



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

Project **FP7-318652 / BioASQ**

Deliverable **D5.2**

Distribution **Public**



<http://www.bioasq.org>

Challenge Evaluation Report

Prodromos Malakasiotis, Ion Androutsopoulos and
Ioannis Pavlopoulos

Status: Final (Version 1.0)

December 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	http://www.bioasq.org
Project start	October 2012
Project duration	2 years
EC Project Officer	Martina Eydner

Deliverable

Deliverable type	Report
Distribution level	Public
Deliverable Number	D5.2
Deliverable title	Challenge Evaluation Report
Contractual date of delivery	M15 (December 2013)
Actual date of delivery	December 2013
Relevant Task(s)	WP5/Task 5.2
Partner Responsible	AUEB-RC
Other contributors	
Number of pages	29
Author(s)	Prodromos Malakasiotis, Ion Androutopoulos and Ioannis Pavlopoulos
Internal Reviewers	Sergios Petridis
Status & version	Final
Keywords	Challenge evaluation, questionnaires

Executive Summary

This deliverable evaluates the first cycle of the BIOASQ challenge both quantitatively and qualitatively. The evaluation was conducted by measuring the number of participants both in the challenge and the workshop, the number of visitors in our websites, and the downloads of the benchmark data. We also used questionnaires distributed to the participants of the challenge and the workshop and to the team of biomedical experts. The analysis showed that we managed to successfully organize the first cycle of the challenge. The participation was satisfactory, especially in Task 1A, and most of the participants are not only willing to participate in the second cycle of BIOASQ, but they also intend to recommend it to other research groups. We also managed to organize a successful workshop leaving a good overall impression to the participants. Finally, we had a very good cooperation with the team of biomedical experts, providing them with all the help and tools they needed for the creation of the benchmark datasets for Task 1B and the evaluation of the systems' responses, again for Task 1B.

Contents

1	Introduction	1
2	Evaluation via web–site statistics and workshop participation	2
3	Evaluation via questionnaires	5
3.1	Challenge evaluation	5
3.2	Workshop evaluation	5
3.3	Interaction with the team of biomedical experts evaluation	7
A	Questionnaires used	13

List of Figures

2.1	Unique visitors at BIOASQ official site from November '12 until November '13.	4
2.2	Unique visitors at the evaluation platform from April '13 until November '13.	4
3.1	Overall impression for BIOASQ by the participating teams.	6
3.2	Willingness to participate in the second cycle of BIOASQ by the participating teams of the first cycle.	6
3.3	Willingness to recommend BIOASQ by the participating teams of the first cycle.	6
3.4	Difficulty to understand BIOASQ tasks.	7
3.5	Assesment of the presentation of BIOASQ at the workshop.	7
3.6	Assesment of the invited talks at the workshop.	8
3.7	Assesment of the scientific presentations at the workshop.	8
3.8	Percentage of participation of the workshop participants in the challenge.	8
3.9	Willingness to participate in the second cycle of BIOASQ by the participants of the workshop.	9
3.10	Overall impression for the annotation tool by the team of biomedical experts.	9
3.11	Willingness of the team of biomedical experts to use the annotation tool again.	9
3.12	Willingness of the team of biomedical experts to use the annotation tool for their work (e.g., to organize a search).	10
3.13	Willingness of the team of biomedical experts to recommend the annotation tool.	10
3.14	Overall impression for the assessment tool by the team of biomedical experts.	10
3.15	Willingness of the team of biomedical experts to use the assessment tool again.	11
3.16	Willingness of the team of biomedical experts to recommend the assessment tool.	11
3.17	Assessment of the interaction of the team of biomedical experts with us.	11
3.18	Assessment of the saving procedure of the annotation tool.	12
3.19	Assessment of the concepts search procedure of the annotation tool.	12
3.20	Assessment of the RDF triples search procedure of the annotation tool.	12

List of Tables

2.1	Summary of the challenge statistics.	2
2.2	Teams participating in both tasks of the first cycle of BIOASQ challenge.	3
2.3	Tasks 1A and 1B dataset downloads.	3

Introduction

This deliverable evaluates the first cycle of the BIOASQ challenge. For that purpose we provide figures measuring the number of the participants both in the challenge and the workshop, the visitors in our web-sites, and the downloads of the benchmark data. We also distributed questionnaires to the participants of the challenge and the workshop, as well as to the team of biomedical experts. The questionnaires aimed to assess the quality, appropriateness, and diversity of the challenge benchmarks and evaluation measures, the quality of the support to the participants, and the adequacy and quality of the challenge evaluation infrastructure.

