



Intelligent Information Management  
Targeted Competition Framework  
ICT-2011.4.4(d)

Project **FP7-318652 / BioASQ**

Deliverable **D2.6**

Distribution **Public**



<http://www.bioasq.org>

## **Social Network**

Norman Heino and Axel-Cyrille Ngonga Ngomo

Status: Final (Version 1.0)

October 2013

**Project**

Project ref.no.	FP7-318652
Project acronym	BioASQ
Project full title	A challenge on large-scale biomedical semantic indexing and question answering
Project site	<a href="http://www.bioasq.org">http://www.bioasq.org</a>
Project start	October 2012
Project duration	2 years
EC Project Officer	Ms Martina Eydner

**Deliverable**

Deliverable type	Report
Distribution level	Public
Deliverable Number	D2.6
Deliverable title	Social Network
Contractual date of delivery	M12 (September 2013)
Actual date of delivery	October 2013
Relevant Task(s)	WP2/Task 2.2
Partner Responsible	ULEI
Other contributors	
Number of pages	17
Author(s)	Norman Heino and Axel-Cyrille Ngonga Ngomo
Internal Reviewers	
Status & version	Final
Keywords	project templates, latex, reports

---

## Executive Summary

---

The BioASQ Social Network is designed as a platform for the experts to communicate on, publish results about and access the BioASQ data. Beyond the end of the project, it will facilitate the compilation of novel versions of the BioASQ benchmarks, as well as new datasets. During the course of the project, it will be used as a platform to present the disseminate the BioASQ questions, mirror the project activities as well as detect experts for the second batch of questions.

---

## Contents

---

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Concepts</b>	<b>3</b>
2.1	Interaction Model . . . . .	3
2.2	Objects in the BioASQ Social Network . . . . .	4
2.2.1	Questions . . . . .	4
2.2.2	Users . . . . .	4
2.2.3	Activities . . . . .	4
2.2.4	Messages . . . . .	5
2.2.5	Notifications . . . . .	5
<b>3</b>	<b>Implementation</b>	<b>6</b>
3.1	Architecture . . . . .	6
3.2	Data Schema . . . . .	7
3.3	Deployment . . . . .	8
<b>4</b>	<b>Existing Functionality</b>	<b>10</b>
4.1	Authentication . . . . .	10
4.2	Home . . . . .	10
4.3	Messages . . . . .	13
4.4	Questions . . . . .	13
4.5	Timeline . . . . .	14
4.6	User/Preferences . . . . .	15
<b>5</b>	<b>Envisioned Extensions</b>	<b>17</b>

---

## List of Figures

---

1.1	User story for the BioASQ Social Network . . . . .	2
2.1	An expanded question as displayed in BioASQ Social Network . . . . .	4
2.2	Various activities as displayed by BioASQ Social Network. . . . .	5
3.1	Architecture of the BioASQ Social Network . . . . .	7
3.2	Data schema of the Activity, Message, and User entities. . . . .	8
4.1	Sign in . . . . .	11
4.2	Registration and password reset. . . . .	11
4.3	Home after registration (top) and with activities (bottom). . . . .	12
4.4	Messages tab . . . . .	13
4.5	Question list . . . . .	14
4.6	User . . . . .	15
4.7	Preferences . . . . .	16

---

## Introduction

---

In efforts such as BioASQ it has often been difficult to ensure sustainability after the end of the project, leading to 1) benchmarks not being updated to novel trends in technologies as well as 2) systems being overtrained on static benchmarks but not performing well in real use. This phenomenon can be clearly seen in several domains of research such as Named Entity Recognition, where old benchmarks are still being used to train systems. In the domain of question answering, novel efforts that generate up-to-date benchmarks (e. g., QALD) are most commonly driven by small teams composed of committed scientists. Yet, when dealing with specialized content such as bio-medical data (as in the case of BioASQ), large teams of experts need to be coordinated. Moreover, the results of their efforts need to be coordinated and compiled effectively. Finally, the existing resources need to be reused effectively.

To address these sustainability requirements, we have developed the BioASQ Social Network. The main idea behind this tool is to serve as a platform for the experts to communicate on, publish results about and access the BioASQ data. Beyond the end of the project, it will facilitate the compilation of novel versions of the BioASQ benchmarks, as well as new datasets. During the course of the project, it will be used as a platform to disseminate the BioASQ questions, mirror the project activities as well as detect experts for the second batch of questions. In more detail, the BioASQ Social Network is targeted towards the following requirements from BioASQ's Description of Work:

- R1. Easy dissemination of questions
- R2. Platform for the experts to communicate on, publish results about and access the BioASQ data
- R3. Mirror BioASQ activities into the platform at an early stage of the project.
- R4. Broadcast the latest results all around BioASQ's core topics
- R5. Enable experts to work on the BioASQ resources
- R6. Track changes to the dataset and facilitate the compilation of novel versions of BioASQ benchmarks

