



ENTITY EVALUATION TUTORIAL

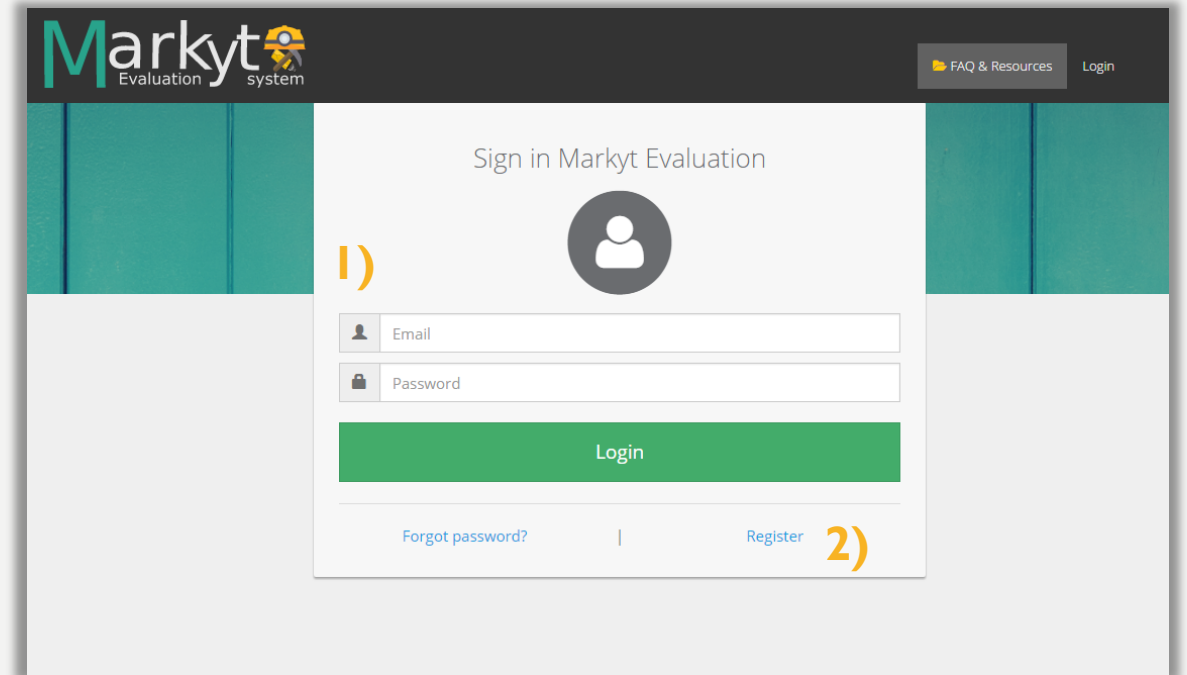
CONTENTS

- Register and login
- User's area
 - Entity file submission
 - Single and multiple entity evaluation
 - Final task submissions
- Contacting Markyt
Evaluation support

LOGIN PANEL

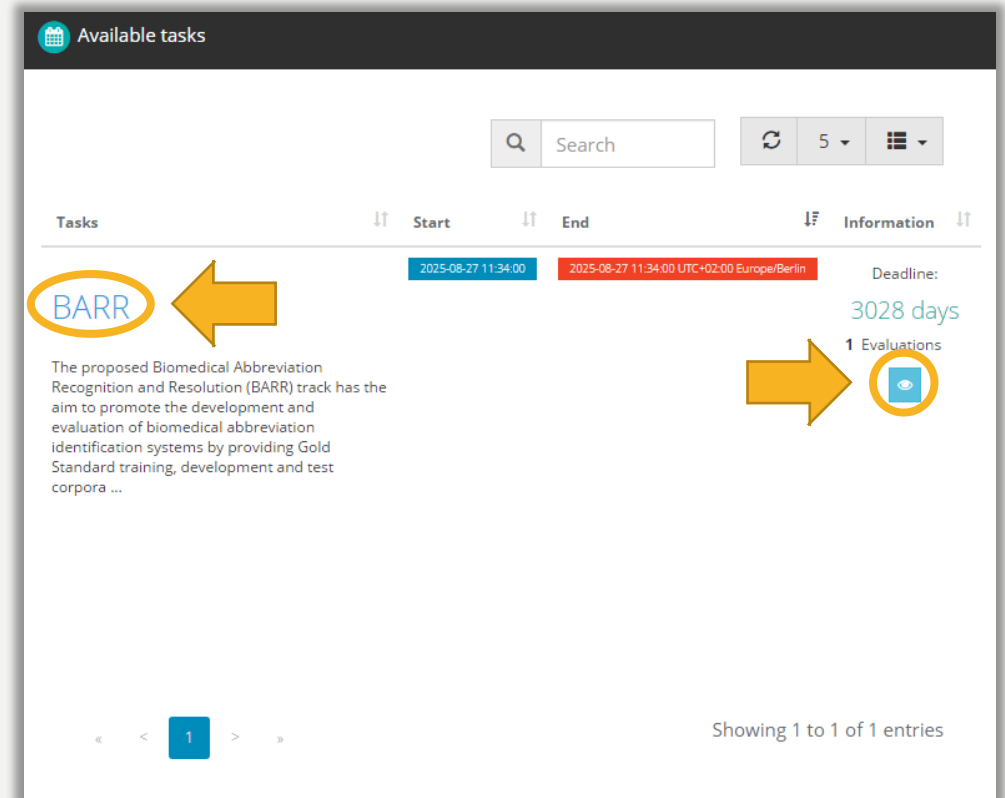
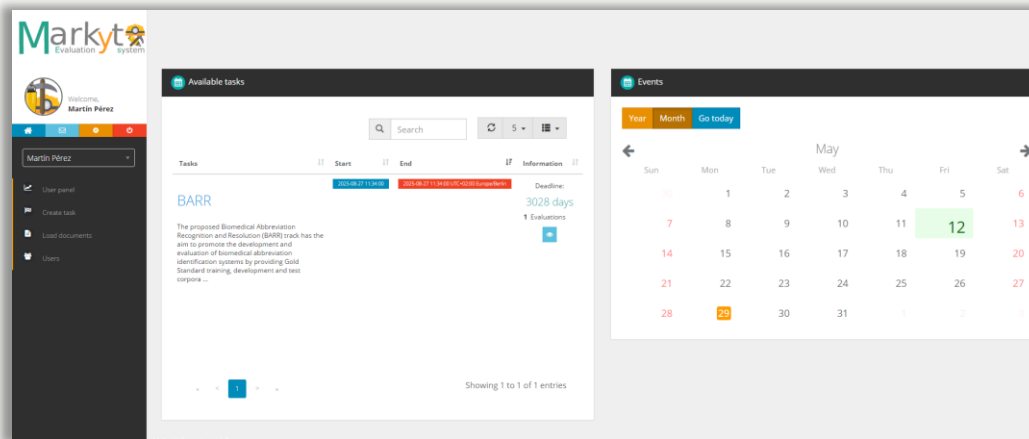
To access the platform go to the participant sign in area:

