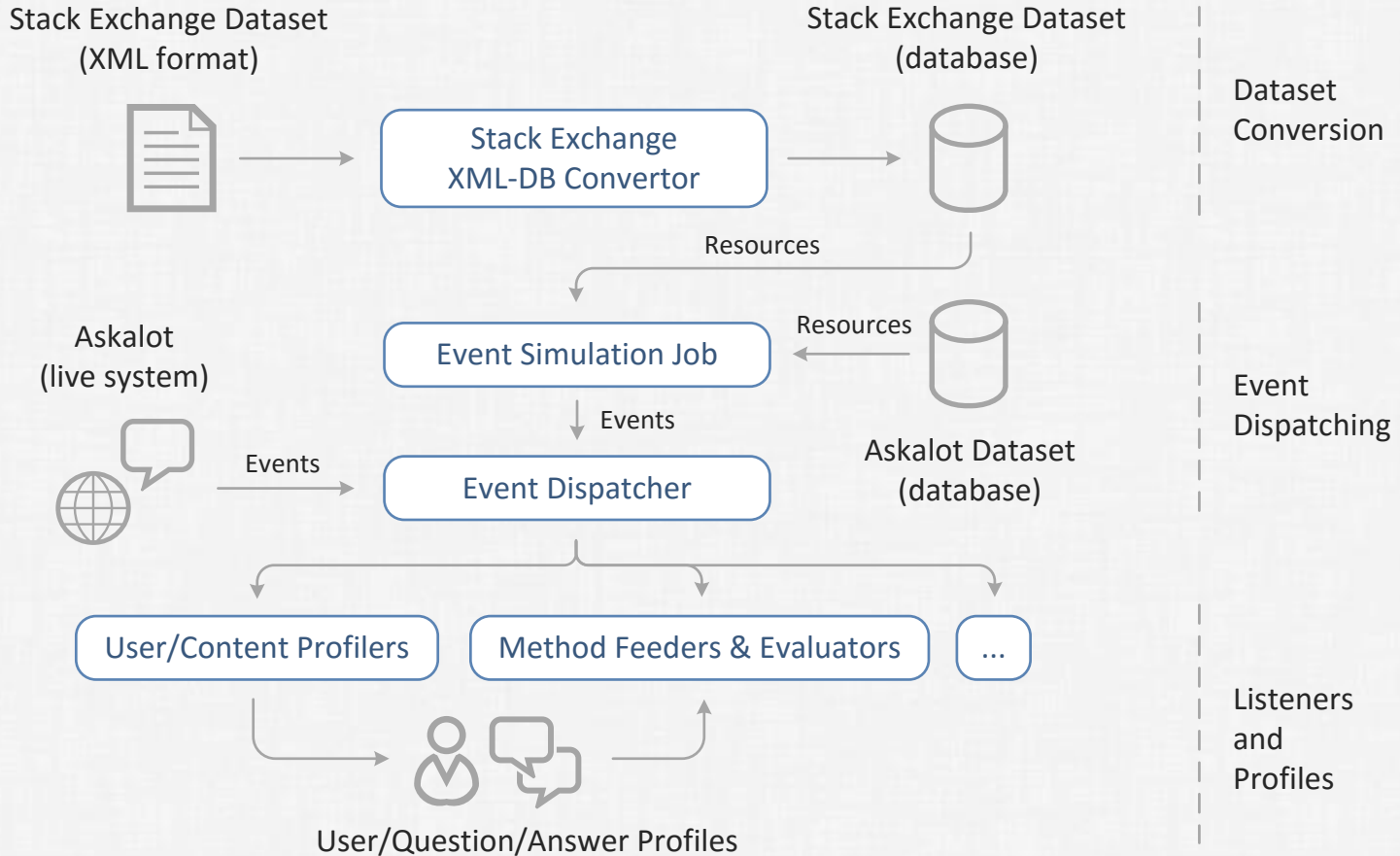
There is information in the minimal requirements on the project that it is necessary to use at least two datasets. Are these requirements valid for each algorithm or for overall project? \*[An example of question which tackles with...]