These requirements are described in the following user story (see Figure 1.1: The creation of a new benchmark (see as the first version of BioASQ benchmark) is both the starting point and the aim of

the process. To this end, questions of different types need to be created, cleaned, post-processed and finally compiled to and published as a benchmark. The BioASQ Annotation Tool and the corresponding backend scripts support this functionality. Now, a benchmark created with the BioASQ Annotation Tool can be loaded directly into the social network (thus fulfilling the requirements R1). These questions can be accessed by the users and build the basic set of resources upon which the users of the network can communicate on (see next section for technical details). Moreover, the users can publish results (e. g., through messages and posts) about the questions or the whole benchmark. By these means, we fulfill requirement R2. The mirroring of the BioASQ (or future project) activities is solved by the administrators of the project by creating a user account through which they create posts about the project, the current benchmark creation process, etc. (R3, R4). One of the main goals of the BioASQ Social Network is to enable the identification of first-class experts for the creation of new benchmark sets. Here, we opted to use the tracking abilities provided by the BioASQ Social Network (R6) to detect active users, which can be invited to the BioASQ Annotation Tool to create or curate questions (R5).

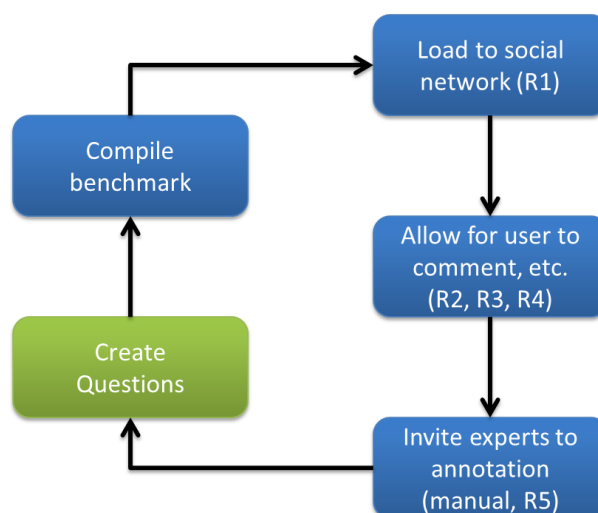


Figure 1.1: User story for the BioASQ Social Network. The steps on the right occur in the BioASQ Social Network while those on the left are carried out in the BioASQ Annotation Tool.

In the following, we present the concepts and implementation underlying the BioASQ Social Network. We begin by presenting the basic concepts underlying our implementation. Thereafter, we present implementation details such as the frameworks used and the schema used in the backend of the tool. The existing functionality is then described extensively. After a presentation of possible extensions of the BioASQ Social Network, we conclude with a summary of this report.

---

## Concepts

---

This chapter gives a short overview of interaction models in existing social networks and describes abstract concepts employed in the BioASQ Social Network.

### 2.1 Interaction Model

Existing social networks have established different models of interaction between users.

**Synchronous interaction.** In such an interaction model users need to mutually agree on a relationship before it is sanctioned by the system. Establishing such a relationship is usually called “friending” since it is often used in social networks that model personal relationships. Popular examples for social networks based on synchronous interaction are Facebook and LinkedIn.

**Asynchronous interaction.** An asynchronous interaction model does not require mutual consent. Instead the expressed will of one user to interact with another user is enough for the system to establish the relationship. Social networks adopting the asynchronous model call the creation of such a relationship “following”, “subscribing”, or “starring”. The asynchronicity of the relationship is also reflected in the notifications generated by the system. The user who enacted the relationship (called *follower*) receives updates and information of the user he chose to interact with (called *followee*).

The most popular example of an asynchronous social network is Twitter.

Asynchronous relationships are not limited to person-to-person interaction. It is also possible to “follow” other entities such as video channels (YouTube), projects (GitHub) or online discussions.

Due to the need to interact with both questions and people the model adopted for the BioASQ Social Network is an *asynchronous model*.



## 2.2 Objects in the BioASQ Social Network

### 2.2.1 Questions

In the BioASQ Social Network questions are the most important items. They are subject to discussion and evaluation by the networks participants with the ultimate goal of achieving a consensus on which questions to include in a given challenge.

Experts using the BioASQ Annotation Tool formulate questions. In regular intervals or upon user initiation they can be imported from the Annotation Tool into the BioASQ Social Network. Direct access of the Annotation Tool's question database is not possible because the BioASQ Social Network maintains additional metadata that need to be created upon import. Editing questions or modifying them in any way other than metadata is also not supported by the BioASQ Social Network. The idea here is that the BioASQ Social Network serves as a platform where all can discuss about the questions while the annotation tool is reserved for a dedicated group of people who can make use of the users' suggestions to create a new benchmark dataset.

Questions are displayed as depicted in [Figure 2.1](#). The left-hand side is always visible while the answer on the right-hand side is collapsed by default and can be expanded by clicking on the question title.