- 1) Login in to the system with your credentials; or,
- 2) Create a new account.



USER'S MAIN PAGE

After logging in, you are granted access to the available competitions (e.g. BARR task).



TASK VIEW

- For example, this are the general panels of the BARR task.
- Mainly, you can upload entity files to your private repository (1) and run entity evaluations (2).
- You can also check current overall/average performance metrics (3).

3) Metrics

The image shows two side-by-side panels for evaluation metrics. The left panel is titled 'Entity prediction evaluation' and the right panel is titled 'Relation prediction evaluation'. Both panels display a table of metrics including Avg, Min, and Max for Micro and Macro F-scores, along with a Total evaluations count. The Entity panel shows a value of 1 for most metrics, while the Relation panel shows 0.

Entity prediction evaluation			
Avg	Micro	F-score	1
Min	Micro	F-score	1
Max	Micro	F-score	1
Avg	Macro	F-score	1
Min	Macro	F-score	1
Max	Macro	F-score	1
Total evaluations			1

Relation prediction evaluation			
Avg	Micro	F-score	0
Min	Micro	F-score	0
Max	Micro	F-score	0
Avg	Macro	F-score	0
Min	Macro	F-score	0
Max	Macro	F-score	0
Total evaluations			0

1) Upload here your entity files

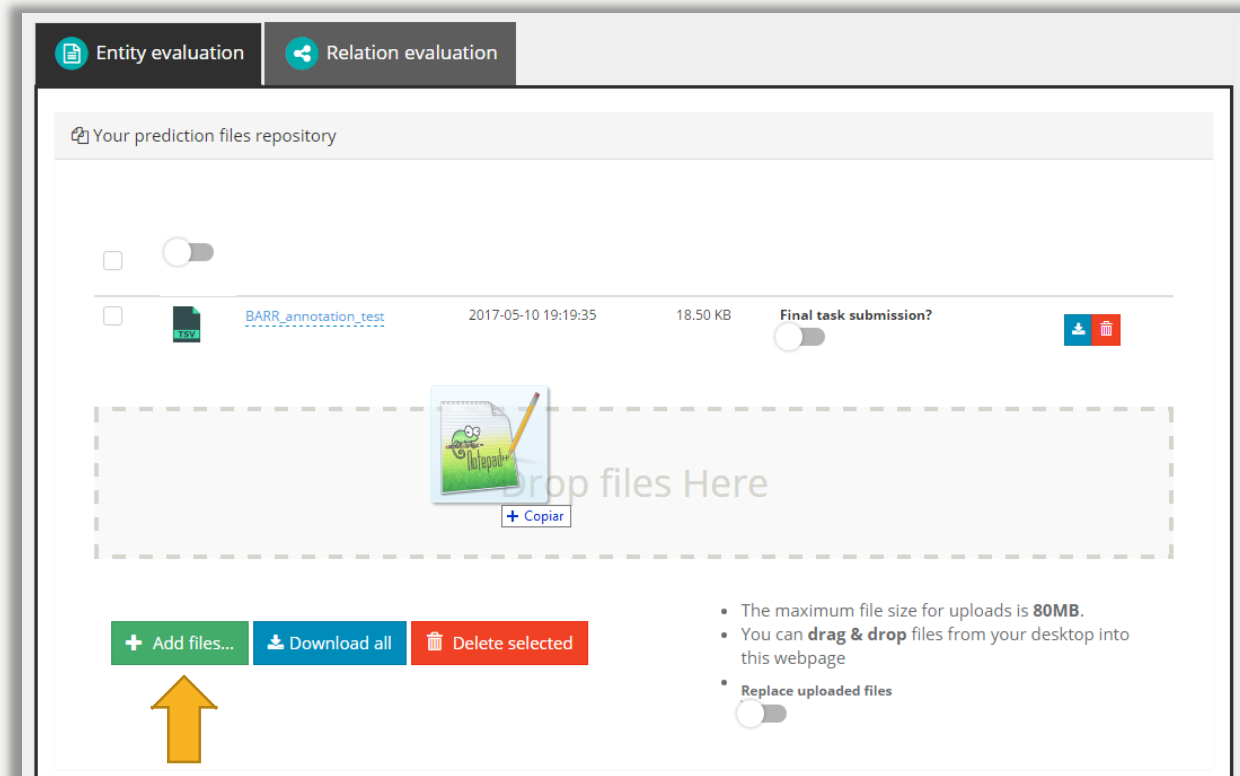
2) Evaluate uploaded entity files here

The image shows the 'Entity evaluation' interface. It includes a section for uploading files to the 'Your prediction files repository' with a table listing files like 'BARR_annotation_test'. Below this is a section for 'Your entity evaluations' with a table showing the status of evaluations, including 'BARR_annotation_test' which is 'SUCCESSFULLY TERMINATED'. The interface also includes buttons for 'Add files...', 'Download all', 'Delete selected', and 'Run evaluation'.

