

MT-EQuAl: a Toolkit for Human Assessment of Machine Translation Output

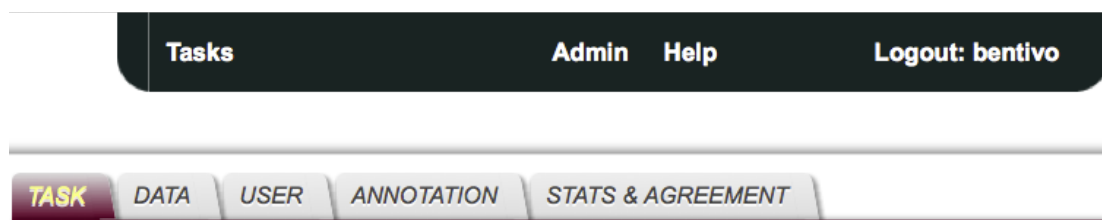
PROJECT MANAGEMENT INSTRUCTIONS

Luisa Bentivogli, Christian Girardi
FBK, Trento - Italy
{bentivo,cgirardi}@fbk.eu

Accessing the project management interface

To work with the MT-EQuAl tool, you must use one of the supported browsers: *Chrome, Safari, Firefox*.

To access the project management interface, you have to enter your Username and Password. Once you have logged in, you have to click on the “Admin” tab. The project management interface is composed of five tabs: TASK, DATA, USER, ANNOTATION, STATS & AGREEMENT.



1. TASK Tab

In this tab the project manager creates a new task. The task type must be selected (quality rating, error annotation, word alignment) and the title and description must be inserted in the corresponding fields. Specific instructions for annotators can also be added, and they will be displayed in the annotation interface. Then, it is possible to set the specific features of the task. The manager can choose if the MT outputs are presented for annotation always in the same order or randomized. Moreover, s/he can choose to adopt a default task setting or to customize it according to specific annotation needs.

TASK

DATA

USER

ANNOTATION

STATS & AGREEMENT

Task type:

Error Annotation

Task name:

Errors_EN_IT

(es. TEST_Errors_EN-AR-ZH)

Short description:

Identification of transation errors in MT output

Instructions:

enable HTML editor

In this task you are presented with a source sentence and 2 automatic translations. Since the source sentence is presented in isolation, a reference human translation is also given with the only purpose of better understanding the source text if necessary.

For each automatic translation, you are asked to

Show systems output randomly:

☐

Task customization?

Use default setting?

value (integer)	label	color
0	No errors	C8E8B8
1	Too many errors	F0D3A8
2	Reordering errors	69B4FF
3	Wrong lexical choice	B296FF
4	Missing word(s)	E9FF24
5	Superfluous word(s)	91F0FF
6	Morphology errors	FF9EFF
7	Casing and punctuation err	A6445B
		FFFFFF

Create

Quality Rating tasks

- F_Quality_EN-AR
- F_Quality_EN-RU
- F_Quality_EN-ZH
- Quality1_EN-IT
- Quality2_EN-IT
- Quality_EN-DE
- test_Quality_EN-IT

Error Annotation tasks

- Errors10_EN-IT
- Errors11_EN-IT
- Errors12_EN-IT
- Errors13_EN-IT
- Errors14_EN-IT
- Errors15_EN-IT
- Errors1_EN-IT
- Errors2_EN-DE
- Errors2_EN-IT
- Errors3_EN-DE
- Errors3_EN-IT
- Errors4_EN-DE
- Errors4_EN-IT
- Errors5_EN-IT
- Errors6_EN-IT
- Errors7_EN-IT
- Errors8_EN-IT
- Errors9_EN-IT
- Errors_EN-DE
- F_Errors_EN-AR
- F_Errors_EN-RU
- F_Errors_EN-ZH

Quality Rating Task. The default setting is a 5-point rating scale. It is possible to customize it by choosing the number of points in the rating scale. Points can be added and deleted by clicking on the “+” and “-” buttons. For each point added to the scale:

- the “value” field must be an integer and corresponds to the value associated to that point in the annotation export file.
- the “label” field must be a number (typically from 0 or 1 to n), and correspond to the number that appears in the rating scale in the annotation interface.
- in the “color” field it is possible to set the color displayed for that point in the scale.

Error Annotation Task. The default error typology includes 6 classes: reordering errors, wrong lexical choices, missing word(s), superfluous word(s), morphology errors, casing and punctuation errors. It is possible to customize the task by choosing the number and type of errors to annotate. For each error type added to the personalized tagset:

- the “value” field must be an integer and corresponds to the value associated to that error in the annotation export file.
- the “label” field must be filled with the name of the error, which is displayed in the error menu in the annotation interface.




















































































































- in the “color” field it is possible to set the color for the error: in the annotation interface, the word(s) annotated with errors are displayed underlined with the color associated to the error.

Word Alignment. The default setting for the word alignment task allows to annotate 2 alignment types, distinguishing between "possible" and "sure" alignments. It is possible to set different types of alignment to be annotated. For each alignment type added to the personalized task:

- the “value” field must be an integer and corresponds to the value associated to that alignment type in the annotation export file.
- the “label” field must be filled with the name of the alignment type, which is displayed in the annotation interface.
- in the “color” field it is possible to set the color for the error: in the annotation interface, at each click in the matrix cell the color of each given alignment type is displayed following the order set for the task.