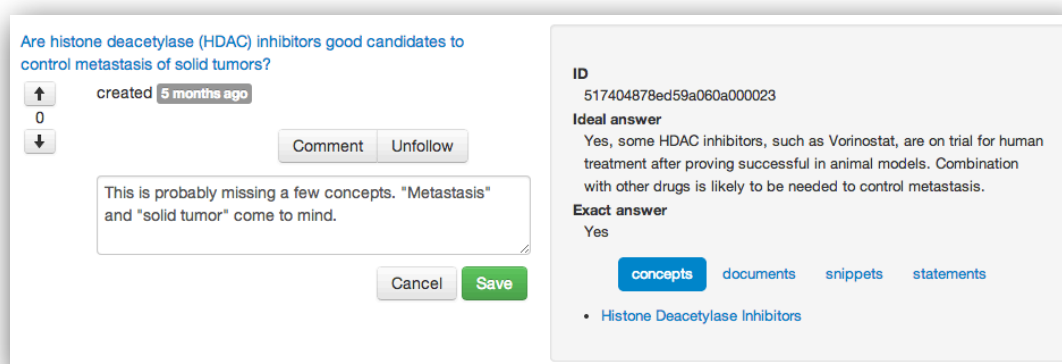


Figure 2.1: An expanded question as displayed in BioASQ Social Network while composing a comment on it.

### 2.2.2 Users

Generally speaking, users are people who participate in the network. The cohort of users in the networks is comprised of question authors (experts) and interested individuals. In order for users to be able to participate in the network they are required to register an account. To do so they must provide an invitation code. See [section 4.1](#) for a more detailed description of the sign up process.

### 2.2.3 Activities

All actions of users (with the exception of sending private messages) on to the BioASQ Social Network generate activity objects. Activity objects can be compared to a log entry in a log file produced by a system. They are used to propagate user actions through the network in a consistent manner. Actions

generating activity objects are considered public and can be seen by all participants of the network. Those actions include:

- importing a question,
- commenting on a question,
- placing a vote on a question,
- following a question, and
- following a user.



Figure 2.2: Various activities as displayed by BioASQ Social Network.

The four types of activities are depicted in [Figure 2.2](#). Questions associated with the activity can be inspected in-place by clicking on the link labeled *a question*. Such expanded questions have all the functionality that questions in the *Questions Tab* can have. That is, users can comment, vote or follow/unfollow expanded questions. Comment activities allow composing a reply directly from the *Activities Tab*. In addition, activities provide links to the user that created it. Activity features described above work in every view that displays activities.

## 2.2.4 Messages

Private messages are for user-user communication. As the name suggests, private messages do not generate public activity. A message can only be seen by its sender and its receiver.

## 2.2.5 Notifications

BioASQ Social Network automatically sends email notifications to user *A* when the following events occur:

- user *A* receives a message from another user,
- someone starts following user *A*,
- someone writes a comment to a question user *A* is following,
- someone replies to a comment made by user *A*.

---

## Implementation

---

The BioASQ Social Network has been developed as a modern Web application based on the MEAN stack<sup>1</sup>. This software stack includes the following:

- MongoDB<sup>2</sup>, a document-oriented database;
- Express<sup>3</sup>, a server-side web framework;
- Angular<sup>4</sup>, client-side data-binding and view rendering framework;
- Node<sup>5</sup>, a JavaScript environment for server-side development.

In addition, the Bootstrap<sup>6</sup> CSS framework was used to to achieve a pleasant GUI that is consistent with the Annotation Tool, which also used that framework.

### 3.1 Architecture

Architectural components can be separated into two subsets: frontend and backend components. The former reside on the client while the latter are executed on the server. Data exchanged between frontend and backend is comprised entirely of JSON<sup>7</sup> objects.

The frontend components include view controllers, services, view models, and views. View controllers orchestrate interaction of client-side objects. With the help of services, which are orthogonal to controllers, they load data from or send data to the backend API. In creating the view model, they can transform data if necessary. The Angular rendering system interpolates views whenever a change in the

---

<sup>1</sup><http://mean.io>

<sup>2</sup><http://mongodb.org>

<sup>3</sup><http://expressjs.com>

<sup>4</sup><http://angularjs.org>

<sup>5</sup><http://nodejs.org>

<sup>6</sup><http://getbootstrap.com>

<sup>7</sup><http://json.org>

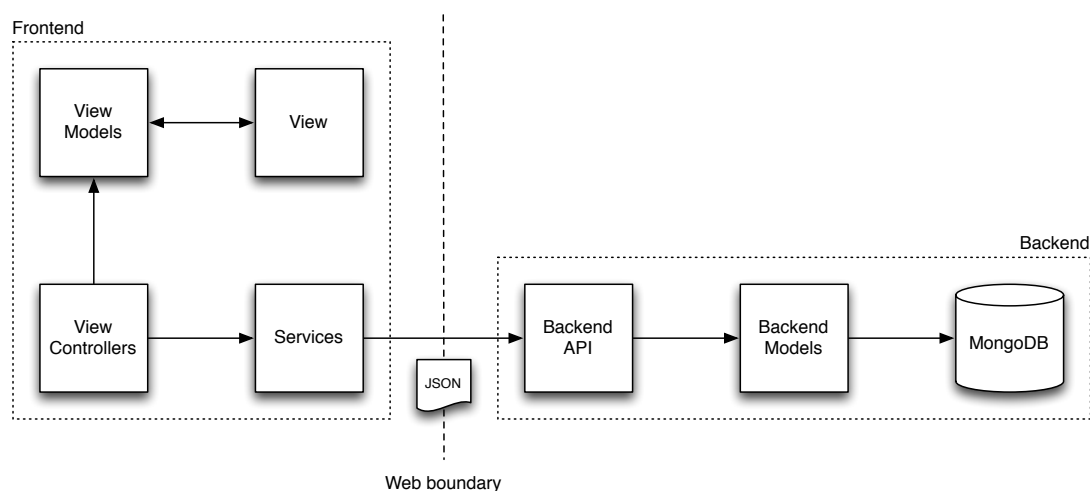


Figure 3.1: Architecture of the BioASQ Social Network. Arrow directions depict control flow. Note the arrow directions between view models and view that denote control flow in both directions (two-way data binding).