Predictions	Created	F-score	Progress	State	Actions
BARR_annotation_test	2017-05-10 19:19:42	1.000	100%	SUCCESSFULLY TERMINATED	[Icons]

PREDICTION REPOSITORY

- Each participant has a private entity repository.
 - Simply drag files to your browser or press the 'Add files...' button in order to upload files.
- Prediction files must be in TSV format and the encoding should be UTF-8.
 - Examples are provided in [Resources section](#) (inside Markyt Evaluation website).
 - Note that you have to download the example and upload it into the repository before using it (i.e. running any evaluation).



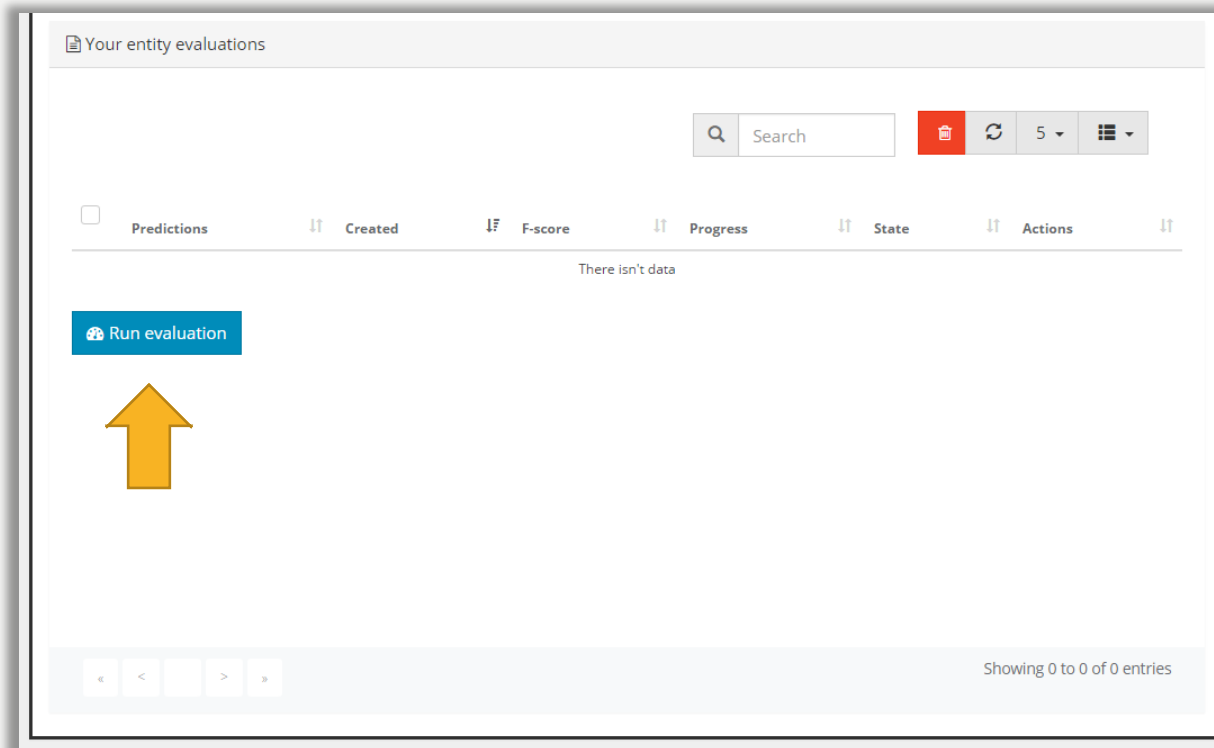
ENTITY PREDICTIONS FILE EXAMPLE

- Download an example [here](#)

```
DOCUMENT_ID→LANGUAGE→SECTION→INIT→END→ANNOTATED_TEXT→TYPECRLF
100005→es→T→63→67→PCEA→GLOBALCRLF
100224→en→A→689→692→IgE→GLOBALCRLF
100224→en→A→679→684→INF-γ→GLOBALCRLF
100224→en→A→480→483→IgE→GLOBALCRLF
100224→en→A→452→457→ELISA→GLOBALCRLF
100224→en→T→29→39→Tyrophagus→CONTEXTUALCRLF
100428→es→A→1946→1948→EM→MULTIPLECRLF
100428→es→A→1820→1822→EM→MULTIPLECRLF
100428→es→A→1571→1573→RM→GLOBALCRLF
100428→es→A→1526→1528→EM→MULTIPLECRLF
100428→es→A→1256→1258→RM→GLOBALCRLF
100428→es→A→1141→1143→RM→GLOBALCRLF
CRLF
```