**B**

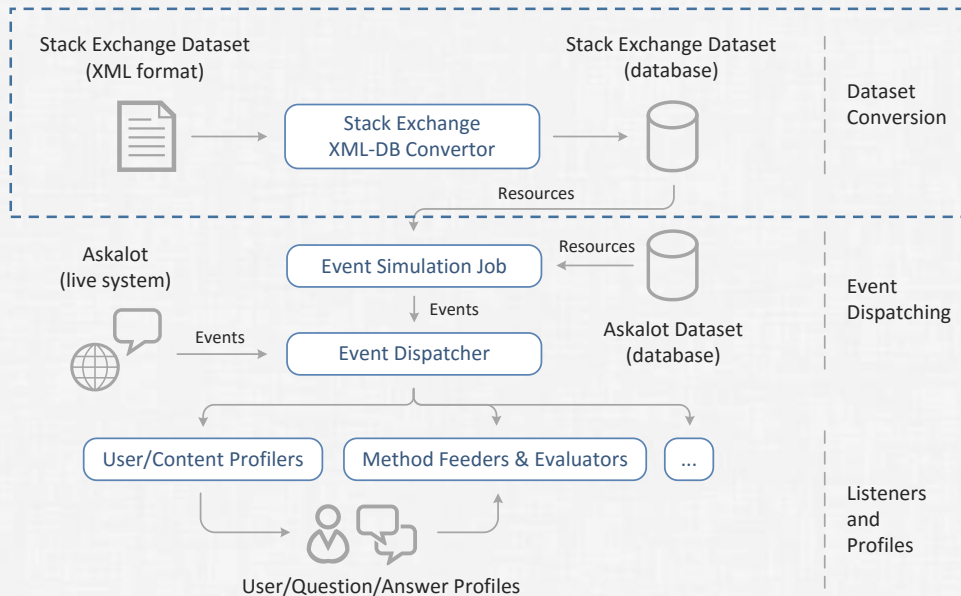
# \* Designing Universal Experimental Infrastructure

- Modular approach where all methods are loosely coupled from other methods or system itself
- Possibility to combine training/evaluation of methods on offline datasets with live experiments

# \* Three main parts of experimental infrastructure



# \* Part 1: Dataset Conversion



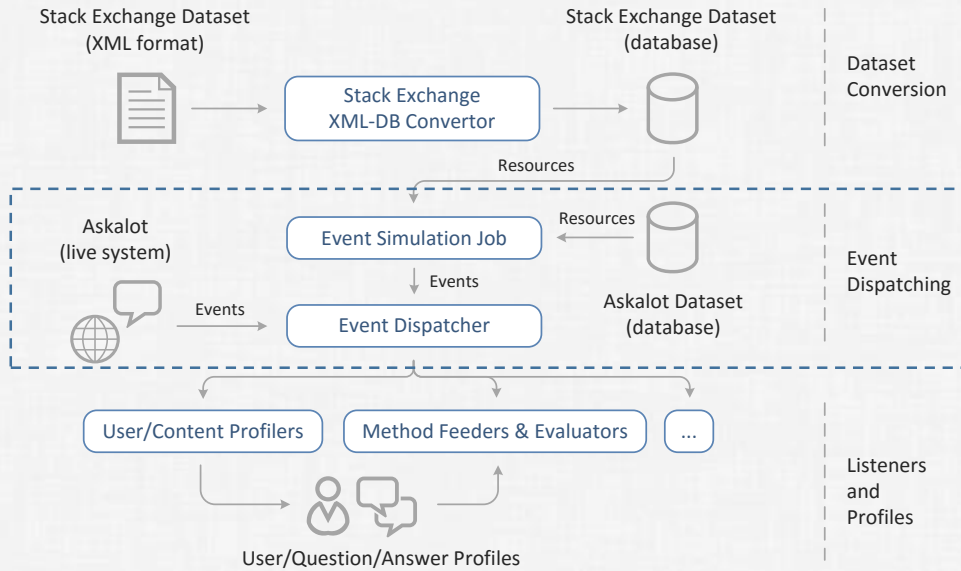
## • Utilities to convert

- any datasets from CQA systems
- to a dedicated experimental database with the same database schema as Askalot system

## • Existing convertors

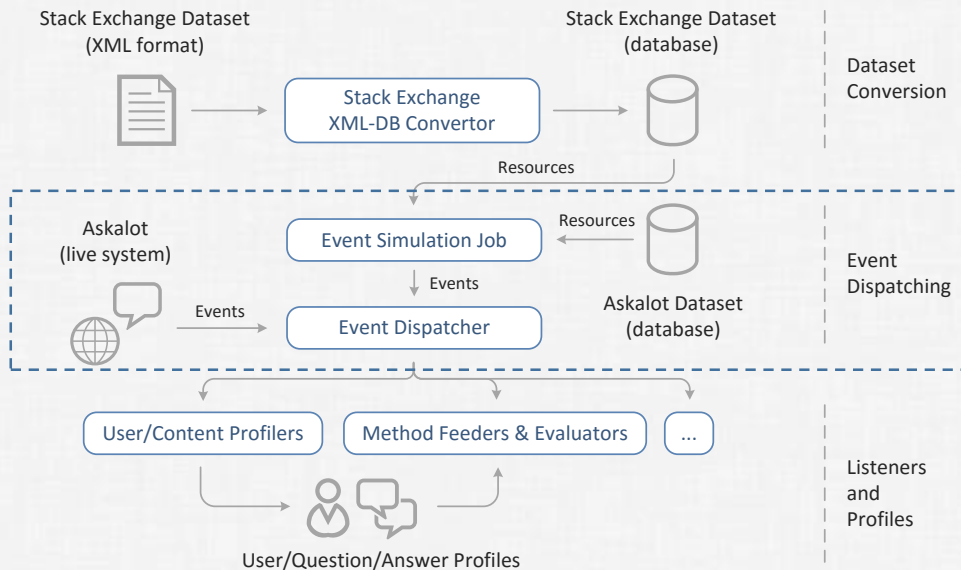
- Stack Exchange datasets
- edX datasets (in progress)