associated view model is detected and updates the model whenever the view changes due to user interaction or other events. Those model changed can also observed in controllers. Alternatively controller functions can be called by the view directly.

Due to the backend not containing any views or presentation logic it is kept quite simple. Each request is directed to the respective controller that can use models to interact with the database and retrieve the data desired by the client. Backend controllers can request cookie-based or password authentication. Authentication request are fulfilled before the controller takes action. In case of authentication failure the controller is not even executed.

## 3.2 Data Schema

The BioASQ Social Network's data schema is depicted in [Figure 3.2](#). The three main entities are Activity, Message, and User. Both, Activity and Message are linked to User via their *creator* attributes. In addition, Message is linked to another user via its *to* attribute. Each instance has a *type* attribute that denotes its entity name and an *id* attribute that is allocated by the database (not shown).

**Activity.** The Activity entity subsumes several entities with a similar structure. These are *Follow*, *Comment*, *Question*, and *Vote*. Some of the sub entities need additional properties that are stored directly in the Activity collection. For *Comment* these are `content`, `reply_of`, and `reply_count` while *Vote* stores an additional property called `dir` (direction).

**Message** The Message entity stores `content` and `title` as well the aforementioned relationships to *User*. In addition, it keeps a self-referencing relation called `reply_of` in case it is a reply to another message.

**User** The User entity stores standard user information like `first_name`, `last_name`, `email`, `img`, and some user preferences, currently only `notifications`. Authentication related properties like

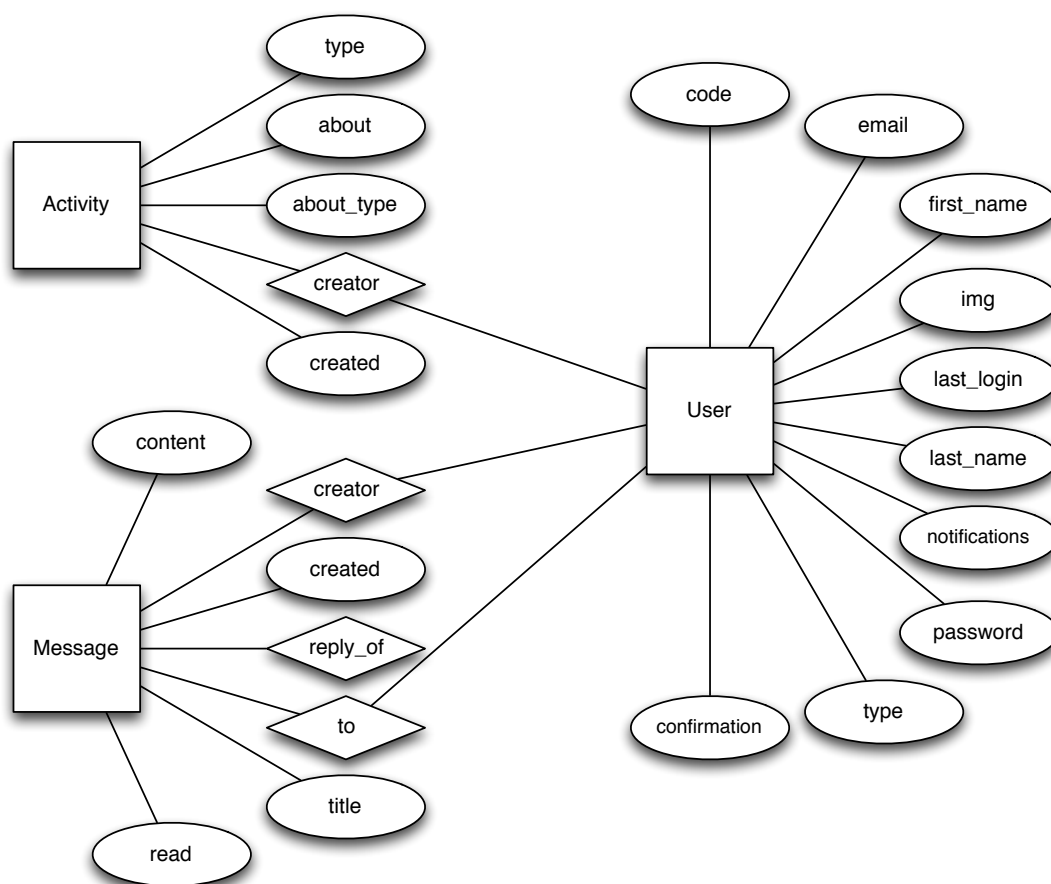


Figure 3.2: Data schema of the Activity, Message, and User entities.

code, confirmation, as well as an encrypted password property are also stored in the User entity.