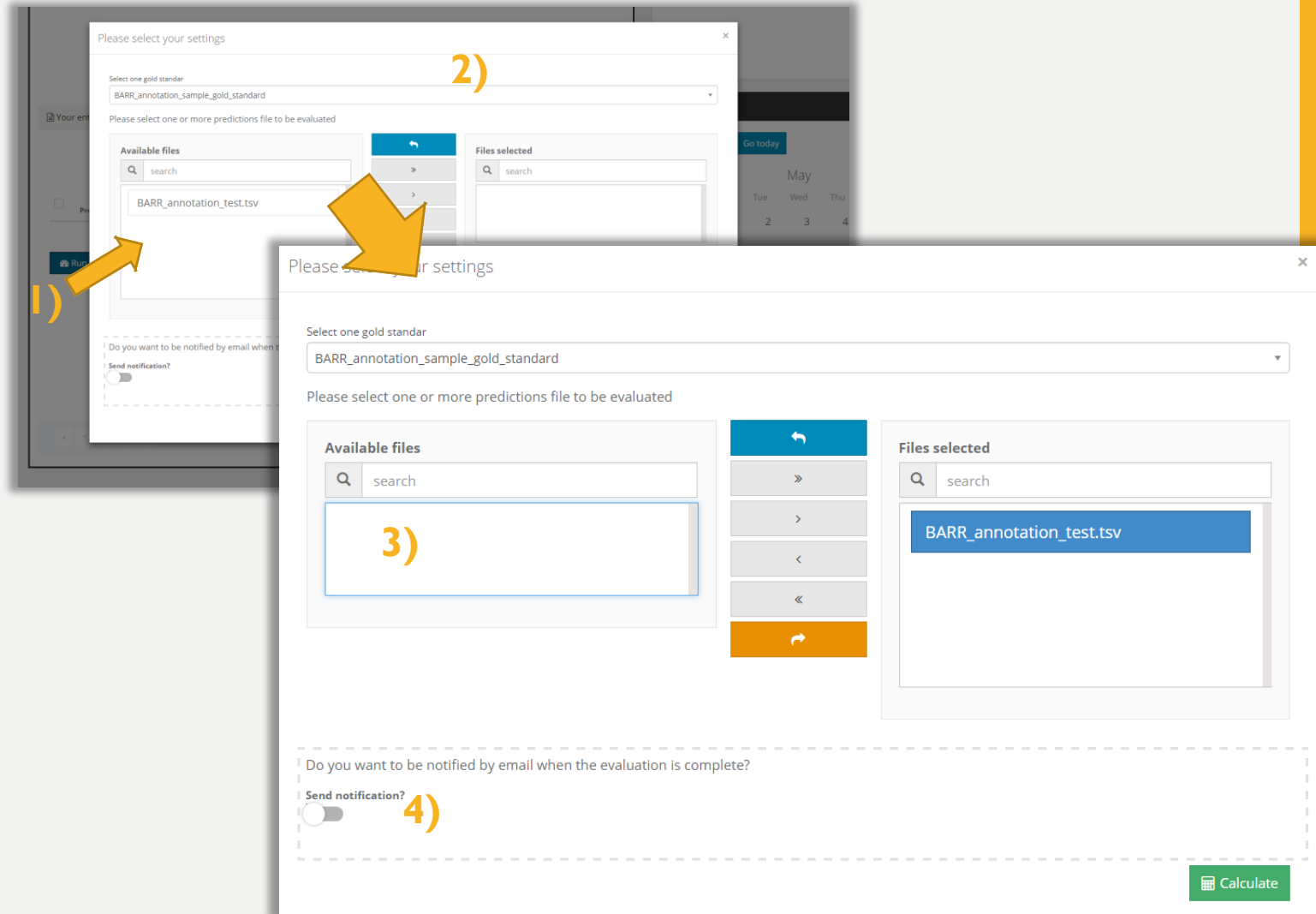
RUN ENTITY EVALUATION (I)

- Click in the “Run evaluation” button to select entity files from your repository and run evaluations.

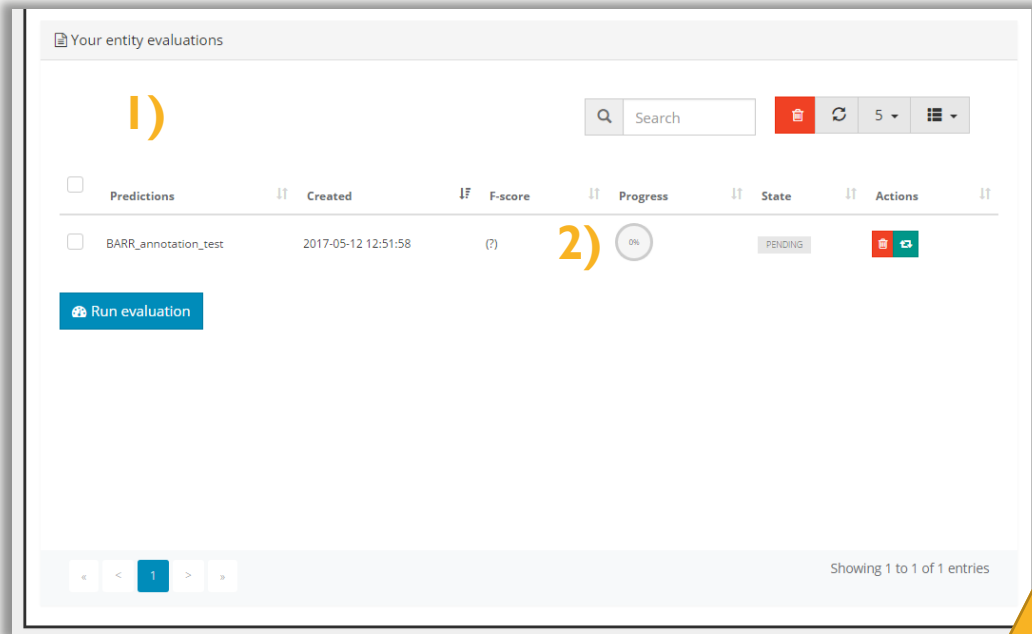


RUN ENTITY EVALUATION (II)

- Once you click “Run evaluation” (1) a new window opens (2).
- Select the entity files that you wish to evaluate by clicking on their name (3).
 - Here, the file BARR_annotation_test.tsv was selected.
- You can also choose to receive an email notification when the evaluation is completed (4).
- Click on “Calculate” to run the evaluation.