Chapters 2 and 3 below present the evaluation of the first cycle of the BIOASQ challenge based on web statistics and questionnaires respectively. Finally, Appendix A presents the questionnaires used for the qualitative evaluation of BIOASQ.

Evaluation via web-site statistics and workshop participation

One of the most important features for a successful challenge is attracting participants. Table 2.1 summarizes the most important statistics concerning the evaluation of the challenge. In more detail, we had 117 users registered on the evaluation platform.¹ Moreover, several teams around the globe participated in the first cycle of the BIOASQ challenge, especially in Task 1A, including key players, like Mayo Clinic, NLM, and Toyota Technological Institute. Table 2.2 lists the participating teams of the first cycle of the BIOASQ challenge. We have also organized a successful Workshop following CLEF 2013 as a post event in Valencia, with 30 participants and important invited speakers.²

Another important objective of the challenge is to establish BIOASQ as a reference point for the biomedical community. A look at the statistics from our websites indicates that we are moving towards the right direction. Figures 2.1 and 2.2 show that even after the end of the challenge there was high traffic on our websites indicating the community's interest in BIOASQ. In addition, Task 1A continued to run in a “non-challenge” mode helping towards that direction. Table 2.3 shows the numbers of downloads for the datasets of Tasks 1A and 1B.

¹<http://bioasq.lip6.fr/>

²Consult <http://www.clef2013.org/> for information on CLEF 2013. For more details about the BioASQ workshop, visit <http://www.bioasq.org/workshop>.

Registered users	117
Datasets downloads	697
Teams	11
Workshop participants	30
Number of biomedical experts	10

Table 2.1: Summary of the challenge statistics.

Task 1A participation (46 systems, 11 teams)	
Mayo Clinic	USA
University of Alberta	CANADA
Aristotle University of Thessaloniki + Atypon	GREECE
University of Vigo	SPAIN
University of Colorado	USA
NCBI, NLM	USA
Universite de Rouen	FRANCE
Fudan University	CHINA
UCSD	USA
Toyota Technological Institute	JAPAN
Imran	PAKISTAN
Task 1B participation	
Phase A (4 systems, 2 teams)	
Mayo Clinic	USA
University of Alberta	CANADA
Phase B (7 systems, 2 teams)	
University of Alberta	CANADA
Toyota Technological Institute	JAPAN

Table 2.2: Teams participating in both tasks of the first cycle of BIOASQ challenge.

Dataset downloads	
Task 1A	584 (including 78 after the end of the challenge)
Task 1B	113

Table 2.3: Tasks 1A and 1B dataset downloads.

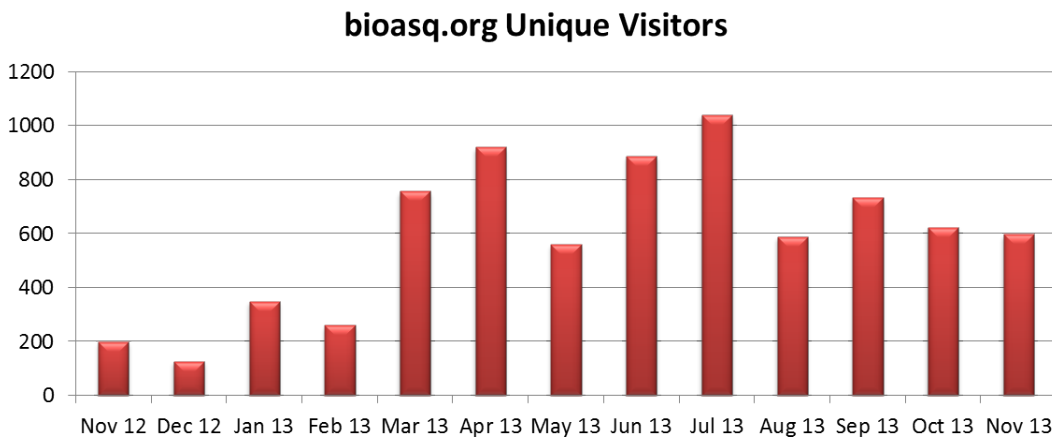


Figure 2.1: Unique visitors at BIOASQ official site from November '12 until November '13.

