ML4MICROBIOME WORK PACKAGE 1
State-of-the-art evaluation and update

Jaak Truu
University of Tartu
Estonia

08.07.2021

Objectives

To continuously evaluate the state-of-the-art ML/statistics methods, and to ensure that every action member is "on the same page" in terms of their robustness and suitability for microbiome research and how well they address the specific challenges in 1.1.1, separately and combined.

Task 1.1: Technology watch

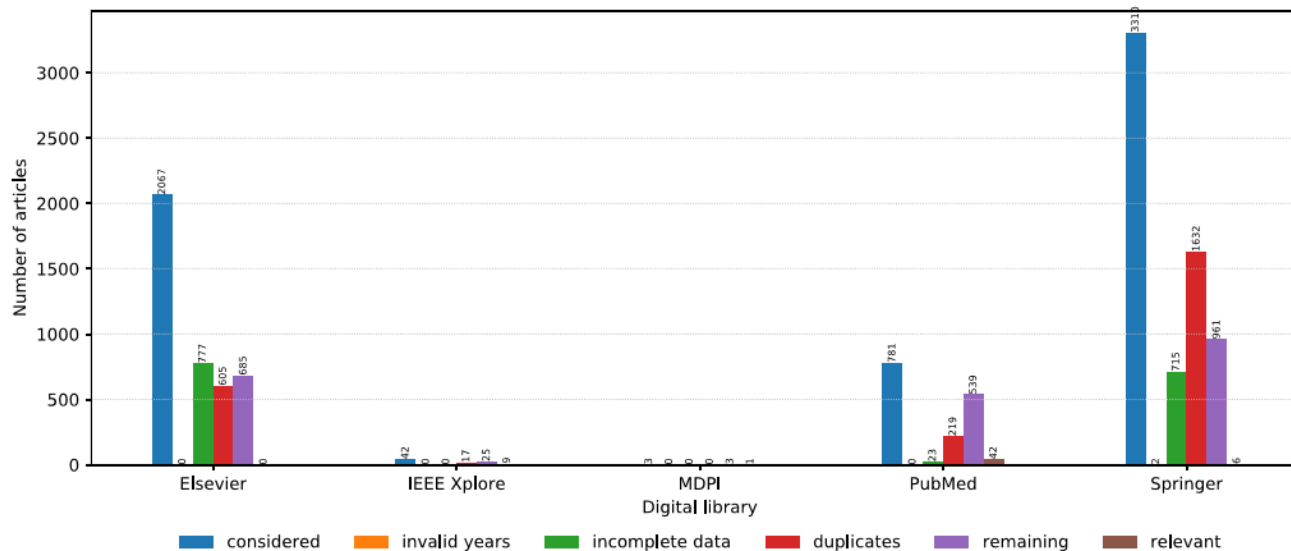Task 1.2: Evaluation of ML/statistics methods currently used in microbiome research

Task 1.3: Define priority areas for novel ML/statistics applications for microbiome data
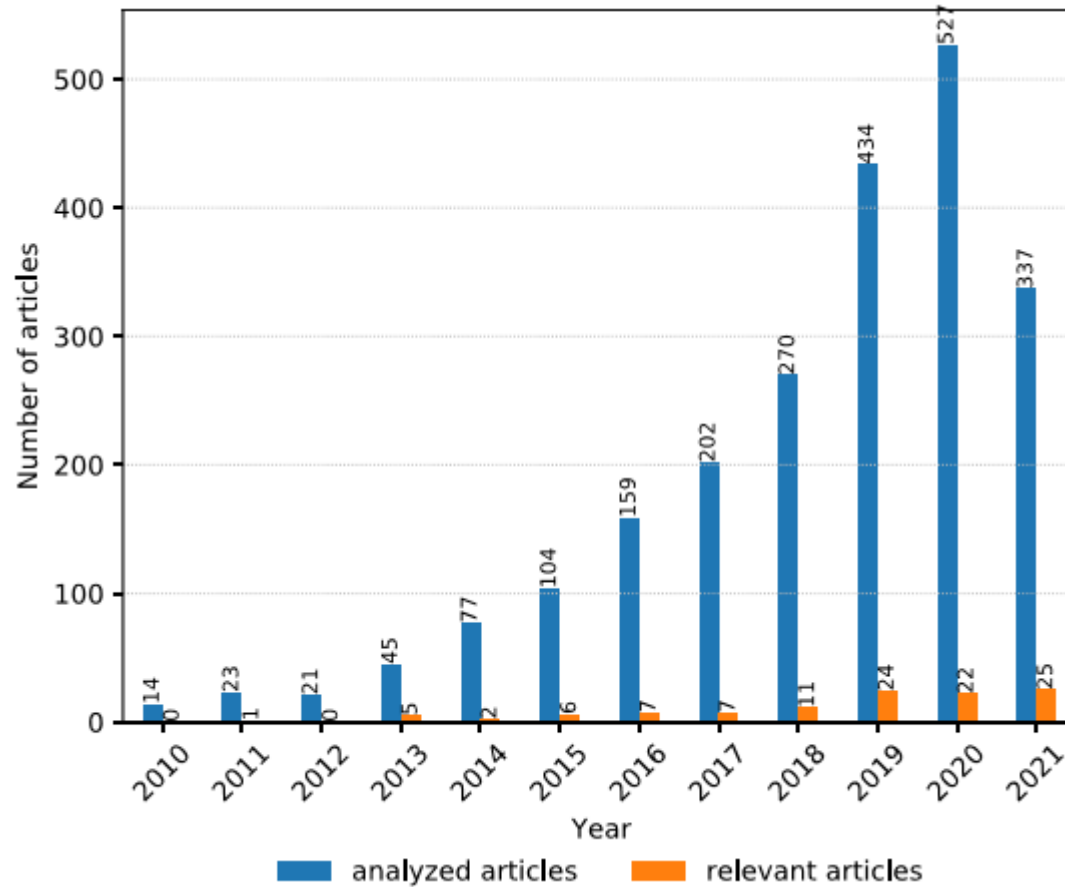
WG1 activities in 2021

