



Overview of BioASQ task a

Large-Scale Online Biomedical Semantic Indexing

A. Nentidis, A. Krithara, G. Katsimpras, and G. Paliouras



NATIONAL CENTRE FOR
SCIENTIFIC RESEARCH "DEMOKRITOS"

16 November 2022

6th workshop on computer-assisted subject indexing, online

Introduction

What is BioASQ

A competition

- ▶ BioASQ is a series of **challenges** on **biomedical semantic indexing** and **question answering (QA)**.
- ▶ Participants are required to semantically index content from **large-scale** biomedical resources (i.e. MEDLINE) and/or
- ▶ to assemble data from **multiple heterogeneous sources** (e.g. scientific articles, knowledge bases, databases)
- ▶ to compose **informative answers** to biomedical natural language questions.

Introduction

What is BioASQ

Structured into Shared Tasks

- ▶ **Task A:** Large-Scale Online Biomedical Semantic Indexing.
- ▶ **Task B:** Biomedical Semantic Question Answering.
- ▶ **Task C:** Funding Information Extraction.
- ▶ **Task Synergy:** Question Answering for Open Developing Issues.
- ▶ **Task MESINESP:** Medical Semantic Indexing in Spanish.
- ▶ **Task DisTEMIST:** Disease Text Mining and Indexing.
- ▶ **Task MedProcNER:** Medical Procedure Text Mining and Indexing.

Introduction

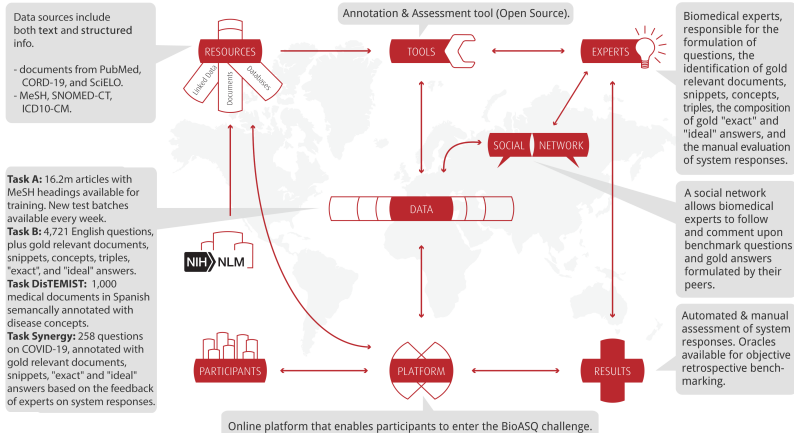
What is BioASQ

Some facts

- ▶ The challenge has run ten times, in **ten years**.
 - ▶ Annual evaluation cycle and **workshop**.
 - ▶ More than **100 teams** from **28 countries**.
- ▶ **Participation** can be **partial** (any task, subtask, response type).
- ▶ **Prizes** for each task/subtask.
- ▶ BioASQ datasets, infrastructure, evaluation services etc.
 - ▶ Available as **open-source** projects.