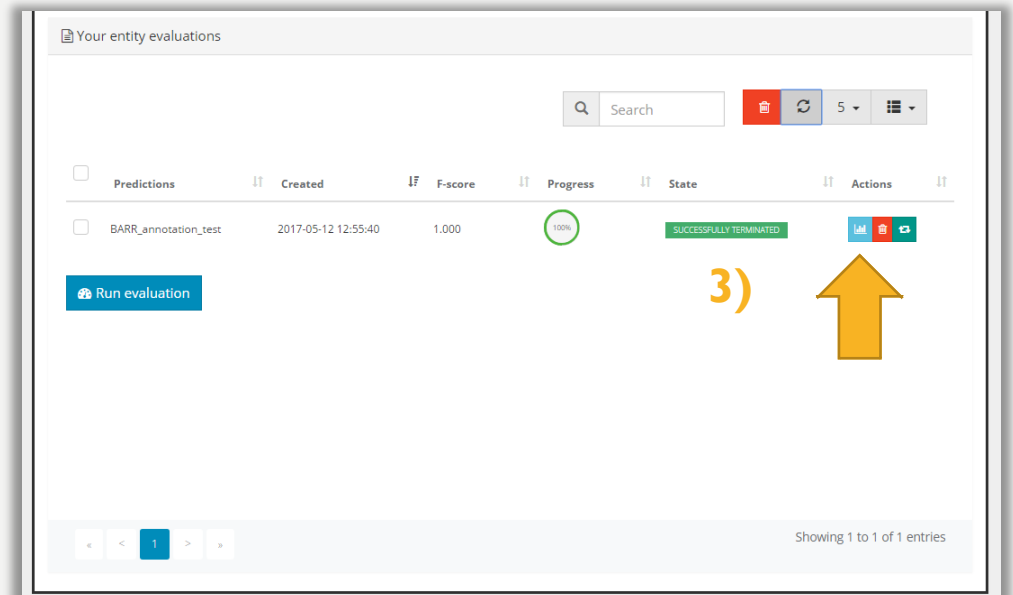


RUN ENTITY EVALUATION (III)



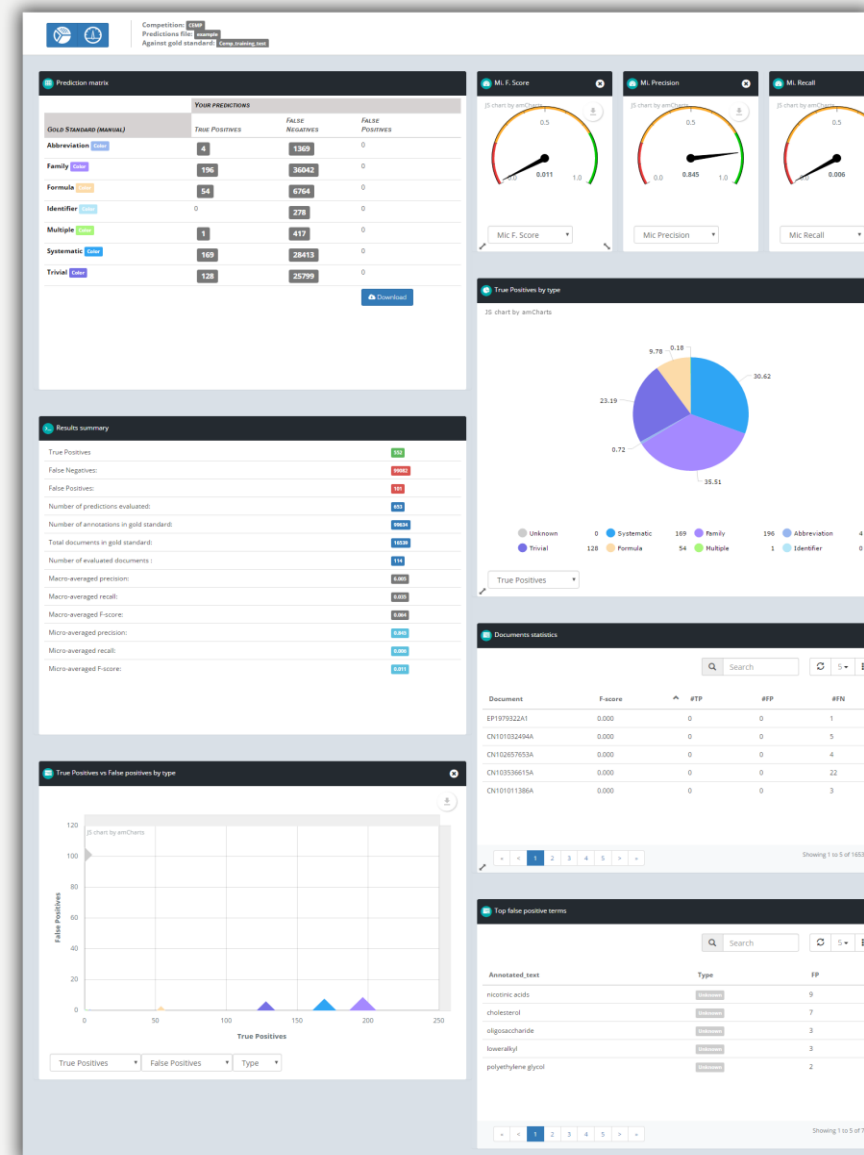
- When the evaluation ends you may click on the blue button to visualize evaluation statistics (3).

- Your evaluation request will be inserted into the Markyt Evaluation queue (1).
- You can check its current status at any point (2).