- Updated annual report for 2021
  - *Will be delivered in October 2021*
  - *Will be based on revised first annual report*
  - *Compilation of the report will follow the same process as last year*
- Additional task
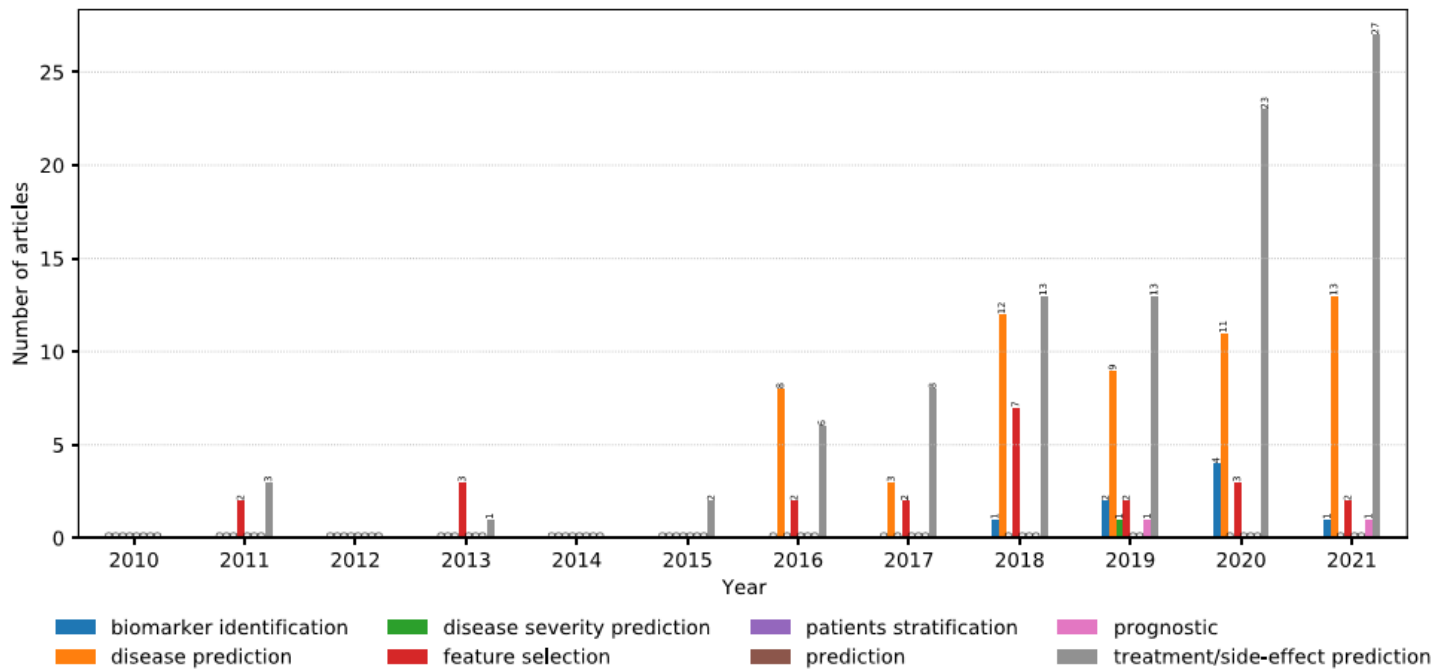 - *Next action/WG1 publication – will be discussed among WG1 members*