**Question** For the schema of the Question entity please refer to Deliverable D3.3 for the Annotation Tool.

### 3.3 Deployment

The BioASQ Social Network is developed as an open source project. All source code is hosted on GitHub and can be accessed under

<https://github.com/AKSW/BioASQ-SN>.

External dependencies *Node* and *MongoDB* must be installed first. A script to download *Node* dependencies is provided and can be executed on a shell with the `npm install` command.

When all dependencies are installed a configuration file must be created. A configuration template is provided under the name `config.js-dist`. This template must be copied to a file named `config.js` in the same directory. Most default values can be left as is unless they are incompatible with the deployment requirements (e.g. if BioASQ Social Network is to be run under HTTP port

80 the respective configuration entry must be changed). Values that need changing in any case are `mail.user` and `mail.pass`. They must contain valid credentials for a Gmail account from which notification emails will be sent.

After that procedure the server can be started by executing the start script located in `bin/start`.

---

## Existing Functionality

---

### 4.1 Authentication

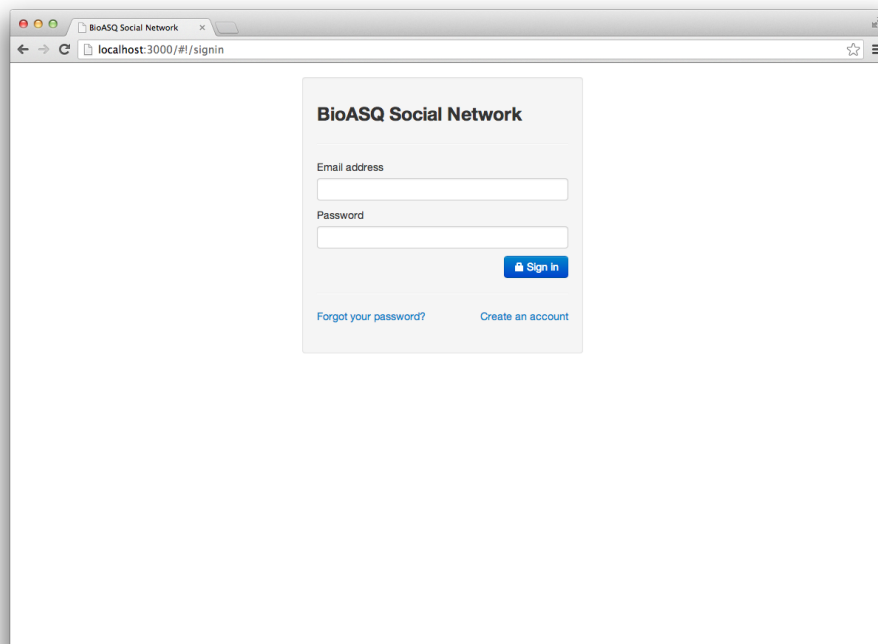
Before using BioASQ Social Network the user must be authenticated. There is no public information that can be seen without authentication. In order to authenticate, the user needs to sign in using the provided form and enter his email address and password (see [Figure 4.1](#)). In addition to standard sign in there are two more options provided: setting a new password in case the old one was lost and creating an account with BioASQ Social Network.

**Creating an account.** By clicking on *Create an account* the registration form as depicted in [4.2\(a\)](#) is presented. Once the form is completed and submitted by clicking on the button labeled *Register* an email is sent to the email address that was provided. This email contains a link that needs to be followed in order to activate the account. After that the user is automatically signed in. Alternatively, he can use the sign in form as described above.

**Setting a new password.** In case a user has forgotten her password and cannot sign in she should click the link labeled *Forgot your password?* The form depicted in [4.2\(b\)](#) is presented and the user should enter the same email address that she registered with. After clicking on the button labeled *Send Link* and if the provided email address was indeed one of a registered user a link will be sent to that address. The link contains a security code that can only be used once. If clicked it leads to a form similar to the registration form but only providing password fields where a new password can be created.

### 4.2 Home

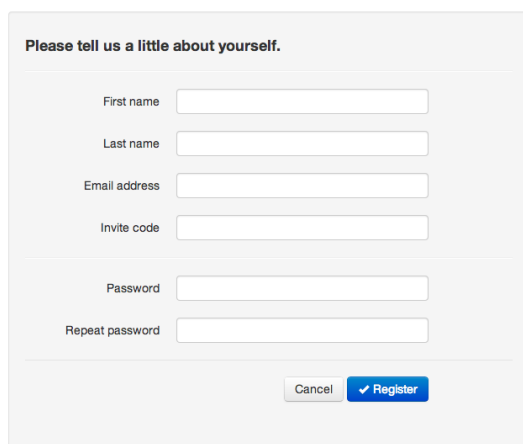
A user's *Home* tab shows a list of activities related to users and questions that the particular user follows. The top part of [Figure 4.3](#) shows an example of a user's home screen right after he signed into the BioASQ Social Network for the first time. Since he does not follow any questions or people yet, the home screen is empty. Once he starts to follow things he is interested in, his home screen starts to fill up with activity items (see [Figure 4.3](#)). This way, the user only sees activity for objects he cares about.