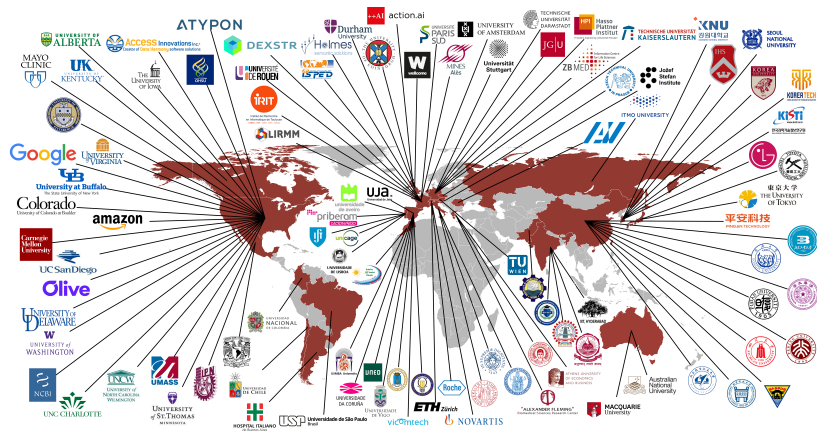
Introduction

BioASQ ecosystem



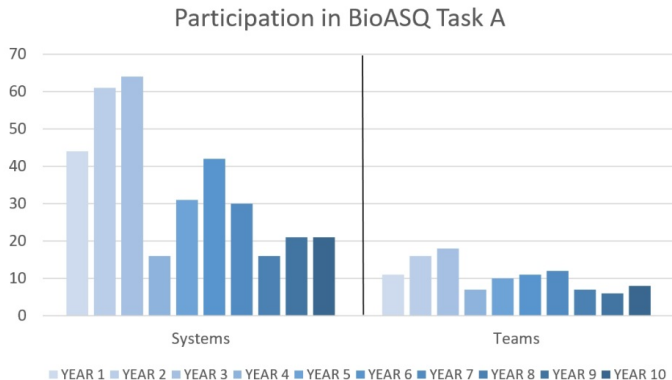
Challenge Participation

Overall



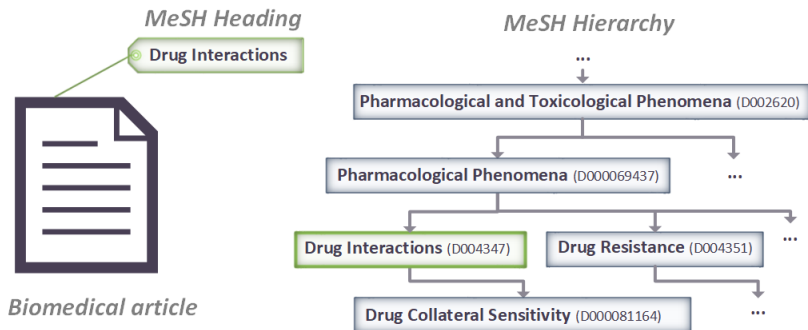
Task A

Participant Teams and Systems



Task A

Semantic indexing example

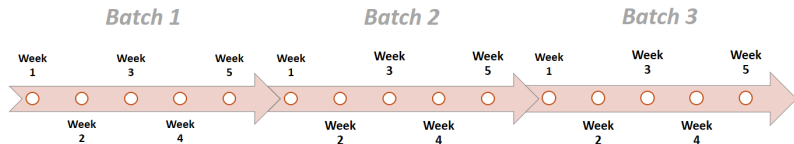


Task A

Structure

Task A: Hierarchical text classification

- ▶ Organizers distribute **new unclassified MEDLINE articles**.
- ▶ Participants have 21 hours to assign **MeSH terms** to the articles.
- ▶ **Evaluation** based on annotations of **MEDLINE curators**.



Task A

Datasets

► Training data

Year	Articles in millions	Labels in thousands	Labels per article
2022	16.2	29.7	12.68
2021	15.6	29.4	12.68
2020	14.9	29.1	12.68
2019	14.2	28.9	12.69
2018	13.5	28.3	12.69
2017	12.8	27.8	12.66
2016	12.2	27.3	12.62
2015	11.8	27.1	12.61
2014	12.6	26.8	12.72
2013	10.9	26.6	12.55

► Test data: Organized into independent batches

- One test set per week: About 6,000 articles
- Five test sets per batch: About 30,000 articles
- Three batches per year: About **150,000** articles

Task A

Evaluation Measures

Flat measures

- ▶ Accuracy (Acc.)
- ▶ Example Based Precision (EBP)
- ▶ Example Based Recall (EBR)
- ▶ Example Based F-Measure (EBF)
- ▶ Macro Precision/Recall/F-Measure (MaP, MaR, MaF)
- ▶ **Micro Precision/Recall/F-Measure** (MiP, MiR, **MiF**)

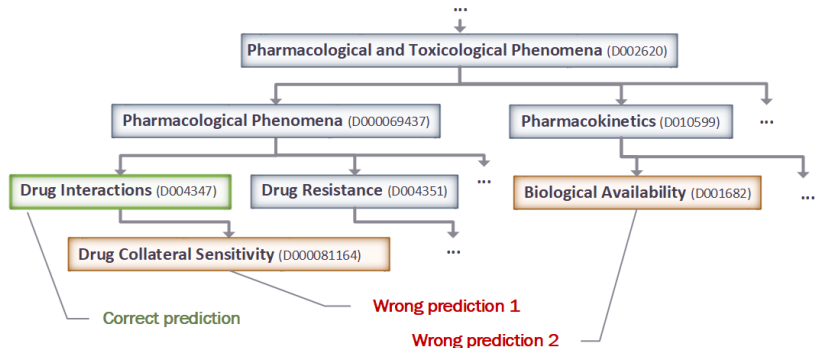
Hierarchical measures

- ▶ Hierarchical Precision (HiP)
- ▶ Hierarchical Recall (HiR)
- ▶ Hierarchical F-Measure (HiF)
- ▶ Lowest Common Ancestor Precision (LCA-P)
- ▶ Lowest Common Ancestor Recall (LCA-R)
- ▶ **Lowest Common Ancestor F-measure** (LCA-F)

A. Kosmopoulos, I. Partalas, E. Gaussier, G. Paliouras and I. Androutsopoulos: Evaluation Measures for Hierarchical Classification: a unified view and novel approaches. Data Mining and Knowledge Discovery, 29:820-865, 2015.