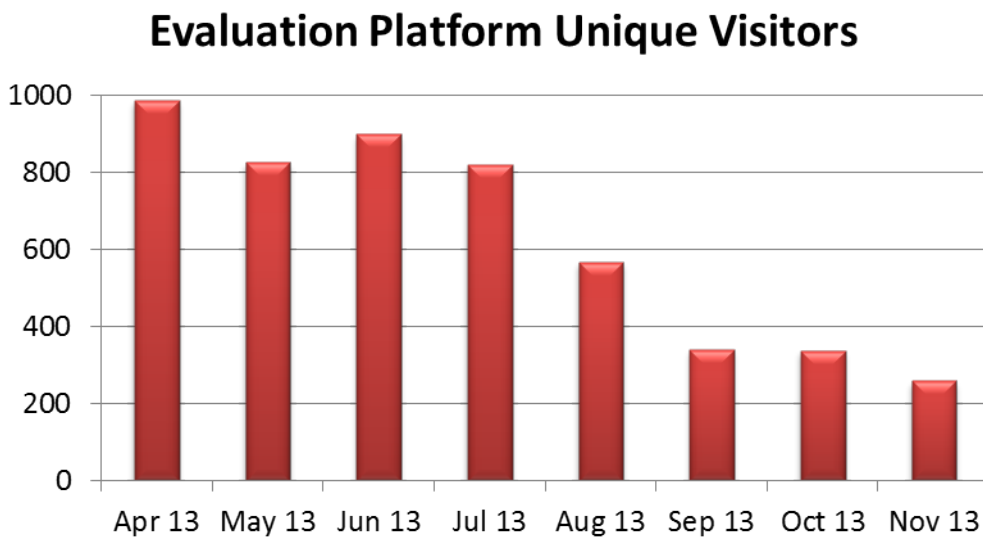


Figure 2.2: Unique visitors at the evaluation platform from April '13 until November '13.

Evaluation via questionnaires

Having evaluated BIOASQ via web–site statistics, we move on to the results collected via questionnaires. We created and distributed questionnaires to the workshop participants, the participating teams of the challenge, and the team of biomedical experts. The questionnaires aim, among other goals, to assess the quality, appropriateness, and diversity of the challenge benchmarks and evaluation measures, the quality of the support to the participants, and the tools and the support provided to the team of biomedical experts for the creation of the benchmarks. For a more detailed view of the questionnaires see Appendix A.

3.1 Challenge evaluation

The questionnaire distributed to the teams participating in the first cycle of BIOASQ included questions targeting several aspects of the individual tasks, like the quality of the datasets, the technical support etc. Additionally, more general questions were provided to capture the overall impression of the participants for BIOASQ. Figures 3.1–3.3 summarize the results for the most general questions. According to these figures the participants are in general satisfied with the challenge and as a consequence, not only are they willing to participate in the next cycle of the challenge, but they are willing to recommend BIOASQ to other research groups as well. However, there were some difficulties in the general understanding of the challenge as indicated by Figure 3.4. To alleviate this problem we modified the guidelines of the challenge for both tasks in order to make them clearer for the potential participants. Another problem we had to face was the rather small participation in Task 1b. For that we came up with two strategies. The first one was to identify all the research areas that relate to BIOASQ and modify the first page of our official website to include this information. The second one was to create a list of potential participants and invite them, through personal contact, to participate in the challenge.

3.2 Workshop evaluation

Figures 3.5 – 3.9 summarize the statistics collected by the questionnaires distributed to the participants of the workshop. Overall, the participants were satisfied with the different aspects (scientific presentations,



Figure 3.1: Overall impression for BIOASQ by the participating teams.

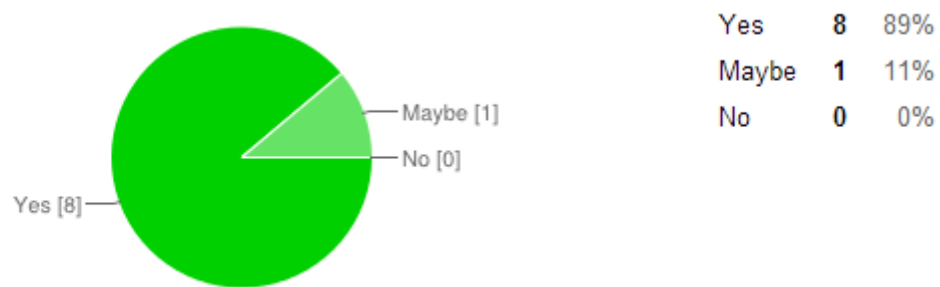


Figure 3.2: Willingness to participate in the second cycle of BIOASQ by the participating teams of the first cycle.