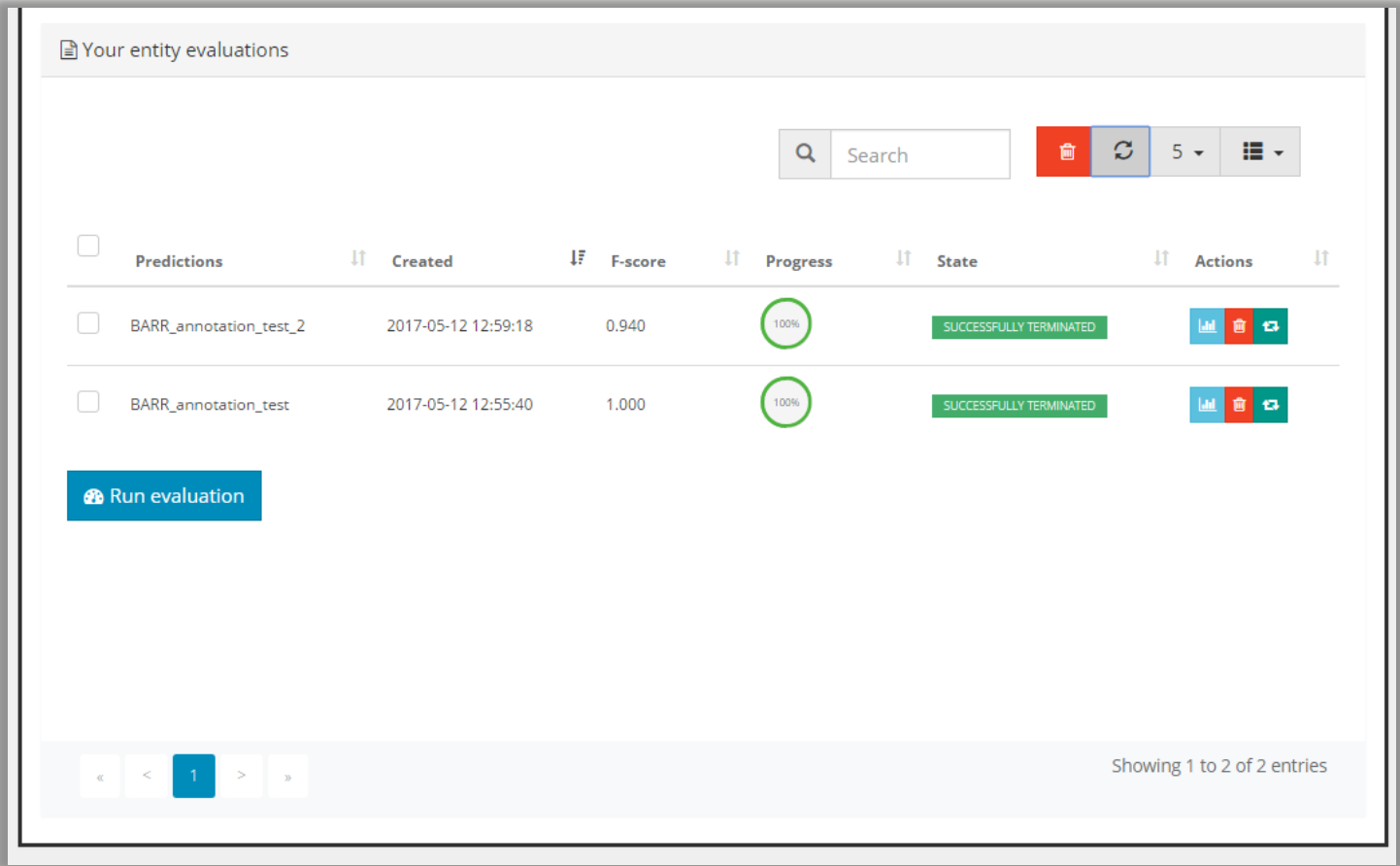
EVALUATION RESULT VIEW

- This is an example of the general view of entity performance report.
- Typically, this report details performance in terms of F-score, precision, and recall (micro- and macro-averaged), and lists top false positives and negatives (by type and document).



COMPARE EVALUATIONS(I)

- Once you have evaluated two or more entity files you may compare results. That is, evaluate whether your prediction ability has improved.
- For example, create a local copy of the prediction file that you previously uploaded and randomly remove some prediction entries from the file.
- Upload this new prediction file and evaluate it.



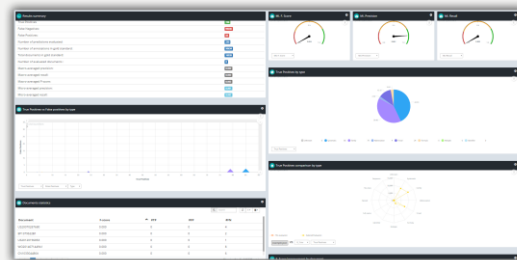
The screenshot displays a web interface titled "Your entity evaluations". It features a search bar and a table with columns for Predictions, Created, F-score, Progress, State, and Actions. Two evaluation entries are listed, both showing 100% progress and a "SUCCESSFULLY TERMINATED" state. A "Run evaluation" button is located below the table. The footer indicates "Showing 1 to 2 of 2 entries".

☐	Predictions	Created	F-score	Progress	State	Actions
☐	BARR_annotation_test_2	2017-05-12 12:59:18	0.940	100%	SUCCESSFULLY TERMINATED	  
☐	BARR_annotation_test	2017-05-12 12:55:40	1.000	100%	SUCCESSFULLY TERMINATED	  