The screenshot shows a web browser window titled "BioASQ Social Network" with the address bar displaying "localhost:3000/#/signin". The main content area features a sign-in form with the following elements:

- BioASQ Social Network** (Section Header)
- Email address** (Text label) with an input field.
- Password** (Text label) with an input field.
- Sign in** (Blue button with a lock icon).
- [Forgot your password?](#) (Text link)
- [Create an account](#) (Text link)

Figure 4.1: Sign in



**Please tell us a little about yourself.**

First name

Last name

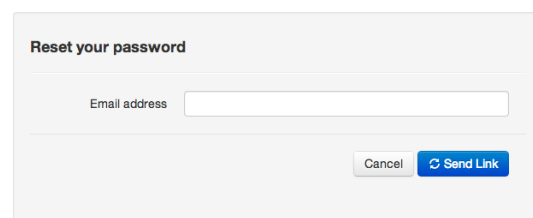
Email address

Invite code

Password

Repeat password

(a) Registration



**Reset your password**

Email address

(b) Reset

Figure 4.2: Registration and password reset.



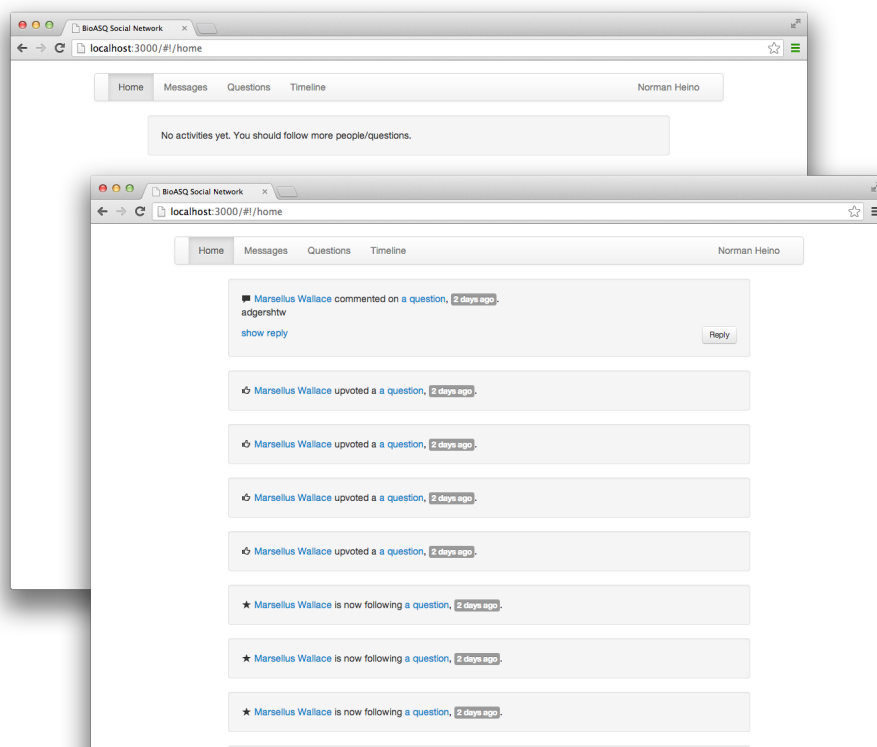


Figure 4.3: Home after registration (top) and with activities (bottom).

## 4.3 Messages

The BioASQ Social Network allows users to send each other short messages. The *Messages Tab* is where both received (*Inbox*) and sent (*Outbox*) messages are displayed. For composing a new message, several options exist:

- The button labeled *New Message* on the *Messages* tab can be clicked, which will display the message compose GUI as depicted in [Figure 4.4](#). The message can be sent to an arbitrary person from the list of possible recipients.
- When visiting another user's page the button labeled *Message* under the user's profile picture can be clicked which will open a message compose GUI similar to that in [Figure 4.4](#). The receiver of the message is automatically set to the currently visited user and cannot be changed.

Once a message is sent the recipient will be informed via email, if she has not decided to opt out of that feature.