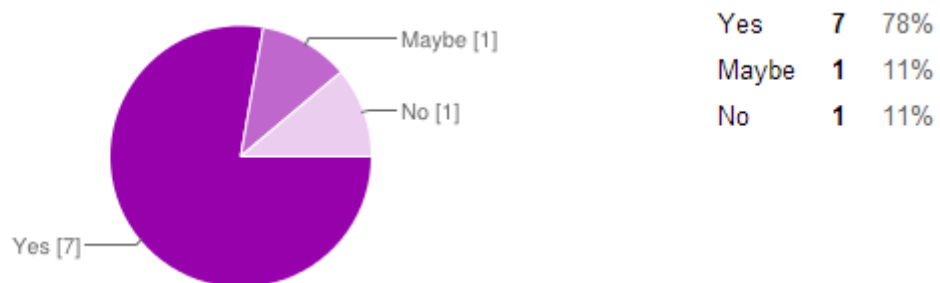


Figure 3.3: Willingness to recommend BIOASQ by the participating teams of the first cycle.



Figure 3.4: Difficulty to understand BIOASQ tasks.

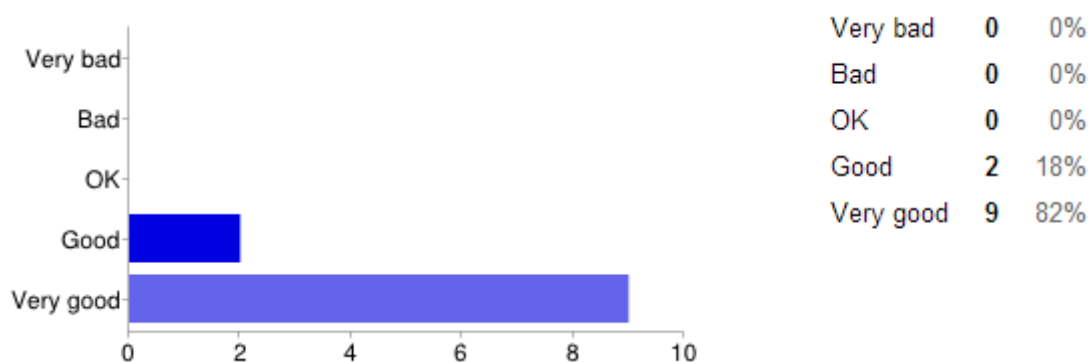


Figure 3.5: Assessment of the presentation of BIOASQ at the workshop.

invited speakers, etc.) of the workshop. Apart from that, even though half of the workshop participants did not participate in the first cycle of BIOASQ, almost all of them are showing a tendency to participate in the second cycle.

3.3 Interaction with the team of biomedical experts evaluation

One of the most difficult goals of BIOASQ was the creation of the benchmark data for Task 1B. For that purpose we had to coordinate a team of biomedical experts and provide them with tools and technical support that would help them in the creation of the benchmarks. In addition, the team of biomedical experts manually assessed the responses of the participating teams in Task 1B, again with the assistance of a tool. After the end of the first cycle of the challenge, we distributed questionnaires to the biomedical experts, in order to assess the quality of the tools and their interaction with us. Figures 3.10 – 3.17 show that we had a very good cooperation with the team of biomedical experts. The experts were not only satisfied by the tools, but they are also willing to use them again in the future and even recommend them to others. Particularly for the annotation tool they are willing to use it for their own work, especially if it is improved. One of the improvements to be made is the saving procedure since, according to the questionnaires, this was the most problematic feature (Figure 3.18). Another issue raised was the search for concepts and RDF triples (Figures 3.19 and 3.20). Taking into consideration all the issues raised by the team of biomedical experts we made all the necessary improvements to the annotation tool to make the creation of the second cycle benchmarks as smooth as possible.

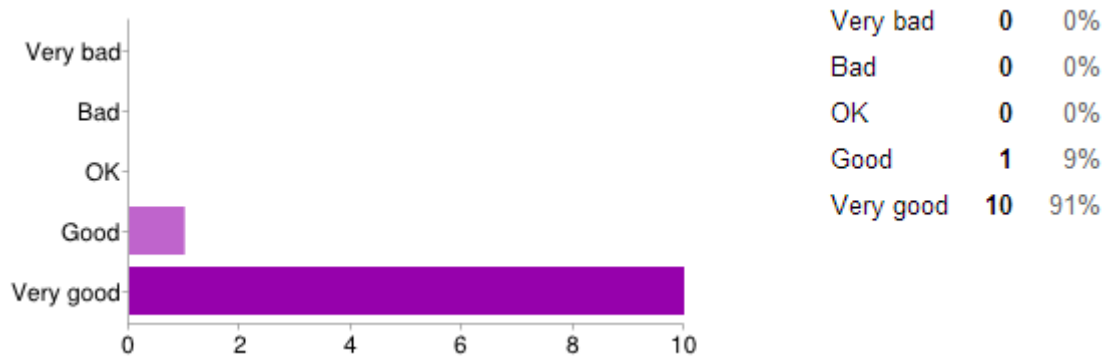


Figure 3.6: Assesment of the invited talks at the workshop.