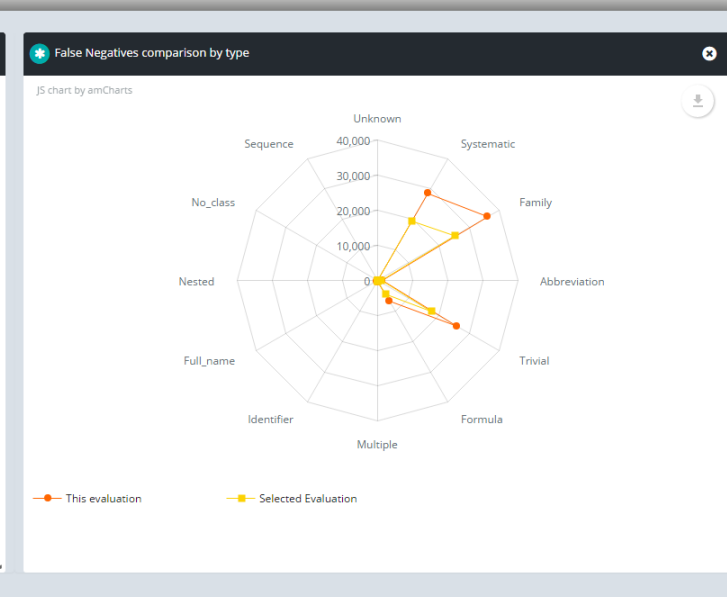
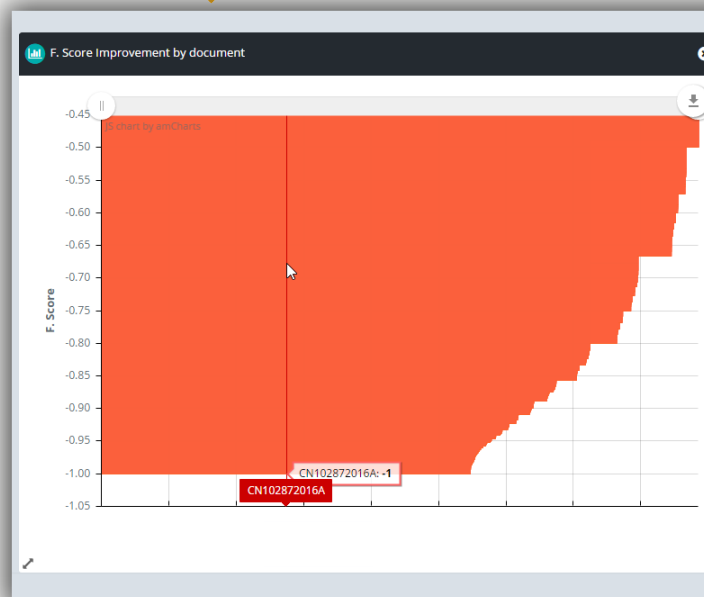
 Run evaluation

Showing 1 to 2 of 2 entries

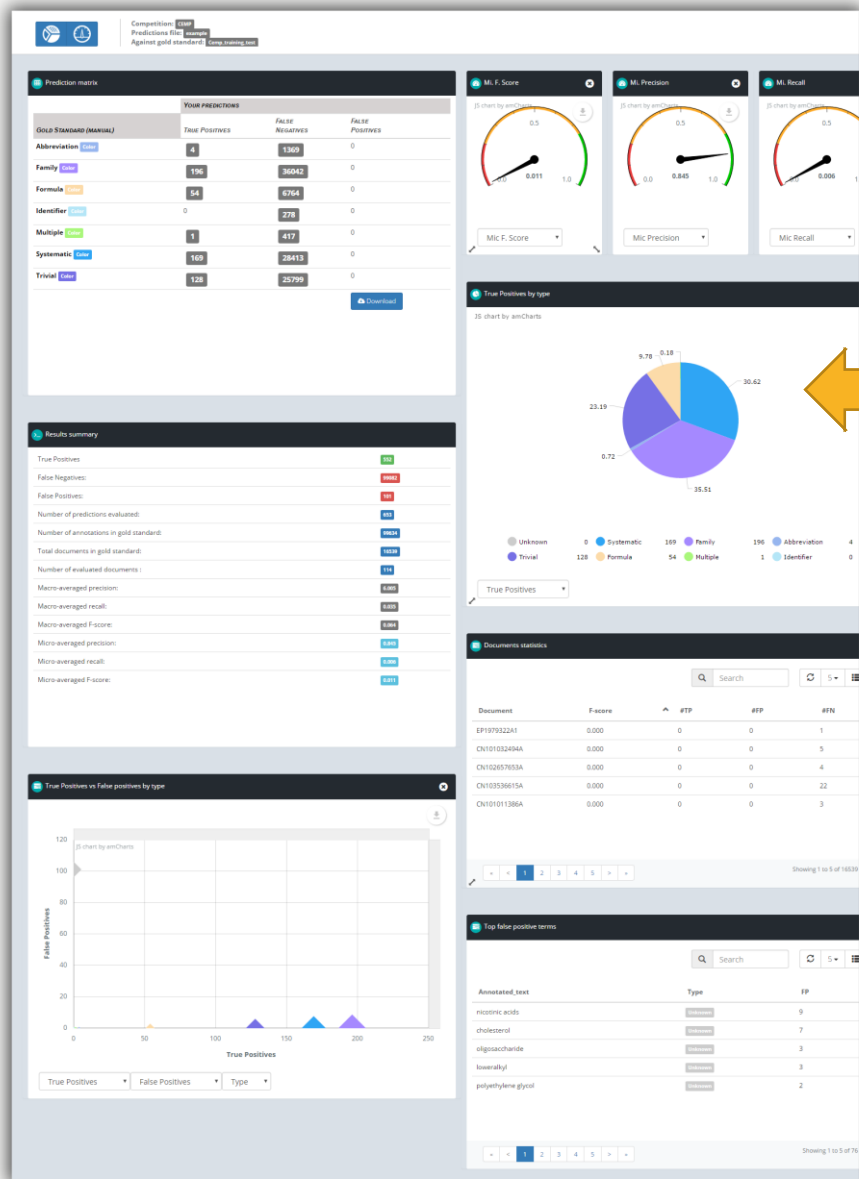
COMPARE EVALUATIONS (II)



- By accessing the prediction report (as before), you have access to current performance as well as comparative figures.
- Some panels enable the comparison of the present statistics against the statistics obtained by any other entity file previously evaluated (e.g. check the entity evolution panel).

