



At BioCreative-CHEMDNER Challenge

FREE ANNOTATION SOFTWARE

NEXT GENERATION COMPUTER SYSTEMS GROUP (SING)
DEPARTMENT OF COMPUTER SCIENCE, UNIVERSITY OF VIGO


PREDICTION ANALYSIS TOOL

<http://www.markyt.org/biocreative/analysis>

1. Markyt credentials

When you register as participant of the challenge ([register for track 2](#) of BioCreative V), you will receive the below email with your credentials to use Markyt. If you do not receive this email, please contact us.

Markyt



Exciting news!


Dear CHEMDNER participant,

We're delighted to introduce our brand new feature to help you analyze your predictions with Markyt application. Using the credentials presented below you will be able to access the Markyt CHEMDNER prediction analysis tool in this url <http://markyt.org/biocreative/analysis>


Your username is this email **email@example.com** <- This will be your participant username

Your password is: **1af543bc** <- This will be your participant code

Note that the use of this tool is entirely optional.



You can see a thorough analysis of your predictions in a visually appealing way!



Your results were safe from spies! we will be encrypted your results and only you can see them through the tool.

Please do not reply to this mail.

© Markyt 2015

2. .TSV file structure

The .tsv file with predictions to be uploaded for evaluation should follow the structure specified by the competition, i.e.:

- ❖ Patent identifier.
- ❖ Offset string consisting in a triplet joined by the ':' character. You have to provide the text type (T: Title, A:Abstract), the start offset and the end offset.
- ❖ The rank of the entity returned for this document.
- ❖ A confidence score.
- ❖ The string of the entity mention.

```
CN103371975A A:271:274 1 0.99 RGD
CN103371975A A:276:306 2 0.98989 Arginine
US20090312385 A:100:112 1 0.99 CB2
US20090312385 T:0:11 2 0.98989 Cannabinoid
WO2014144455A1 A:616:621 1 0.99 CARM1
WO2014144455A1 T:53:58 2 0.98989 carm1
```

3. Load predictions

You may load your predictions at <http://markyt.org/biocreative/analysis>. Follow the instructions in the picture below:

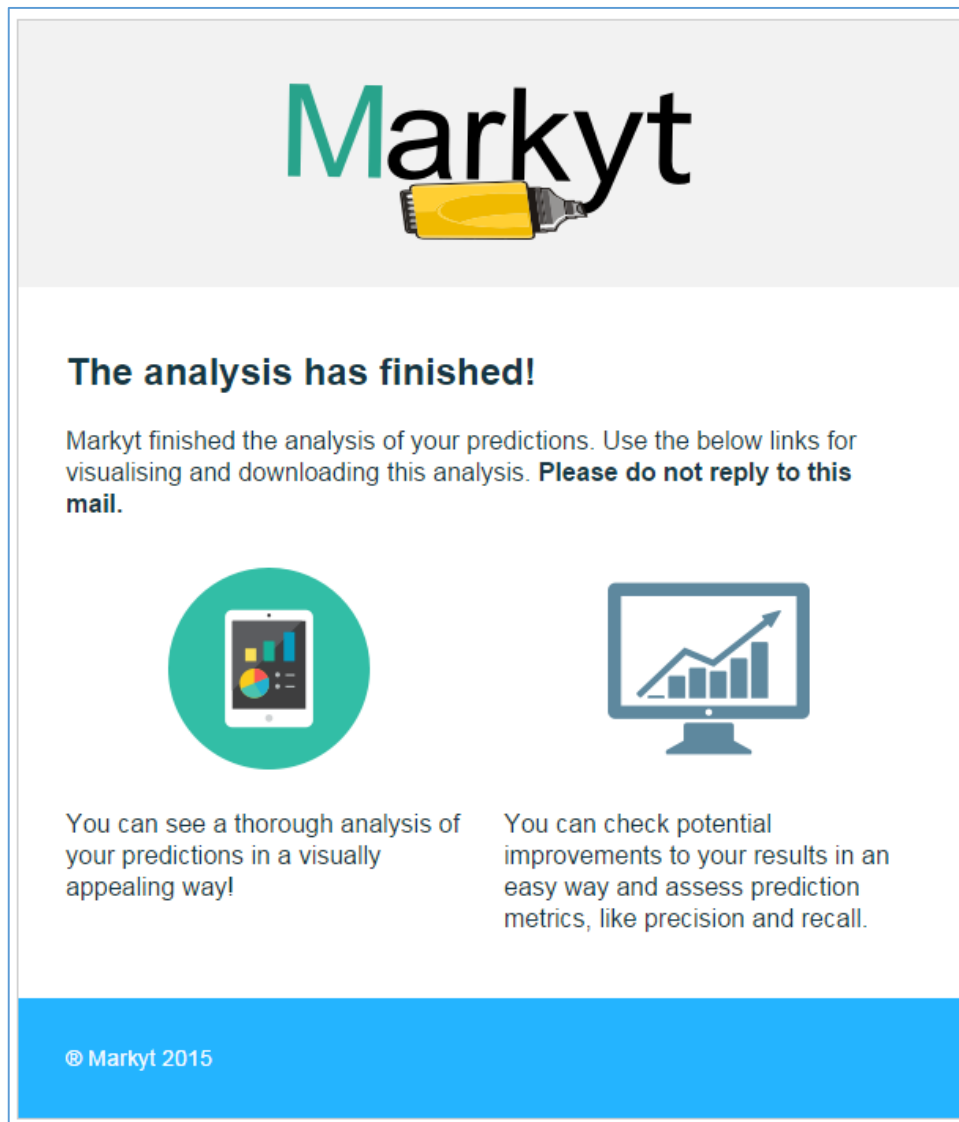
The screenshot shows a web form for loading predictions. At the top right, there are links for 'About | Help | contact'. The form contains several input fields and buttons, each highlighted with a red box and accompanied by a numbered instruction:

- Email***: A text input field containing 'email@example.com'. Instruction: '1. Put your participant email here'.
- Your participant code***: A text input field containing '1af543bc'. Instruction: '2. Put your participant code here'.
- Remember your credentials?**: A checkbox labeled 'Yes' which is currently unchecked. Instruction: '3. Select 'Load predictions' tab'.
- Load predictions**: A button with a folder icon, highlighted in red. Instruction: '3. Select 'Load predictions' tab'.
- Load analysis**: A button with a bar chart icon.
- Project**: A dropdown menu showing 'Not found'. Instruction: '4. Select the project to make the predictions (If your credentials don't have a project, this message will be shown)'.
- Load team prediction**: A section containing a file input field with 'file:\prediction.tsv' and a blue 'Select file' button with a cloud icon. Instruction: '5. Select your .tsv file with data'.
- Submit**: A green button at the bottom left. Instruction: '6. Click here to start the prediction!'.

A light blue informational box in the center of the form reads: 'Here you can submit your **predictions.tsv** as many runs as you want, but only one at the time (check the predictions file format [here](#)). For security reasons, the results will be sent to you by email.'

4. MARKYT results

When the evaluation is finished, you will receive an email like the one shown in the picture below. The email will have an attached file with *.mtmp* extension. This file contains the evaluation of your system's predictions, which you can then visualize in Markyt. Notice that these results will be private, i.e. only you will receive the email and will be able to see the results.



5. Load analysis

You may load the .mtmp file at <http://markyt.org/biocreative/analysis>. Follow the instructions provided in the picture below.

[About](#) | [Help](#) | [contact](#)

Email*
email@example.com 1. Put your participant email here

Your participant code*
1af543bc 2. Put your participant code here

Remember your credentials?
 Yes

Load predictions Load analysis 3. Select 'Load analysis' tab

Here you can load your analysis results (**file.mtmp**).

Load analysis results
file:\\demo_cemp.mtmp
Select file 4. Select your .mtmp to analyze

5. Click here to start the analysis!

6. Check analysis

Here is an example of the prediction analysis. You will have general counts of matches and mismatches and performance metrics. You may also explore the lists of mismatches per annotation type and document.

[About](#) | [Help](#) | [contact](#)

Chemdner cemp sample

Results by type

Your predictions vs Gold Standard (manual)

Prediction result details

	You	
<i>GOLD STANDARD (MANUAL)</i>	<i>TRUE POSITIVES</i>	<i>FALSE NEGATIVES</i>
Systematic Color	289	91
Trivial Color	277	40
Family Color	214	60
Multiple Color	9	9
Formula Color	119	45
Abbreviation Color	20	5
Identifier Color	2	0
<i>FALSE POSITIVES</i>	220	Download

Micro and Macro results

Micro-averaged results		Macro-averaged results	
Precision	0.81	Precision	0.79
Recall	0.79	Recall	0.8
F-score	0.8	F-score	0.79

Resume of data

Predictions

True Positives	930
False Negatives:	250
False Positives:	220
Number of predicted terms:	1150
Number of Golden predictions:	1180
Evaluated Documents	165
Not in golden documents	0

Additional information

Team

- ❖ Dr. Anália Lourenço
- ❖ Dr. Florentino Fdez-Riverola
- ❖ Martín Pérez Pérez
- ❖ Gael Pérez Rodríguez

Bibliographic Reference

- ❖ Pérez-Pérez M, Glez-Peña D, Fdez-Riverola F, Lourenço A (2014) Marky: A tool supporting annotation consistency in multi-user and iterative document annotation projects. *Comput Methods Programs Biomed* 118:242–251. doi: 10.1016/j.cmpb.2014.11.005.