Task A

Semantic indexing errors



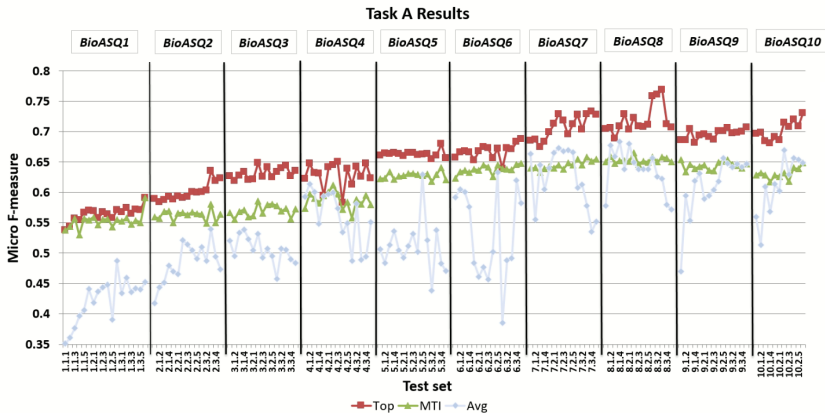
Task A

System approaches

- ▶ **Feature Extraction:** Representing each abstract.
 - ▶ Trend from BOW and tf-idf representations towards neural word and paragraph embeddings.
 - ▶ Also efforts to exploit journal and year information.
- ▶ **Concept Matching:** Finding relevant MeSH labels.
 - ▶ Trend from KNN, SVM and Learning-To-Rank approaches towards Deep Networks and Attention Mechanisms.

Task A

Results



Task A

NLM Transition to Automated Indexing

- ▶ MTI improved by almost 10% during the last 10 years.
 - ▶ largely due to the **adoption of ideas from BioASQ** participant systems.
- ▶ High point: NLM Adopted **fully automated indexing** in 2022.

MEDLINE 2022 Initiative: Transition to Automated Indexing

MEDLINE 2022 Initiative: Transition to Automated Indexing. NLM Tech Bull. 2021 Nov-Dec;(443):e5.

2021 December 01 [posted]

As part of the efforts of the National Library of Medicine (NLM) to transform and accelerate biomedical discovery and improve health and health care, we are transitioning to automated [MeSH](#) indexing of [MEDLINE](#) citations in [PubMed](#). Automated indexing will provide users with timely access to MeSH indexed metadata and allow NLM to scale MeSH indexing for MEDLINE to the volume of published biomedical literature. Human indexers have been and will continue to be involved in the refinement of automated indexing algorithms and will play a significant role in the quality assurance approaches for automated indexing.

In 2018, NLM launched the MEDLINE 2022 initiative, a five-year development plan that aims to ensure that MEDLINE continues to evolve to meet the needs of users in an age of data-driven discovery. A key goal of this initiative involved implementing a range of indexing methods to ensure the timely assignment of MeSH to MEDLINE citations. Based on the successful pilot of automated indexing on a limited scale since 2016, it was determined that fully automated MEDLINE indexing be implemented with quality control, and that human curation and automation be specifically applied to improve the discoverability of chemical and gene information in MEDLINE.

https://www.nlm.nih.gov/pubs/techbull/nd21/nd21_medline_2022.html

Task A

Lessons learnt

- ▶ A task based on a real process (semantic indexing at NLM) may affect the process itself.
- ▶ Unforeseen factors may affect the course of a task (e.g. NLM indexer contracting policies).
- ▶ Being “live” comes at the cost of reduced uniformity in the datasets (e.g. annual changes in MeSH vocabulary).

Task A

Looking into the future

- ▶ Automatic alignment of the different MeSH versions.
- ▶ Other types of semantic annotations (e.g. publication types).
- ▶ Identifying datasets used in literature.

Past BioASQ collaborators

Grateful to ...

- ▶ University of Houston (US)
- ▶ Transinsight GmbH (DE)
- ▶ Université Joseph Fourier (FR)
- ▶ University Leipzig (DE)
- ▶ Université Pierre et Marie Curie Paris 6 (FR)
- ▶ Athens University of Economics and Business – Research Centre (GR)



Sponsors

Grateful for the support by our sponsors

Platinum Sponsor



Silver Sponsors



ATYPON

Past Sponsors



VISEO

Stay Tuned!

Visit www.bioasq.org
Follow [@BioASQ](https://twitter.com/BioASQ)

The **eleventh BioASQ workshop** as part of **BioASQ lab 2023**
in **CLEF 2023** will take place at Thessaloniki, Greece
on **18-21 September 2023!**