2. DATA Tab

In this tab, a table summarizing all the data stored in the database for each task is displayed.

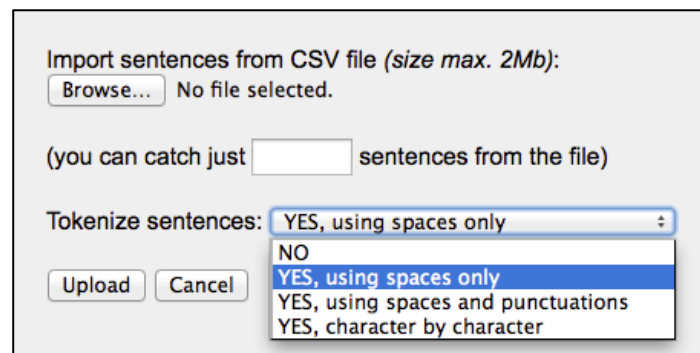
TASK	DATA	USER	ANNOTATION	STATS & AGREEMENT			
Task name	source	reference	sys1	sys2	sys3	sys4	sys5
F_Quality_EN-AR	1479  	1479  	1479  	1479  	1479  		
F_Quality_EN-RU	1479  	1479  	1479  	1479  	1479  		
F_Quality_EN-ZH	1479  	1479  	1479  	1479  	1479  		
Quality1_EN-IT	278  	278  	278  	278  	278  		
Quality2_EN-IT	308  	308  	308  	308  	308  		
Quality_EN-DE	96  	96  	96  	96  	96  	96  	
test_Quality_EN-IT	500  	500  	500  	500  	500  		
Errors10_EN-IT	186  		186  	186  			
Errors11_EN-IT	190  		190  	190  			
Errors12_EN-IT	209  		209  	209  			

By clicking on the “+” button of each cell, the project manager can import the files for the task. Different files are taken in input: source text, reference translation, and one file for each of the MT outputs to be evaluated. This allows the evaluation of single systems as well as the comparison of multiple systems¹.

¹ The input files must be in csv format (UTF-8 encoded). Each row in the file must contain one evaluation item, typically one sentence. For each row, the following info must be specified, separated by a tabular space: sentence ID; language (en, ar, it, zh,...); sentence text. Note that the source and target languages must be declared in the csv, so that the tool can apply the most suitable text visualization (e.g. the text can be displayed left to right and viceversa).

When the “+” button is clicked, a window for uploading the file appears. It is possible to upload the full file or to select only a subset of sentences. It is also possible to decide if and how to tokenize the sentences.

Note that for the error annotation and word alignment tasks, the MT output sentences must be tokenized. The tool offers three types of default tokenization but it is also possible to adopt a personalized tokenization. In that case, you can apply your tokenization to the text and separate the tokens using spaces. Then you can upload the file and choose “YES, using spaces only”.



3. USER Tab

In this tab the project manager can create accounts for users. In addition to the personal data of the user, a user password must be inserted, which will be used by the user to access the interface. Three different types of users can be created, and different permissions are associated by default to each user type:

- *admin*: administrator, who has all permissions
- *advisor*: the advisor of a task can access the “Annotation” tab in the project management interface in order to monitor the annotators’ work. The advisor cannot modify the annotations
- *annotator*: the annotator can only access the annotation interface and see the task(s) s/he has been assigned to.

By selecting one or more tasks in the “Tasks” menu, the user is associated to the tasks and will be able to see the task data and annotate them.

It is also possible to activate or deactivate the user account. An active user can login into the interface and perform annotation, while a non-active user cannot login. Typically it is useful to disable users when an annotation task is concluded. In this way, all the annotations remain stored in the database but cannot be modified anymore.

All the created users are listed in the right side of the interface. Active users are displayed in black, while non-active ones are displayed in struck-through light gray characters. By clicking on the name of a user it is possible to modify his/her profile and assignments.

TASK

DATA

USER

ANNOTATION

STATS & AGREEMENT

Full name:

User name:

E-mail:

Password:

User type:

Annotator

Annotator

Advisor

Admin

Team:

Notes:

Tasks:

(hold CTRL for multiple selection)

all

F_Quality_EN-AR

F_Quality_EN-RU

F_Quality_EN-ZH

Quality1_EN-IT

Quality2_EN-IT

Quality_EN-DE

test_Quality_EN-IT

Active:

☐

Create

ALL USERS

admin - Administrator (tasks: all)

ben (tasks: all)

cgir (tasks: all)

nicv (tasks: 6)

fedr (tasks: 6)

jore (ks: 6)

yuw (asks: 1)

hor (asks: 1)

har (ks: 1)

for (tasks: 1)

ant (tasks: 1)

ras (DINOV (tasks: 1)

deu (tasks: 1)

hou (ks: 1)

jarn (tasks: 1)

war (tasks: 1)

alep (UBNYY (tasks: 1)

mar (F (tasks: 1)

fed (o (tasks: 6)

fara (asks: 6)

vha (tasks: 7)

tina (s: 2)

bae (ks: 2)

tdau (2)

cbo (s: 9)

fostr (8)

a2d (s: 8)

a2b (ks: 6)

frat (asks: 6)

PRIVACY-PROTECTED DATA

4. ANNOTATION Tab

This tab contains the progress monitoring and export functions.