Scoping Review

- In 2021, Springer, IEEE, and PubMed were accessed
- In 2021, two additional digital libraries: Elsevier and MDPI were included
- No Oxford Academic Publishers included

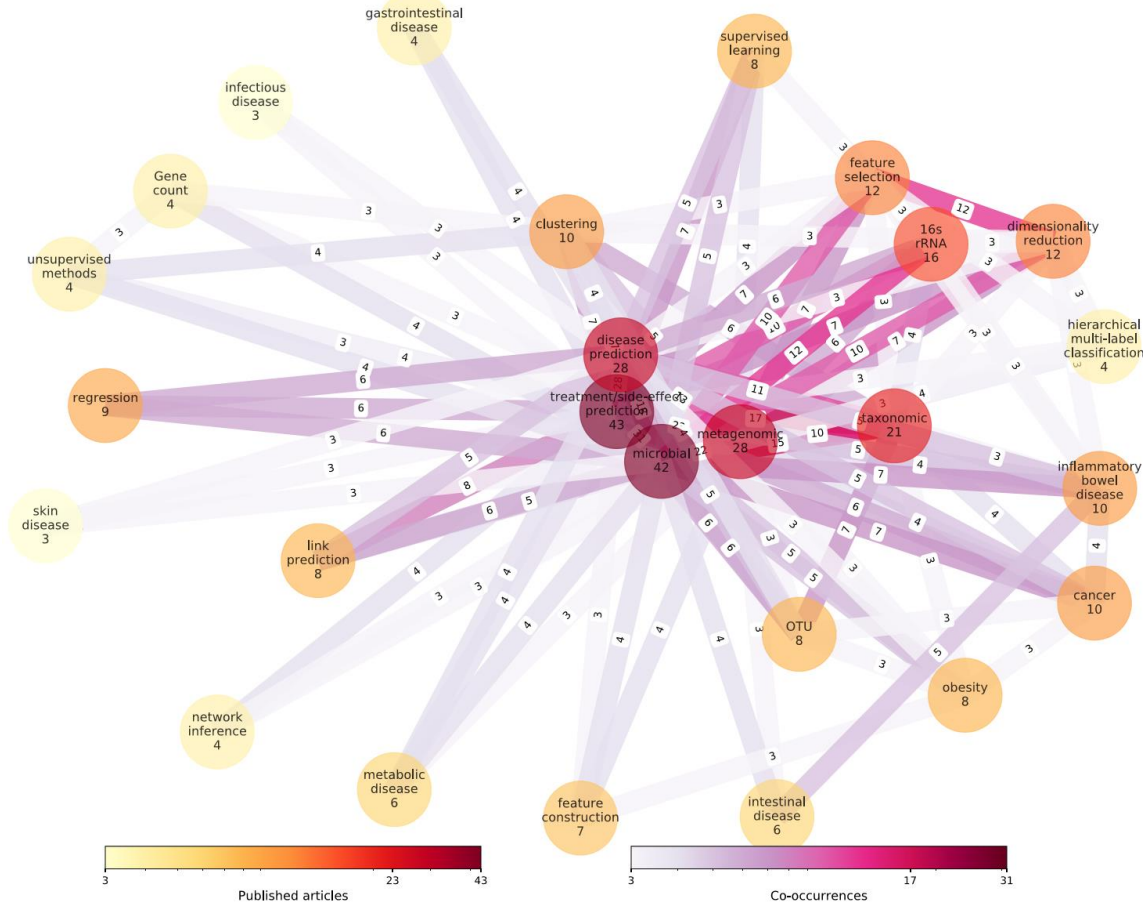NLP toolkit (Zdravevski et al., 2019)

Publication search results

Publication search results

Co-occurrence network

Shown keywords:26 Shown pairs: 109 Minimum co-occurrence in articles: 3 (75-th percentile)

Microbiome research in the source code repositories

Short summary (from September)

- Number of repositories 1855 (old), 2465 (new).
- New repositories 1014.
- Repositories deleted 404.

- Update (October)
  - As of today we have slightly more repositories (2589).
- Updated info on github list
  http://microbiome.przymus.org/
- Next step – identify new (unique) publications for 2021

Next steps

1. Data base curation – Scoping review
2. Data base creation/curation – Github
3. Additional search (OAP)
4. Merging the obtained data bases
5. Updating the data base by action members
6. Report compilation

Other activities

- Next action/WG1 publication – will be discussed among WG1 members
- Presentations
  - *Enrique Carrillo - Machine Learning & Microbiome for Precision Nutrition. GOBLET & EMBnet AGM 2021*
  - *Jaak Truu*

FEMS Conference on Microbiology
in association with **Serbian Society for Microbiology**
Belgrade 2022
30 June - 2 July 2022 • Serbia

Thank you!