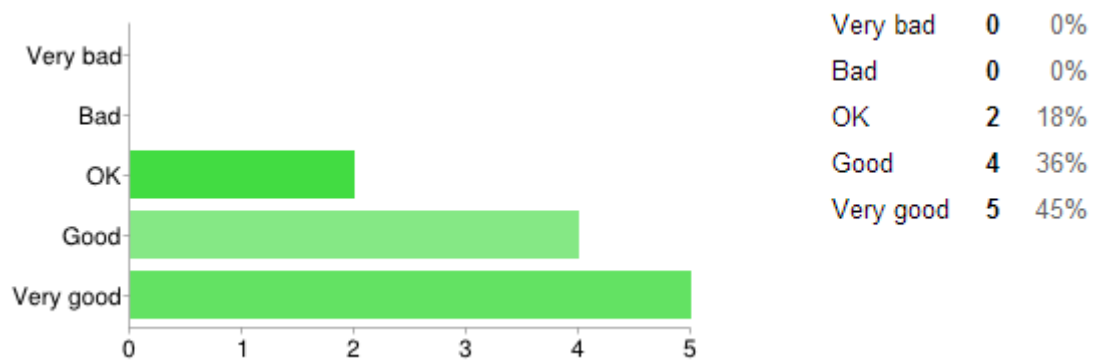


Figure 3.7: Assesment of the scientific presentations at the workshop.

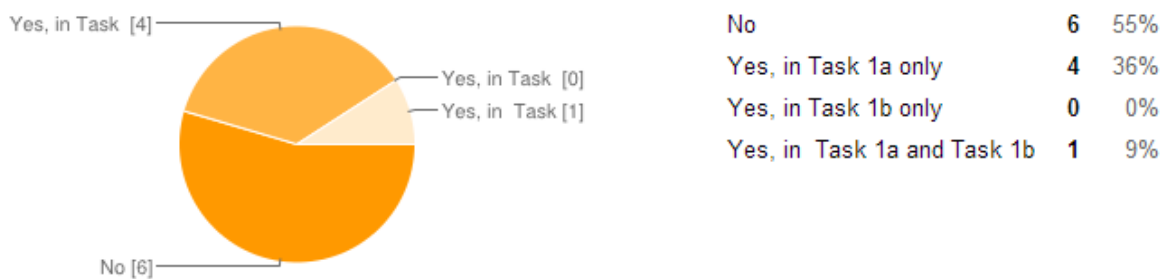


Figure 3.8: Percentage of participation of the workshop participants in the challenge.

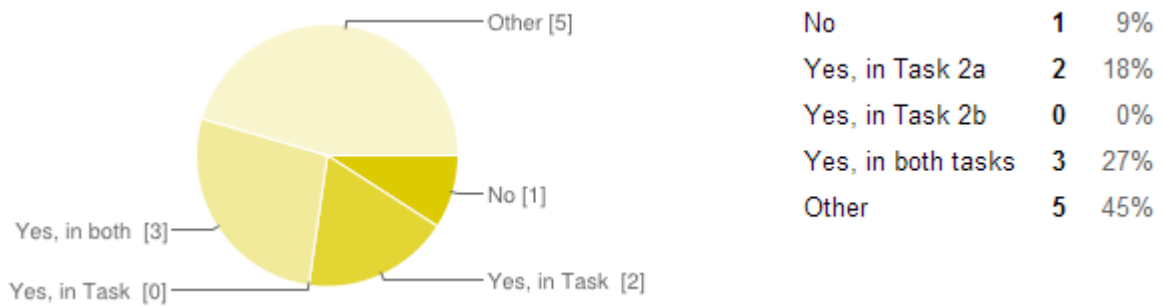


Figure 3.9: Willingness to participate in the second cycle of BIOASQ by the participants of the workshop.



Figure 3.10: Overall impression for the annotation tool by the team of biomedical experts.