Progress monitoring. For each user, a report containing real-time information about the progress of the annotation is displayed, namely: the total number of annotation performed (“N.ann”), the total number of evaluation items annotated and confirmed (i.e. the annotator has clicked on the “Done” button for the given evaluation item) (N. of evaluated items), a summary of annotations divided by task (“Annotated tasks”). Moreover, the IDs of the more recent annotated items are listed (“Last annotations”), and by clicking on them it is possible to monitor user activity through the visualization of the remote client interface in read-only mode. This feature is particularly useful as it addresses the typical problems related to training and supervision of remote annotators.

Copy of annotations. The MT-Equal tool allows the user not only to annotate data from scratch, but also to revise other users’ annotations. To activate the revision mode, the annotations of another user can be copied into a given user account by selecting them in the menu “Duplicate annotation from task” which appears in the “Annotated tasks” column.

Export of annotated data. Data can be exported both in *xml* and *csv* format. By clicking the Export buttons at the top-left corner of the interface, the task list is

presented, and the manager can export the annotated data for all the tasks selected by the manager. By clicking the Export buttons within the “Annotated tasks” column, it is possible to export all the annotations carried out by each single user.

Furthermore, data annotated by one user can be copied into another user account. For each user, the “Annotated tasks” column contains the “Duplicate annotation from task” menu. By selecting one or more annotations available in the menu, they are copied into the given user account. This function allows the user to revise an existing annotation.

TASK	DATA	USER	ANNOTATION	STATS & AGREEMENT
Export all annotations: CSV XML				
User	N.ann	N.pages	Last annotations	DB Check
#1 admin: Administrator	0	0		✓
#2 bentivo: Luisa Bentivogli	2109	4	<p>67318, task: 17 - 2 weeks ago</p> <p>74944, task: 32 - 2 weeks ago</p> <p>30687, task: 7 - 2 weeks ago</p> <p>30686, task: 7 - 2 weeks ago</p> <p>9408, task: 5 - 3 months ago</p> <p>9424, task: 5 - 3 months ago</p> <p>9439, task: 5 - 3 months ago</p> <p>9456, task: 5 - 3 months ago</p> <p>9473, task: 5 - 3 months ago</p> <p>9488, task: 5 - 3 months ago</p> <p>9504, task: 5 - 3 months ago</p> <p>9521, task: 5 - 3 months ago</p> <p>9537, task: 5 - 3 months ago</p> <p>9553, task: 5 - 3 months ago</p> <p>9569, task: 5 - 3 months ago</p> <p>9585, task: 5 - 3 months ago</p> <p>9601, task: 5 - 3 months ago</p>	<p>Check these incompleted annotations or errors: 74993 (2014-05-20 12:26:34),</p>
Annotated tasks				
Duplicate annotation from task: F_Quality_EN-AR (id: 1) done by cgirardi				
<ul style="list-style-type: none"> 3. F_Quality_EN-ZH : 3 5. F_Errors_EN-RU : 2088 7. test_Quality_EN-IT : 6 8. test_Errors_EN-IT : 5 17. Errors1_EN-IT : 5 32. test_WordAlign_EN-IT: 2 				
Export: CSV XML				
Duplicate annotation from task: F_Quality_EN-AR (id: 1) done by cgirardi				
<ul style="list-style-type: none"> F_Quality_EN-AR (id: 1) done by cgirardi F_Quality_EN-AR (id: 1) done by doubad F_Quality_EN-AR (id: 1) done by houqis F_Quality_EN-AR (id: 1) done by jamyam F_Quality_EN-AR (id: 1) done by federico F_Quality_EN-RU (id: 2) done by cgirardi 				

5. STATS & AGREEMENT Tab

For each chosen task, this tab contains a table with annotation statistics: the total number of annotations for each annotation type and for each evaluated MT output is given. By clicking on [#], the number of annotations for each user is presented

Furthermore, a sentence-based visualization of annotators’ agreement is presented. For each evaluation item, the MT outputs are presented and above each sentence a raw for each user’s annotation is displayed. In this way it is possible to see where annotators are in agreement and where they are not.

Choose a task: Errors10_EN-IT

This task has been annotated by 2 users

Annotation type	Systems	
	sys1	sys2
No errors	81[#]	286[#]
Too many errors	-	-
Reordering errors	61[#]	-
Lexicon errors	178[#]	3[#]
Missing word(s)	2[#]	84[#]
Morphology errors	219[#]	-
Casing and punctuation errors	8[#]	-
Superfluous	10[#]	-
	559	373

Show sentence annotations: prev | next

75562

Dobbiamo imparare a queste persone a venire e parlare con voi .

Dovete imparare come far venire queste persone a parlare con voi .