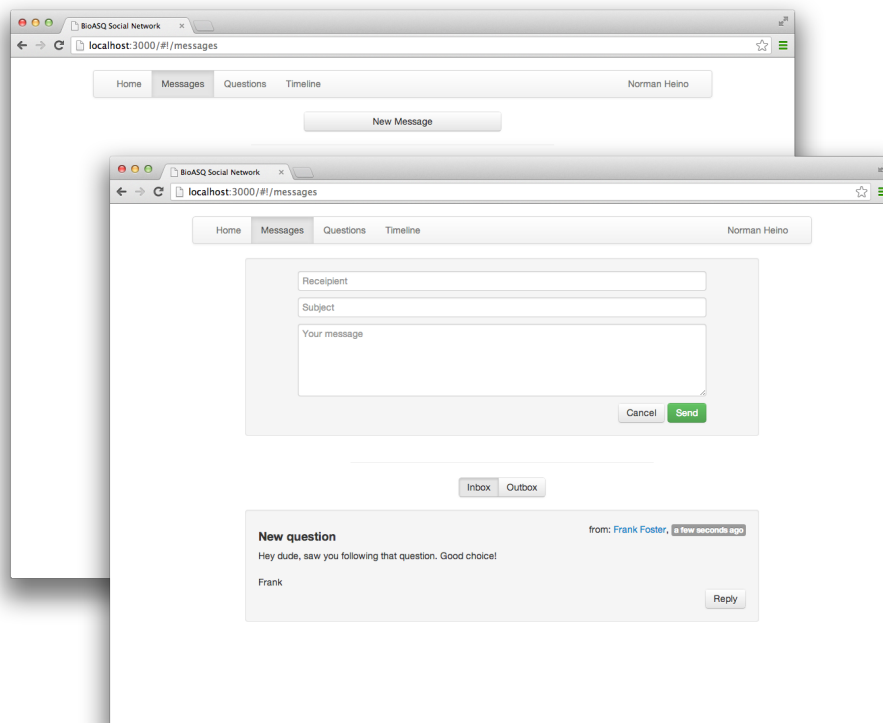


Figure 4.4: *Messages* tab showing a collapsed (top) and an expanded composing GUI (bottom).

## 4.4 Questions

Questions being the object of discourse in BioASQ Social Network have their own view. The Questions tab lists all questions currently stored in the network. By clicking on the title of a question the answer

of the question is revealed along with *concept*, *document*, *snippet*, and *statement* annotations (see [Figure 4.5](#)). Concept and document annotations are linked to their source documents (e. g. MeSH concept description or PubMed article page).

The *Questions* tab empowers users of the network to express their opinion on a particular questions by

- voting in favor of or against a question,
- *following* a question to be informed about activity that takes places,
- *commenting* on a question.

The *Questions* tab can be sorted by creation date or number of votes. This allows users to easily see the latest or most preferred items and participate in the discussion.

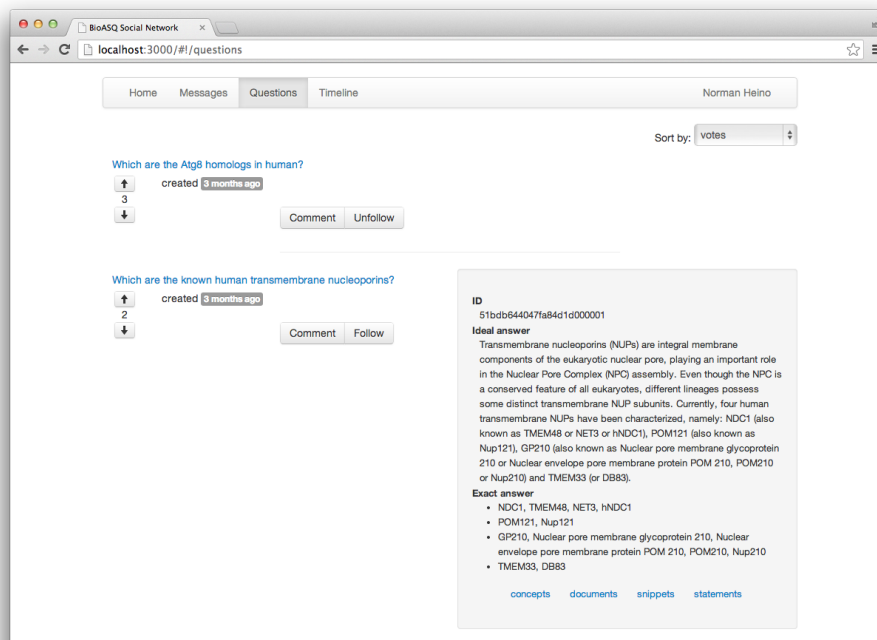


Figure 4.5: Question list showing a collapsed question at the top and an expanded question at the bottom.

## 4.5 Timeline

An important topic in every network is that of discoverability. In BioASQ Social Network we tackle discoverability by providing a global view of all activity that takes place in the network. Taking inspiration from Twitter, this view is called *Timeline*. This way, users are able to discover new questions they might be interested in or people they might want to follow.

The Timeline view is automatically sorted by creation date with the most recently created item being displayed at the top. Since it looks very similar to the *Home* view it is not depicted separately.

## 4.6 User/Preferences

Each user has her own profile page (see [Figure 4.6](#)) that is partially visible to all other users. This page shows some basic information about the user such as name and a profile picture. In addition, the users activities, followings or followers by clicking the buttons labeled *Activity*, *Following* or *Followers*. The options under the profile picture differ depending on whether the profile being watched is that of the currently signed in user or that of another user.

**Profile page of another user.** When visiting the profile page of another user two options are provided: sending a message to the user whose profile page is being displayed (see [section 4.3](#)) or following/unfollowing that user. Both functions are available by clicking the button labeled *Message*, or *Follow/Unfollow*, respectively.