# \* Part 2: Event Dispatching



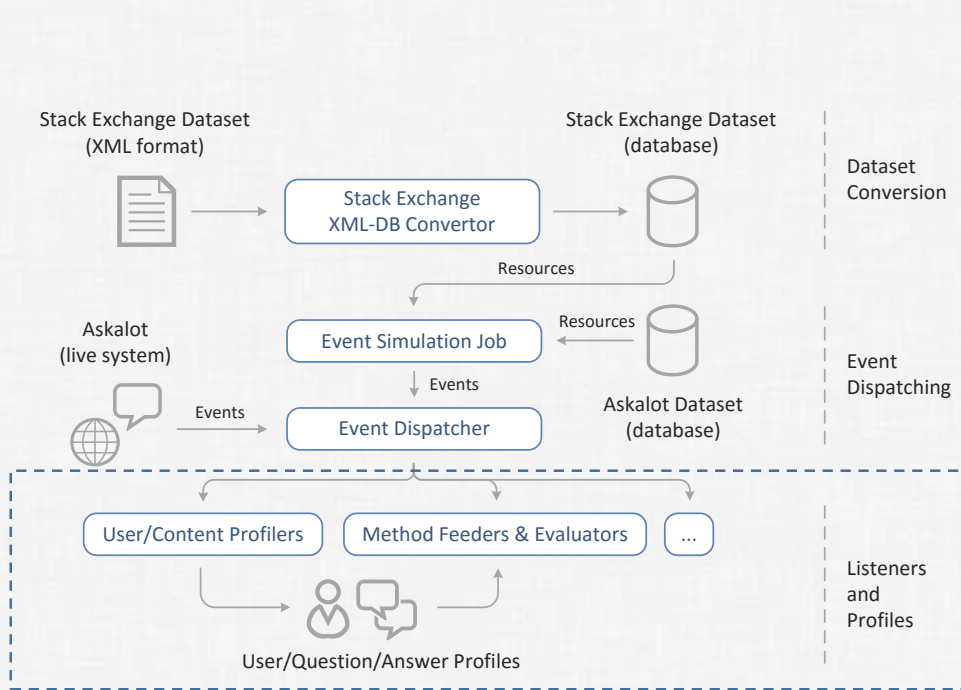
- Event is represented by
  - a resource
    - question, answer, comment, view, vote
  - an action type
    - create, update, delete
  - an initiator
    - who performed this action
- Event sources
  - live system in online experiments
  - datasets in offline experiments
    - either from Askalot itself
    - or from other CQA systems

# \* Part 2: Event Dispatching



- **Event simulation job**
  - selects from the database all resources
  - converts them to a list of events
  - sorts events by time when they originally happened
  - sets the current time in the experimental environment to this event time
  - dispatches the event
- **Exact reproduction of events**
  - as they would be created by the live system

# \* Part 3: Listeners and Profiles



## • Listeners


- Profilers - model users/content
  - user expertise, question difficulty
- Method feeders - trigger and evaluate various research methods
  - recommendation of new questions to potential answerers


## • User/question/answer profiles

- Universal data structures to store results of profilers
  - attribute
  - value
  - probability
  - source

# \* Conclusion

- Drawing upon redesign of CQA system Askalot, we proposed several design recommendations
  - how concepts of CQA systems can be adapted to an educational context and organizational environment
  - with achieving high flexibility and scalability
- As the result, Askalot can be
  - deployed in three different environments (and many more if necessary)
  - characterized also as an open platform based on the universal experimental infrastructure

 [askalot.fiit.stuba.sk/demo](https://askalot.fiit.stuba.sk/demo) (Try it!)

 [github.com/AskalotCQA/askalot](https://github.com/AskalotCQA/askalot)

 [askalot@fiit.stuba.sk](mailto:askalot@fiit.stuba.sk)