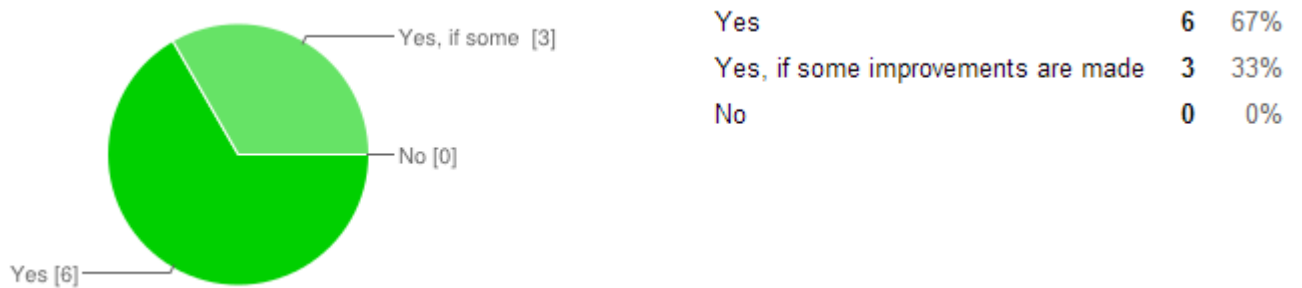


Figure 3.11: Willingness of the team of biomedical experts to use the annotation tool again.

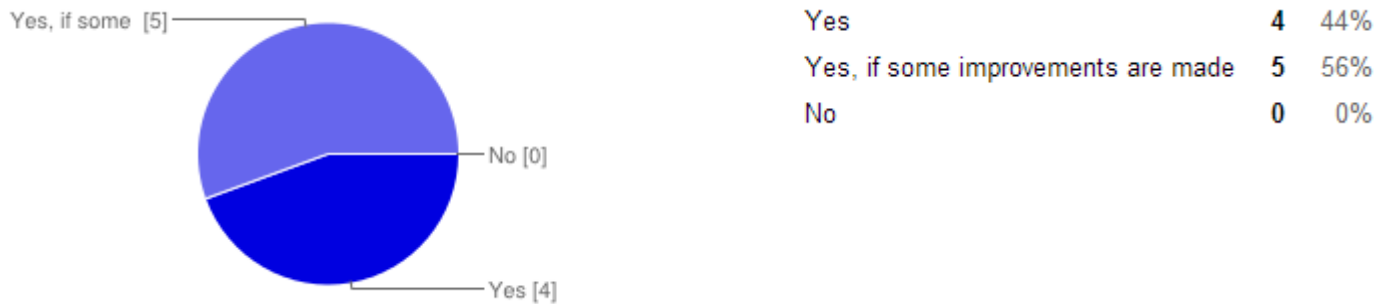


Figure 3.12: Willingness of the team of biomedical experts to use the annotation tool for their work (e.g., to organize a search).

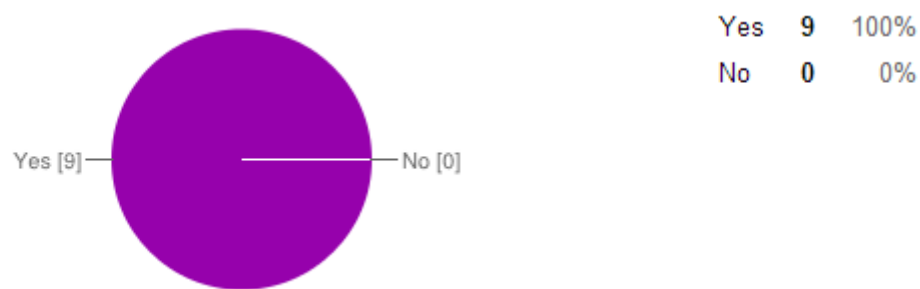


Figure 3.13: Willingness of the team of biomedical experts to recommend the annotation tool.



Figure 3.14: Overall impression for the assessment tool by the team of biomedical experts.

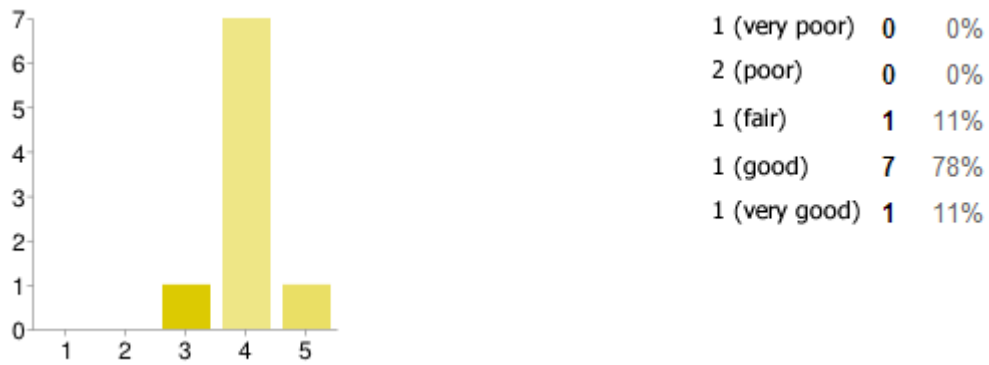


Figure 3.15: Willingness of the team of biomedical experts to use the assessment tool again.