**Signed in user's page.** If the signed in user visits his own profile page, she can sign out or change her profile settings, by clicking one of the buttons labeled *Sign out* or *Preferences*, respectively.

The preferences dialog (see [Figure 4.7](#)) can be used to change the data that was used for signing up. Notification settings can be changed here as well.

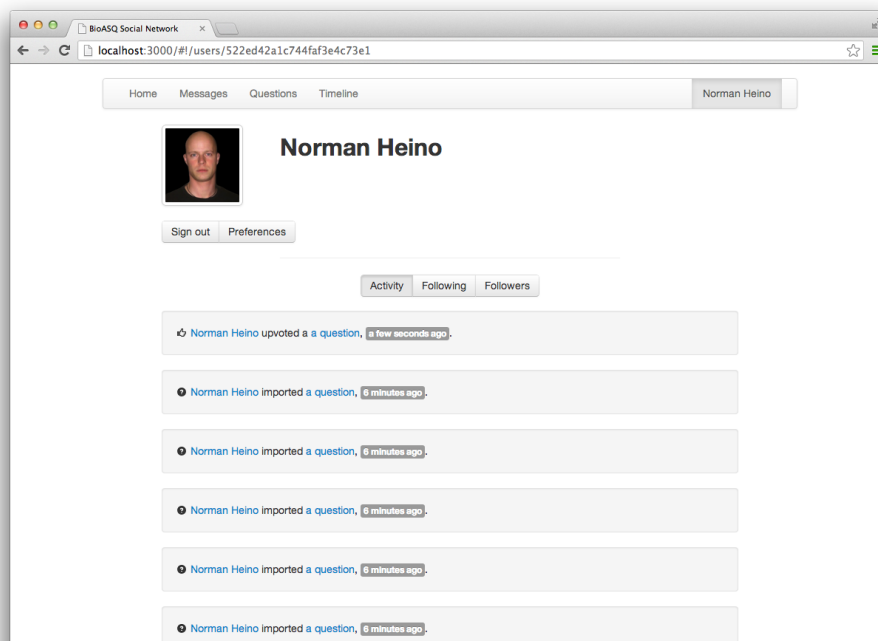


Figure 4.6: User

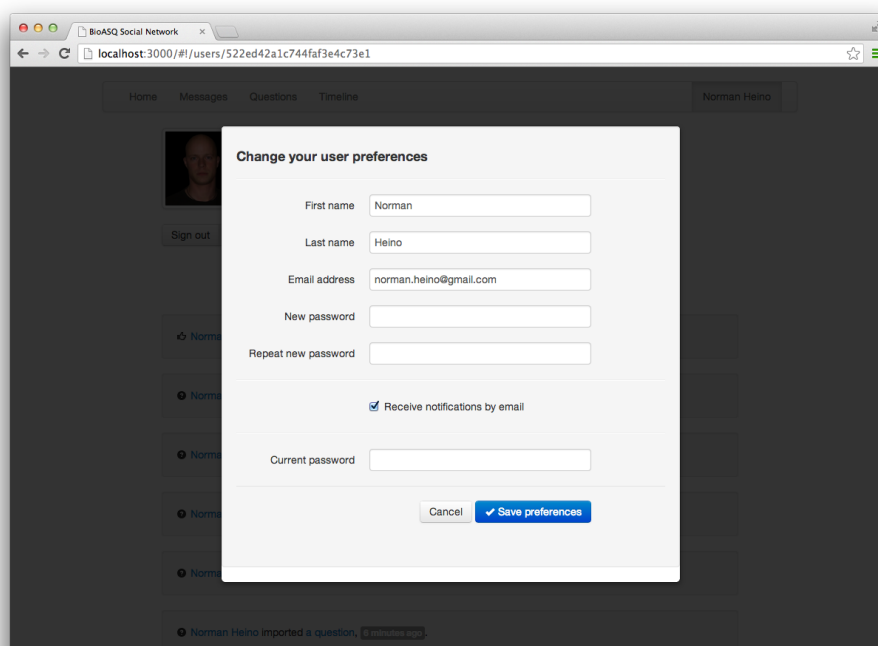


Figure 4.7: Preferences

---

## Envisioned Extensions

---

Though the current implementation state of BioASQ Social Network fulfills all requirements as outlined in [chapter 1](#) some extensions are envisioned that could possibly reduce the amount of manual work required for benchmark creation.

**Integration with the BioASQ Annotation Tool** Currently, importing questions from the BioASQ Annotation Tool is a manual process. Since both BioASQ Social Network and BioASQ Annotation Tool are based on the same technology a tighter integrating is feasible but was not envisioned in the project proposal. The following extensions could provide for such integration:

- Automatic import of questions that are changed in the BioASQ Annotation Tool and creation of associated BioASQ Social Network activity objects.
- Allowing direct deep links to questions in the BioASQ Annotation Tool. Those links can be added to questions displayed in BioASQ Social Network.

**Supporting larger user groups.** A larger number of users can possibly work on several challenges at once. To support such a scenario the notion of *user groups* could be introduced in BioASQ Social Network. In addition, *user roles* would be needed to allow certain users to assume *administrative* functions w. r. t. a group (naming the group, managing its participants, etc.).