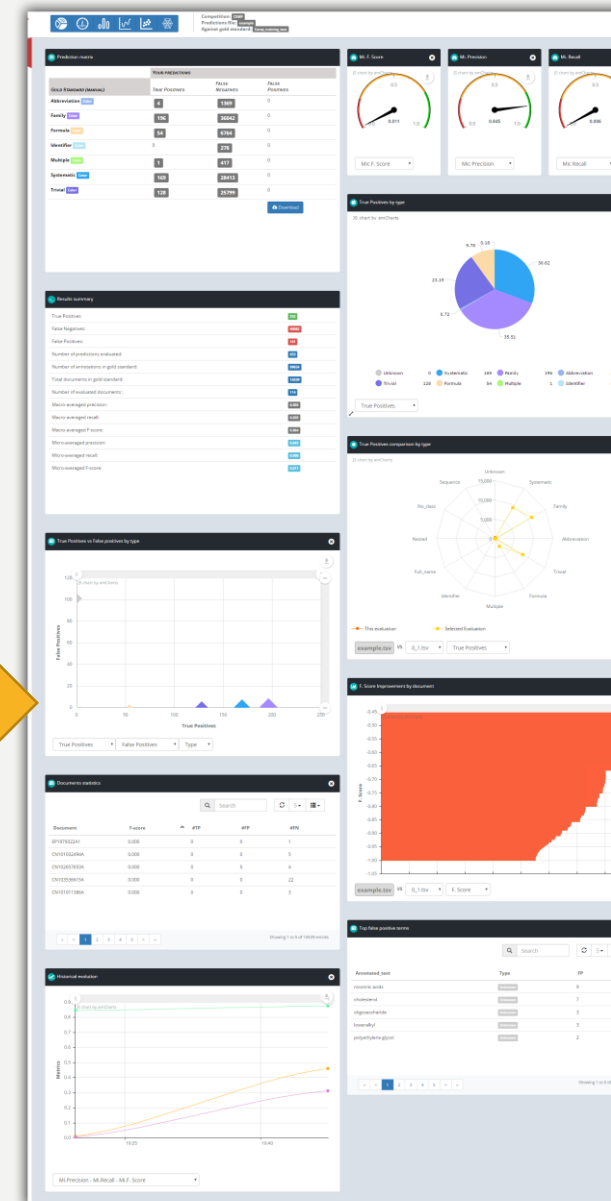


COMPARE EVALUATIONS (III)



These are two general views of the evaluation report:

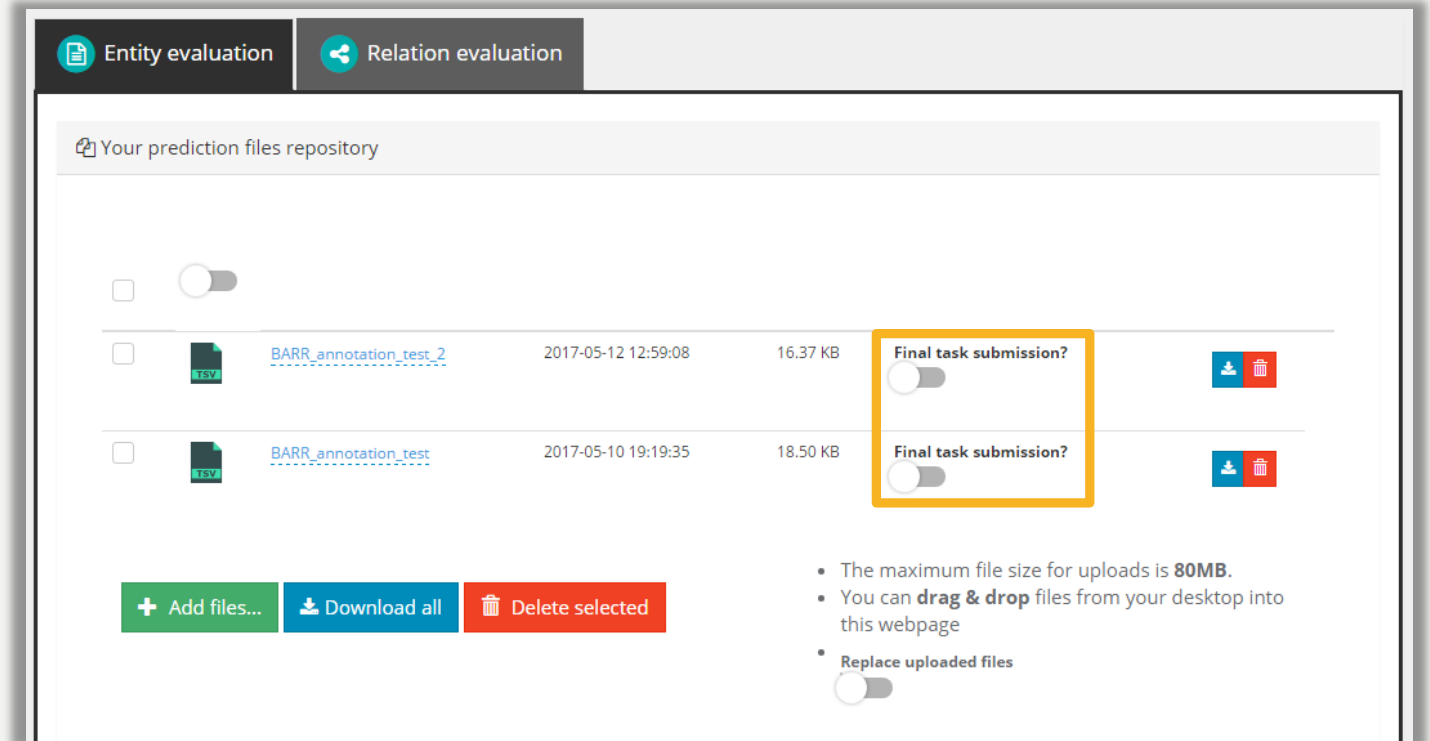
- the left one does not allow comparison while the right one enables entity comparison;
- comparisons are enabled automatically (once there are two or more evaluations you have new graphical artefacts).



FINAL TASK SUBMISSION

From all the uploaded entity files you can choose the final task submissions:

- in the repository panel, switch the submit task button (slide the button to the right) of the file;
- BARR tasks allow a maximum of 5 entity file submissions per participant.



CONTACT US

- All resources are available at <http://www.markyt.org/mmes/pages/resources>
- Contact us by clicking on the button on your account menu or the link at the bottom of the page.
- Help us make Markyt Evaluation functionalities as useful and intuitive as possible!