Figure 3.16: Willingness of the team of biomedical experts to recommend the assessment tool.



Figure 3.17: Assessment of the interaction of the team of biomedical experts with us.



Figure 3.18: Assessment of the saving procedure of the annotation tool.



Figure 3.19: Assessment of the concepts search procedure of the annotation tool.



Figure 3.20: Assessment of the RDF triples search procedure of the annotation tool.

A

Questionnaires used

In the following pages we include the questionnaires used for the qualitative evaluation of the first cycle of BIOASQ.

BioASQ evaluation form for challenge participants

Short survey on the quality of the challenge, from the participant point of view. For more information about BioASQ, please visit <http://bioasq.org>.

* Required



Learning about BioASQ

1. How did you hear about BioASQ? *

Mark only one oval.

- Web site
- Mailing list
- Personal contact
- Twitter
- LinkedIn
- Other:

2. How easy was it to understand the tasks overall? *

Mark only one oval.

	1	2	3	4	5	
Very difficult	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Very easy

Registration

3. Please rate the registration process according to the following. *

Mark only one oval per row.

	Very bad	Bad	Fair	Good	Very good
Registration guidelines	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Registration platform	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Task 1A

Please answer the following questions only if you participated in Task 1A.

4. Please rate Task 1A according to the following.

Mark only one oval per row.

	Very bad	Bad	Fair	Good	Very good
Guidelines	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Data format	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Data quality	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Downloading and uploading procedures	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Scheduling (e.g., release of test sets, time to submit results, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Evaluation methods	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Notifications (e.g., release of test sets, evaluation results, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Technical support	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Overall impression	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

5. Frequency and Size of test sets *

Each batch was formed by 6 test sets, released weekly. In case you participated again, you would prefer that:

Mark only one oval.

- conditions remain the same
- each batch is formed by three double-size test sets, released one every two weeks
- either way is OK
- Other:

6. Additional comments

Anything not covered by the questions above.

.....

.....

.....

.....

.....

Task 1B Phase A

Please answer the following questions only if you participated in Task 1B Phase A.

7. Please rate Task 1B Phase A according to the following.

Mark only one oval per row.

	Very bad	Bad	Fair	Good	Very good
Guidelines	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Data format	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Data quality	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Number of questions	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Diversity of questions	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Difficulty of questions (how difficult were the questions)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Downloading and uploading procedures	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Other supporting software and web-services (e.g., searching for concepts, articles, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Scheduling (e.g., release of test sets, time to submit results, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Evaluation methods	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Notifications (e.g., release of test sets, evaluation results, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Technical support	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Overall impression	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

8. Additional comments

Anything not covered by the questions above.

.....

.....

.....

.....

.....

Task 1B Phase B

Please answer the following questions only if you participated in Task 1B Phase B.

9. Please rate Task 1B Phase B according to the following.

Mark only one oval per row.

	Very bad	Bad	Fair	Good	Very good
Guidelines	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Data format	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Data quality	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Number of questions	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Diversity of questions	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Difficulty of questions (how difficult were the questions)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Downloading and uploading procedures	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Scheduling (e.g., release of test sets, time to submit results, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Evaluation methods	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Notifications (e.g., release of test sets, evaluation results, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Technical support	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Overall impression	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

10. **Additional comments**

Anything not covered by the questions above.

.....

.....

.....

.....

.....

Overall impression

11. **What is your overall impression of BioASQ ***

Mark only one oval.

	1	2	3	4	5	
Very bad	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Very good

12. **Will you participate in the next year's BioASQ Challenge? ***

Mark only one oval.

- Yes
- Maybe
- No

13. **Will you recommend the BioASQ Challenge to others? ***

Mark only one oval.

- Yes
- Maybe
- No

Contact details

14. **Name ***

.....

15. **Email ***

.....

16. **May we contact you in due course? ***

Mark only one oval.

Yes

No

Workshop Assessment Survey BioASQ

Short survey to assess the BioASQ workshop. More information on the project can be found at <http://bioasq.org/>.

* Required

General Information

Please provide us with information about your participation to the workshop

1. Where did you hear about BioASQ *

Mark only one oval.

- Mailing List
- Personal Contact
- LinkedIn
- Twitter
- Other:

2. Did you participate in the challenge? *

Mark only one oval.

- No
- Yes, in Task 1a only
- Yes, in Task 1b only
- Yes, in Task 1a and Task 1b

3. If not, why are you not participating?

Mark only one oval.

- Time restrictions
- Not my area of research
- BioASQ unknown so far
- Other:

4. Are you planning to participate in the BioASQ track at CLEF 2014? *

Mark only one oval.

- No
- Yes, in Task 2a
- Yes, in Task 2b
- Yes, in both tasks
- Other:

Workshop Organization

5. Please evaluate the following parts of the workshop. *

Mark only one oval per row.

	Very bad	Bad	OK	Good	Very good
Scientific presentations	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Keynotes	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
BioASQ presentation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Panel discussion	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

6. Do you have any other suggestions to help us improve the organization of the workshop?

.....

.....

.....

.....

.....

7. How could the promotion of the workshop be improved?

.....

.....

.....

.....

.....

Personal Details

8. Name

.....

9. Email Address

.....

10. May we follow up with you if required?

Mark only one oval.

Yes

No

BioASQ Biomedical experts questionnaire

Please rate your experience as a member of the BioASQ team of biomedical experts

* Required



Annotation tool

Please rate the following functionalities of the annotation tool on a scale of 1 - 5

1. Very poor: I was unable to use it.
2. Poor: Needs substantial improvements.
3. Fair: I could use it but still needs some improvements
4. Good: Almost everything worked smoothly. With a few improvements it would be excellent
5. Very good: Everything worked smoothly. No improvement is needed.

1. Registration *

Mark only one oval.

	1	2	3	4	5	
Very poor	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Very good

2. Question creation *

Mark only one oval.

1	2	3	4	5
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

3. Concepts search *

Mark only one oval.

1	2	3	4	5
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

4. Documents search *

Mark only one oval.

1	2	3	4	5
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

5. Statements search *

Mark only one oval.

1	2	3	4	5
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

6. Snippets annotation *

Mark only one oval.

1	2	3	4	5
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

7. Exact answer creation *

Mark only one oval.

1	2	3	4	5
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

8. Ideal answer creation *

Mark only one oval.

1	2	3	4	5
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

9. Saving your work *

Mark only one oval.

1	2	3	4	5
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

10. Overall impression *

Mark only one oval.

1	2	3	4	5
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

11. **What did you like most in the tool?**

.....

.....

.....

.....

.....

12. **What did you like less in the tool?**

.....

.....

.....

.....

.....

13. **Would you use this tool again in the future? ***

Mark only one oval.

- Yes
- Yes, if some improvements are made
- No

14. Please give a short justification

.....

.....

.....

.....

.....

15. **Would you recommend this tool to others? ***

Mark only one oval.

- Yes
- No

16. Please give a short justification

.....

.....

.....

.....

.....

17. **Do you think you could use the tool in your own work (e.g., to organize a search)?**

Mark only one oval.

- Yes
- Yes, if some improvements are made
- No

18. Please give a short justification

.....

.....

.....

.....

.....

Assessment Tool

Please rate the following functionalities of the assessment tool on a scale of 1 - 5

- 1. Very bad: I was unable to use it.
- 2. Bad: Needs substantial improvements.
- 3. Fair: I could use it but still needs some improvements
- 4. Good: Almost everything worked smoothly. With a few improvements it would be excellent
- 5. Very good: Everything worked smoothly. No improvement is needed.

19. **Ideal answers assessment ***

Mark only one oval.

1 2 3 4 5

20. **Exact answers assessment ***

Mark only one oval.

1 2 3 4 5

21. **Snippets assessment ***

Mark only one oval.

1 2 3 4 5

22. **Concepts assessment ***

Mark only one oval.

1	2	3	4	5
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

23. **Documents assessment ***

Mark only one oval.

1	2	3	4	5
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

24. **Statements assessment ***

Mark only one oval.

1	2	3	4	5
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

25. **Saving your work ***

Mark only one oval.

1	2	3	4	5
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

26. **Overall impression ***

Mark only one oval.

1	2	3	4	5
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

27. **What did you like most in the tool?**

.....

.....

.....

.....

.....

28. What did you like less in the tool?

.....
.....
.....
.....
.....

29. Would you use this tool again in the future? *

Mark only one oval.

1	2	3	4	5
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

30. Please give a short justification

.....
.....
.....
.....
.....

31. Would you recommend this tool to others? *

Mark only one oval.

1	2	3	4	5
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

32. Please give a short justification

.....
.....
.....
.....
.....

Overall impression

33. Please rate your interaction with the BioASQ team. *

Mark only one oval.

	1	2	3	4	5	
Very bad	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Very good

34. Please add ideas/comments of possible ways of improving BioASQ.

.....

.....

.....

.....

.....

Contact details

35. Name *

.....

36. Email *

.....
